

# Beyond the RNA-dependent function of LncRNA genes

Tamer Ali<sup>1,2\*</sup>, Phillip Grote<sup>1\*</sup>

<sup>1</sup>Institute of Cardiovascular Regeneration, Centre for Molecular Medicine, Goethe University, Frankfurt am Main, Germany; <sup>2</sup>Faculty of Science, Benha University, Benha, Egypt

**Abstract** While long non-coding RNA (lncRNA) genes have attracted a lot of attention in the last decade, the focus regarding their mechanisms of action has been primarily on the RNA product of these genes. Recent work on several lncRNAs genes demonstrates that not only is the produced RNA species important, but also that transcription of the lncRNA locus alone can have regulatory functions. Like the functions of lncRNA transcripts, the mechanisms that underlie these genome-based functions are varied. Here we highlight some of these examples and provide an outlook on how the functional mechanisms of a lncRNA gene can be determined.

## LncRNA genes in the genome

The complex genome of eukaryotes is pervasively transcribed and efforts to comprehensively define all transcripts have led to the idea that about half of the genome can be transcribed into RNA in an individual cell (*Djebali et al., 2012*). The units that produce RNAs – the genes – can roughly be categorized into the two main biotypes: protein-coding genes (PCGs) and non-protein-coding genes (NCGs). The largest and most coherent category is the PCG, which encodes RNAs that serve as the template for all the peptides and proteins in the cell. The NCG category is a highly heterogeneous collection and can be sub-grouped into small ncRNA (non-coding RNA) and long ncRNA (lncRNA) genes, where the term long refers to the arbitrary length of 200 nucleotides or longer. In particular, the lncRNA genes have attracted a lot of attention in recent years due to their wide range of action and mostly unexplored functions. While their number was overestimated after their initial discovery, similar to the overestimation of the number of PCGs at the beginning of the human genome project (*Lander et al., 2001*), current and careful curation projects, such as the GENCODE and FANTOM projects, list 17,957 and 27,919 lncRNA genes, respectively (*Figure 1A*), in their most recent data releases of the human genome (*Frankish et al., 2019; Hon et al., 2017*). Hence, the number of lncRNA genes are in the same range, or even a bit higher, than the number of PCGs (19,954). In the future, this currently very heterogeneous class of NCGs may be sub-categorized further into more specific biotypes.

Currently, three major functional principles can be assigned to lncRNA loci (*Figure 1B*): (1) either the RNA is the functional biomolecule and interacts with other components in the cell, for example DNA, proteins or RNAs, (2) a gene regulatory element is embedded in the transcription body of a lncRNA gene and the activity of the lncRNA gene directs the activity of the regulatory element or (3) the process of transcription influences genome and thereby gene activity. A lncRNA locus can have one of these functions or a mixture of them (*Yin et al., 2015*). In this review we will focus on the latter two functional lncRNA properties, in which the RNA is, at least partially dispensable for the lncRNA gene function.

**\*For correspondence:**

ali@med.uni-frankfurt.de (TA);  
grote@med.uni-frankfurt.de (PG)

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 11

**Received:** 02 July 2020  
**Accepted:** 12 October 2020  
**Published:** 23 October 2020

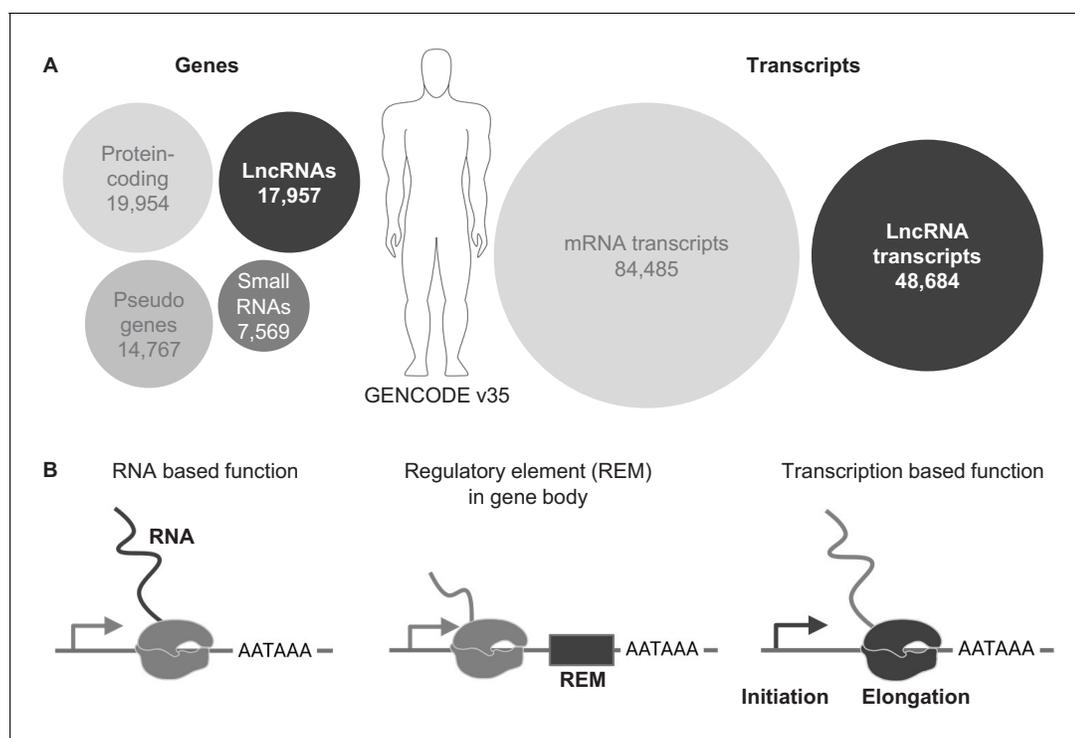
**Reviewing editor:** Didier YR Stainier, Max Planck Institute for Heart and Lung Research, Germany

© Copyright Ali and Grote. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## The transcription of genes

The generation of RNA using the genome as a template, or the process of transcription, depends on certain functional genomic elements (**Figure 2**). The core element of a gene that initiates the production of an RNA is the promoter. A GC-rich element that is accessible (open chromatin) will attract the polymerase machinery and general transcription factors (TFs). This minimal core element serves as a core promoter and can be sufficient to initiate transcription (**Deaton and Bird, 2011**). Transcription of RNA starts at the transcriptional start site (TSS), which is located within the core promoter. Like PCGs, most lncRNAs are transcribed by POL II (RNA polymerase 2, a multiprotein complex), but are more tissue-specific compared to PCGs (for review see **Ransohoff et al., 2018**). Both biotypes (PCGs and lncRNAs) have conserved core promoter sequences with fewer overlapping TF binding motifs in lncRNA promoters, resulting in an overall lower expression level compared to PCGs (**Figure 2; Mattioli et al., 2019**). Thus, the architecture of the core promoter is the first player that defines the degree of lncRNA expression (**Batut and Gingeras, 2017; Mattioli et al., 2019**). The second important element that influences the transcription of genes are enhancers, which are *cis*-regulatory elements that can either have a positive or a negative (which are then often referred to as repressors) impact on their target genes. Consequently, enhancers are genomic regions that encode binding sites for sequence-specific activator or repressor TFs. These elements often confer specificity in spatiotemporal expression. Many lncRNAs can also be generated from such enhancer elements, which contributes to their overall more tissue-specific expression when compared to PCGs (**Mattioli et al., 2019**).

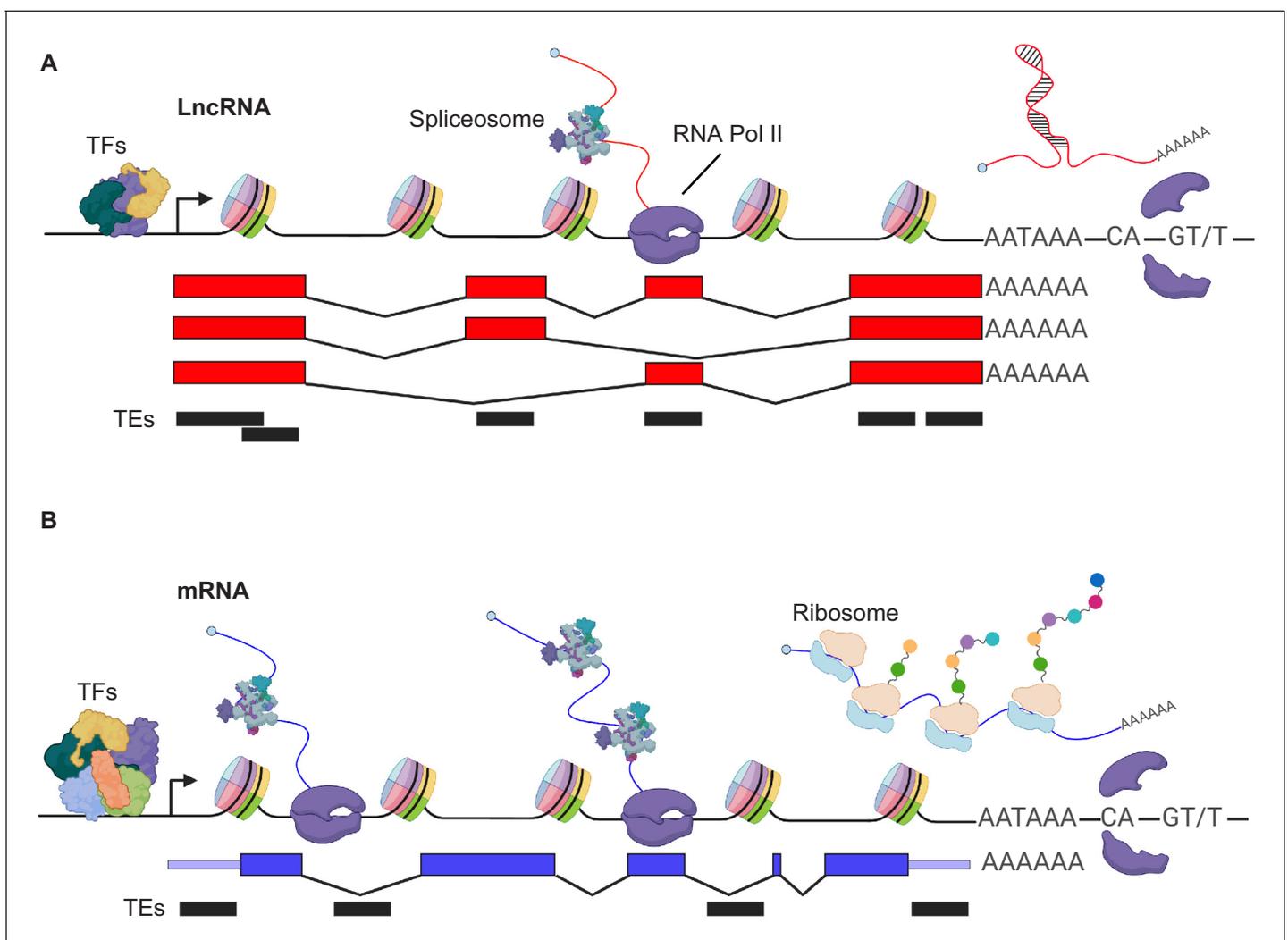
The core promoter initiates transcription and thereby the generation of an RNA that may or may not be further diversified by splicing (**Figure 2**). This depends on whether splice sites are present between the promoter and the transcription termination element, the polyadenylation signal (pA). The mechanism of PCG and lncRNA splicing is similar, although the splicing efficiency of lncRNAs is lower than PCGs, likely due to the loss of proximal RNA POL II phosphorylation over 5' splice sites



**Figure 1.** LncRNA genes in the genome. (A) Overview of genes and transcript numbers in the human genome (GENCODE v35). Circle area represents relative quantities. (B) Schematics of three possible functional properties of lncRNA loci.

(Krcbnáková *et al.*, 2019). In addition, lncRNAs show signs of co-transcriptional cleavage and premature termination with Thr4p PolII enriched over the entire lncRNA body (Schlackow *et al.*, 2017). At some point the transcriptional machinery will run into a termination signal, a DNA sequence element consisting of AATAAA and downstream GU (or U)-rich motifs (Eaton *et al.*, 2020). These elements are ubiquitously present in the genome. In humans, one can find 569,005 elements that meet the criterion of a pA signal (301,001 in mouse and 20,931 in *C. elegans*) (Herrmann *et al.*, 2020). Moreover, this high number likely ensures successful termination of transcription (Eaton and West, 2020).

Another class of genetic elements that play an important role for gene and genome activity are transposable elements (TEs) (for review see Chuong *et al.*, 2017). These mobile genomic elements make up more than 44% of the human genome (Lander *et al.*, 2001) and attracted attention as important regulators of gene and genome activity (Bourque *et al.*, 2018). In this respect, TEs are an important component of lncRNA biology as well (Figure 2A). Approximately, 75% of lncRNA transcripts contain sequence elements from TEs (Kapusta *et al.*, 2013) and some of them represent important sequence elements to direct lncRNA localization (Lubelsky and Ulitsky, 2018). In addition, 25% of TEs are found to overlap with TSS and pA signals of lncRNA genes (Kapusta *et al.*, 2013). Hence, they are an important driving force of lncRNA expression. One recent example is the



**Figure 2.** Distinguishing features of transcript generation of PCGs and lncRNAs (A) lncRNA and (B) mRNAs: lncRNA genes are lowly expressed as fewer transcription factors (TFs) bind the promoter. In addition, lncRNA TSS, exon and/or pA site more often associate with transposable elements (TEs), while TEs contribute mostly to UTRs and/or introns of mRNAs. In addition, mRNAs are more efficiently spliced.

primate-specific lncRNA *XACT* (Table 1), which has been shown to protect the active X chromosome from being silenced (antagonizing *XIST* lncRNA effect) and whose sequence contains elements derived from a TE (Casanova et al., 2019). Interestingly, *XACT* lncRNA is also regulated by a TE-derived enhancer element that harbors pioneer pluripotency factor binding sites. This exemplifies that TEs containing embedded TF motifs can direct tissue-specific expression when they insert next to a promoter element. Several other TE-derived lncRNAs are described elsewhere (Kapusta et al., 2013).

In summary, the genome stores the information required to generate the RNAs that are necessary for a cell's proper function, whether the RNA is protein-coding or not. An elaborate machinery is established that controls the specific activation of genes and whole genomic regions via positive or negative mechanisms. These regulatory mechanisms require energy investment from the cell. It is conceivable that sometimes it can be 'cheaper' for a cell to let spurious transcription of non-harmful transcripts occur, might they be coding or non-coding, than to invest energy in silencing all of these transcriptionally active sites.

## Layers of gene regulation

The expression of genes and whole genomic regions is controlled by several layers of regulation. In addition to the genomic elements described above, DNA is packed with histone proteins into chromatin. These protein components can be modified to act as signaling centers for the transcription machinery (for review see Talbert et al., 2019). In addition, the proteins of the nucleus also regulate the 3D arrangement of genomic DNA in such a way that functionally connected elements of gene regulation come together. In short, each chromosome is composed of sub-megabase units known as topologically associated domains (TADs), the structural and functional unit of the chromosome (for review see Szabo et al., 2019). Such genome arrangements can allow for promoter-enhancer contacts and organize functionally dependent regulatory elements together (Hnisz et al., 2017). The major factors that regulate this organization are CTCF (CCCTC-binding TF) and the cohesin complex (Ali et al., 2016; Rao et al., 2017). CTCF binding frequently co-localizes and interacts with the cohesin complex at TAD borders (Li et al., 2020). Indeed, elimination of cohesin dissolves all chromatin TADs even in the presence of CTCF (Rao et al., 2017). Interestingly, disruption of the TADs either by removal of CTCF or cohesin results in unexpected mild effects on gene expression (Nora et al.,

**Table 1.** Selection of lncRNA genes with RNA independent function.

LncRNA	Relative location of respective TSSs target gene	Literature	Mode of action
Regulatory element located within the transcription unit			
<i>Haunt (Halr1)</i>	40 kb downstream of <i>HOXA</i>	Yin et al., 2015	Activation of <i>HOXA</i>
<i>Lockd</i>	4 kb downstream of <i>Cdkn1b</i>	Paralkar et al., 2016	Positive regulation of <i>Cdkn1b</i> via loop formation
<i>Meteor</i>	80 kb upstream of <i>Eomes</i>	Alexanian et al., 2017	Positive licensing of <i>Eomes</i> expression
<i>ThymoD</i>	844 kb downstream of <i>Bcl11b</i>	Isoda et al., 2017	DNA methylation, CTCF-binding
<i>Pcdha-as</i>	<i>Pcdha</i>	Canzio et al., 2019	DNA methylation, CTCF-binding
<i>GAL10-ncRNA</i>	<i>GAL10</i> antisense transcript	Houseley et al., 2008	<i>GAL10</i> promoter acetylation
<i>AIRN</i>	28 kb Antisense to <i>Igfr2</i>	Latos et al., 2012	Promoter methylation
<i>Upperhand (Hand2os1)</i>	0,1 kb upstream of <i>Hand2</i>	Anderson et al., 2016; Han et al., 2019	Promotes enhancer accessibility for <i>Hand2</i> activation
Activity exerted by transcription initiation or elongation			
<i>Ftx</i>	140 kb upstream of <i>Xist</i>	Furlan et al., 2018	<i>Xist</i> activation independent of <i>Ftx</i> RNA
<i>Chaserr</i>	16 kb upstream of <i>Chd2</i>	Rom et al., 2019	Negative regulation of <i>Chd2</i>
<i>PVT1</i>	52 kb downstream of <i>Myc</i>	Cho et al., 2018	Enhancer boundary element
<i>Handsdwn (Handlr)</i>	11 kb downstream of <i>Hand2</i>	George et al., 2019; Ritter et al., 2019	Transcriptional elongation-based enhancer shielding

2017; Rao et al., 2017). While it has been accepted that gene expression and 3D genome folding are correlated, their functional relevance is still to be elucidated (Ibrahim and Mundlos, 2020).

All of these enhancers and genome organizing regions must be functionally regulated to accurately control gene and genome activity. As many such regulatory sites are associated with lncRNAs, these lncRNA loci might be important functional support elements. The process of transcription can assist in reorganizing chromatin marks (van Steensel and Furlong, 2019), allowing regions to be accessible for other factors or prevent others by diverting/directing the transcription machinery to nearby genes.

### Current annotations in the database are a work-in-progress

Current annotations of genomic databases categorize genes according to various criteria. One that appears, on the surface, to be very simple is the separation of protein-coding genes (PCGs) and non-protein-coding genes (NCGs). It was already found some time ago that RNAs originating from NCGs do actually associate with ribosomes, the machinery that translates mRNAs into proteins (Ingolia et al., 2011; van Heesch et al., 2014). This association is not surprising, as the ribosomes function is to bind RNAs in the cytosol and attempt to translate it into a peptide or protein. However, just because an RNA is bound to a ribosome does not mean it is translated and even if translated, the pure presence of a peptide does not prove a function of this peptide. In more recent in-depth studies, it was found that some lncRNAs do produce peptides and that some of these peptides are even functional (Chen et al., 2020; Ji et al., 2015; van Heesch et al., 2019), including within 5' and 3' untranslated regions (UTR) of mRNAs. Hence, until databases are updated with suitable information that incorporates the presence of peptides derived from expressed RNAs, a peptide coding probability always must be taken into consideration when studying lncRNA function. Equally important, many PCG or NCGs have a high number of splice variants, some of which might encode a peptide and others not.

The revolution of high-throughput sequencing of fragmented cDNA libraries revealed the complexity of expression from the genome. Enrichment of lowly expressed transcripts and subsequent sequence analysis identified an even more complex pattern of splice variants (Mercer et al., 2012). However, these analyses relied on the sequencing of fragmented cDNA libraries and subsequent reconstruction of the transcriptome to a reference genome. The most recent generation of long read sequencers, such as the PacBio or the Nanopore systems, allows the direct analysis of RNAs and eliminates the intermediate step of a fragmented cDNA library. Capturing lncRNA genes specifically and resequencing by Long-read platform (known as Capture Long Sequence or CLS) determined the full variety of splice variants of the mammalian transcriptome (Lagarde et al., 2017). The advantage of this technology is the capability to precisely determine 5' and 3' ends and, ideally, all splice variants of a transcript. For example, the estimated mean number of exons per lncRNA using CLS was 4.27 compared to 3.59 measured by short-reads RNA-seq method (Lagarde et al., 2017). While this approach doesn't eliminate the necessity to carefully determine the splice variants from a lncRNA locus entirely, it does provide a very good starting point for detailed analysis. In particular, when CLS data are not available for your locus-of-interest or your tissue-of-interest, one should determine the full transcript length, splice variants and regulatory elements of the lncRNA-of-interest. Only then can a successful strategy to study the lncRNA be initiated.

### Gene regulation by lncRNA genes – regulatory elements within the transcription unit

Surveying the chromatin and DNA modification landscape led to the annotation of potential regulatory regions across the genome and sometimes even for specific tissues and cell types. Regulatory elements, whether they are promoters or other regulatory elements, can be found within or far away from the transcription unit of a gene. The occurrence of such a regulatory element within a transcription unit, for example of a lncRNA gene, can indicate that the function of this element might be affected by its activity.

One interesting lncRNA gene example that reflects the duality of lncRNA genes with respect to their RNA-based mechanism on one side, and an enhancer element on the other side, is *Haunt*. While the RNA of *Haunt* is thought to be required for negative regulation of *HoxA*, the *Haunt* locus contains regulatory elements to activate the *HoxA* locus during *in vitro* differentiation of pluripotent

stem cells (Yin *et al.*, 2015). While it is shown that these enhancers can interact with *HoxA* directly, the elements are not further defined nor how their function might depend on *Haunt* transcriptional activity.

A similar early example of a lncRNA locus that contains a regulatory element within its transcription unit is the *Lockd* lncRNA locus, which regulates its *cis* gene *Cdkn1b*. The deletion of the entire locus of *Lockd*, including TSS upstream elements, leads to a reduction of *Cdkn1b* expression (Paralkar *et al.*, 2016). While the 5' genomic region of *Lockd* interacts genomically with the promoter of *Cdkn1b*, this interaction is not altered if the transcription of *Lockd* is depleted by a pA signal inserted into the first exon of *Lockd*. Thus, the genomic locus itself is important as a regulatory element rather than its transcriptional activity.

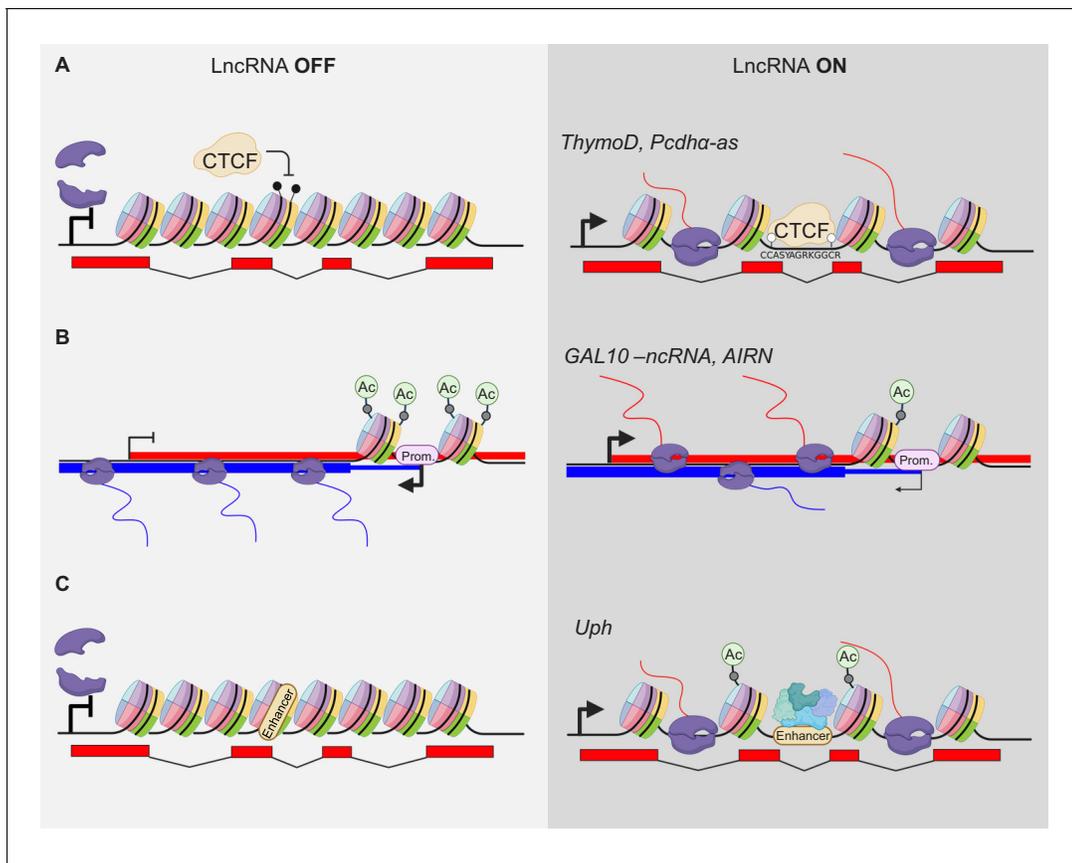
Even if a specific regulatory element cannot be defined, careful analysis and genetic dissection of a lncRNA can point toward such a regulatory principle. The TSS of the *Meteor* lncRNA locus is important to license its *cis*-located gene *Eomes* for activation in the mesendoderm (Alexanian *et al.*, 2017). The lack of *Meteor* expression by TSS deletion causes the loss of *Eomes* activation during mesendoderm differentiation of mouse ESCs. Decreasing levels of *Meteor* RNA during this process did not alter expression of downstream genes, arguing against an RNA-based function of *Meteor*. Interestingly, endogenous activation of *Meteor* is not only licensing *Eomes* gene activation, but other cardiac mesodermal genes as well. Moreover, transcriptional inhibition of *Meteor* using a polyadenylation element insertion downstream of the *Meteor* TSS does not cause the *Eomes* gene to be silenced during mesendoderm differentiation (Alexanian *et al.*, 2017; Engreitz *et al.*, 2016). This finding argues against a transcription-based mechanism of *Meteor* and suggests that the genomic locus *Meteor* harbors important regulatory elements to render the *cis*-located *Eomes* gene activatable during differentiation.

An excellent example of a lncRNA with a defined regulatory element within transcription unit is the *ThymoD* lncRNA locus. Its transcription prevents methylation of a CTCF-binding site located within its transcriptional unit (Isoda *et al.*, 2017; Figure 3A). The binding of CTCF allows looping of the *Bcl11b* transcription unit in the same domain as activating regions of *Bcl11b*. This activation is lost when the transcription of *ThymoD* is blocked by insertion of a pA signal after exon two and before the CTCF-binding site and, consequentially, the CTCF-binding site is methylated (Figure 3A). Therefore, the transcriptional activity has an indirect, structural effect on the regulation of *Bcl11b* while the *ThymoD* RNA is dispensable.

A more complex situation of several antisense transcripts regulating their *cis* gene is the Protocadherin alpha (*Pcdh $\alpha$* ) cluster. The variable, stochastic expression from several Protocadherin clusters provide cell-surface proteins for cellular identity recognition in the neuronal system to allow dendrites and axons to distinguish from self and other neurons. This stochastic expression is partly regulated by a distal enhancer region. The cluster of *Pcdh $\alpha$*  produces three distinct variants from three alternative TSSs to achieve stochastic expression of splice variants from this cluster. The first exon of each of these variants contains an antisense lncRNA transcript (*Pcdh $\alpha$ -as*) (Canzio *et al.*, 2019). The expression of the lncRNAs precedes the expression of the PCGs and positively regulates the most nearby PCG expression. Mechanistically, the *Pcdh $\alpha$*  lncRNAs act similar to the *ThymoD* lncRNA (above) (Figure 3A). Expression of the *Pcdh $\alpha$ -as* variants leads to the demethylation of a CTCF-binding site in the region upstream of the *Pcdh $\alpha$*  PCG, thereby allowing for a stable loop formation with the distal enhancer region and a positive effect on the PCGs expression.

There are also examples of lncRNA genes that reside within a different transcriptional entity from *cis* target genes. Here, it is even more conceivable that their activity has an impact on the gene they are embedded in. One of the first examples was a ncRNA within the *GAL10* gene cluster in yeast *Saccharomyces cerevisiae*. Under 0% galactose, the TF Reb1 binds to the promoter region of *GAL10-ncRNA* antisense to *GAL10* and fully activates its expression (Houseley *et al.*, 2008). The transcriptional unit of *GAL10-ncRNA* overlaps with the TSS *GAL10* and *GAL1*, leading to inhibition of the *GAL10* and *GAL1* gene by promoting high levels of H3K36me3 methylation and hypoacetylation at the *GAL10* and *GAL1* promoters. Addition of galactose to the growth medium blocks *GAL10-ncRNA* expression and hyperacetylation of the *GAL10* and *GAL1* promoters, leading to expression of genes that encode galactose fermenting proteins (Figure 3B).

A similar principle was shown in higher eukaryotes at the *AIRN* (antisense *Igf2r* RNA non-coding) locus. The TSS of the lncRNA *AIRN* is located in the second intron of the *Igf2r* PCG and *AIRN* is transcribed antisense to *Igf2r*. Transcription of *AIRN* negatively regulates *Igf2r* (Santoro *et al.*, 2013).



**Figure 3.** Modulation of gene expression by lncRNA transcription. (A) Transcriptional activity modulates DNA methylation and thereby alters occupation of DNA binding factors within the gene body, for example CTCF. The POL2 complex is indicated in violet. Black drumsticks indicate methylated CpGs, white drumstick non-methylated CpGs. (B) LncRNA expression alters promoter (Prom.) activity by modifying e.g. acetylation of histones at TSS sites. (C) Transcription elongation can activate poised enhancers within their gene body (only acetylation shown).

When transcription of AIRN is blocked by a polyA insertion before the promoter of *Igf2r*, this negative regulation is abolished (Figure 3B). However, if the same pA is inserted after the promoter of *Igf2r*, this negative regulatory effect on *Igf2r* is not observed (Latos et al., 2012). These findings support the hypothesis that the transcription of AIRN, and not the RNA product itself, is important for the transcriptional regulation of the *Igf2r*.

A lncRNA gene transcription that influences an enhancer is *Upperhand*, which is divergently expressed from the *Hand2* protein-coding gene (Anderson et al., 2016). Loss of *Upperhand* transcription leads to a loss of histone acetylation upstream of *Hand2*, including at the cardiac enhancer. As a result, binding of GATA4 to its previously defined enhancer (McFadden et al., 2000) is reduced, and *Hand2* expression in the heart is reduced as well. Hence, the *Upperhand* loss-of-function phenotype is similar to cardiac loss of *Hand2* (Figure 3A). Additional mutants of *Upperhand* draw a more complicated picture of the role of *Upperhand* in activating *Hand2*. A complete deletion of the *Upperhand* transcription unit that encompasses all known regulatory regions of the *Hand2* gene as well, causes loss of *Hand2* 5'UTR expression (Han et al., 2019). These findings assert the presence of important *Hand2* activating genetic elements directly upstream of its TSS, independently of any RNA originating from this region. However, a promoter deletion of *Upperhand* causes a loss of its RNA while leaving all other elements in that region intact, but no effect on *Hand2* expression was observed in this case. Furthermore, a deletion of the last two exons from *Upperhand* has a slight effect on *Hand2* expression. There might be so far uncharacterized enhancer elements in the genomic region of these two exons and their deletion may influence *Hand2* expression. In addition, although the *Upperhand* RNA is suggested to be not required for its *in vivo* function, the RNA

generates peptides that might be functional (*van Heesch et al., 2019*). These somehow conflicting results underline the complexity of regulation of the *Hand2* gene.

These examples highlight the importance of taking a careful look at the whole lncRNA locus that produces an RNA. The occurrence of an annotated regulatory element or the occupation of a genome regulating factor such as CTCF within the transcription unit can be an important indication to look for a genomic function of a lncRNA.

### Gene regulation by lncRNA genes – the act of transcription is functional

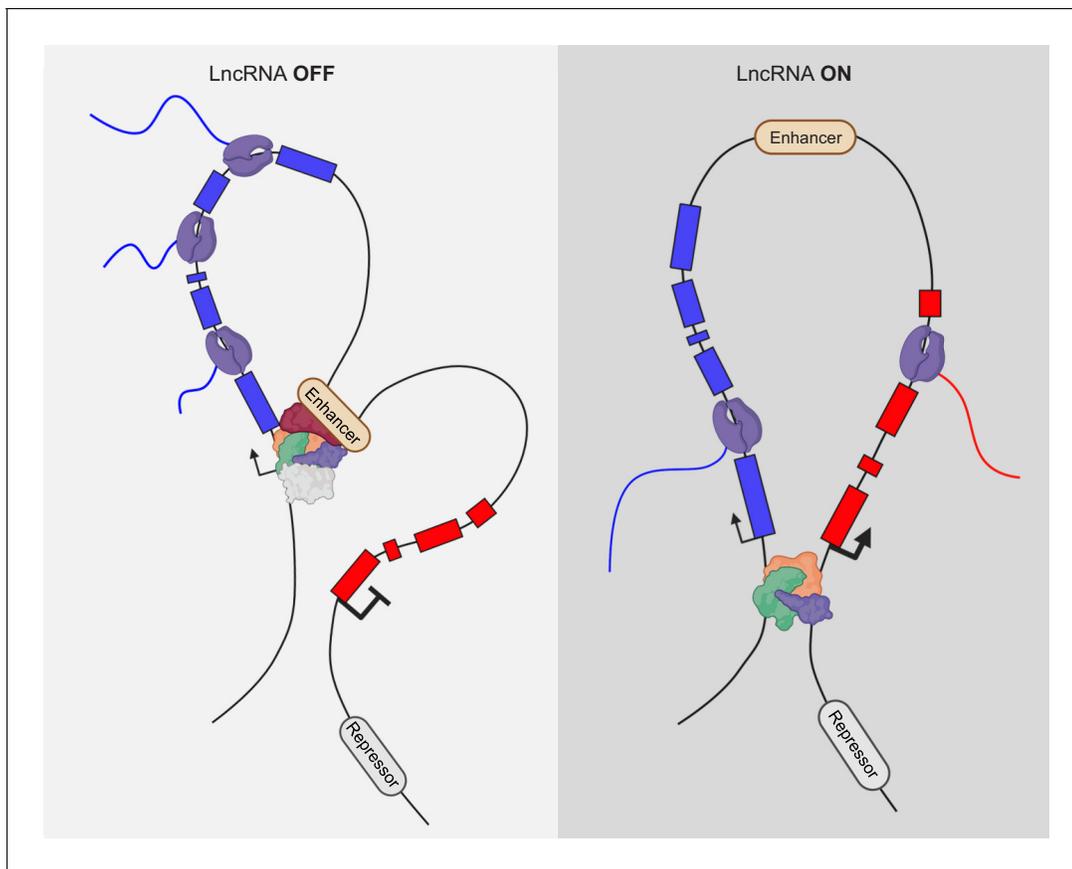
The absence of a regulatory element within the transcription unit could be due to incomplete annotation or a yet unknown factor which binds there, or the act of transcription initiation or transcriptional elongation is important for the function of a lncRNA locus.

One example of such a regulation principle comes from work on the *XIST* lncRNA, which is one of the original lncRNAs that has been extensively studied (*Brockdorff et al., 1992*). While *XIST* acts via the produced RNA (*Brannan et al., 1990; Brown et al., 1992*), the regulation of *XIST*, at least in part, does not. The *XIST* lncRNA locus is flanked by many lncRNAs, and one of them is the *Ftx* locus found 140 kb upstream of *Xist* (*Chureau et al., 2002*). It was initially proposed that the *Ftx* RNA functions to regulate *XIST* (*Chureau et al., 2011*). However, detailed analysis uncovered that the transcription of *Ftx*, and not the produced RNA, is important to regulate *Xist* (*Furlan et al., 2018*). Knockdown of *Ftx* RNA does not cause a loss of *Xist* expression, but deletion of the promoter of *Ftx*, and the consequential loss of *Ftx* transcription, causes the loss of *Xist* expression. CRISPRi of *Ftx* similarly causes loss of *Xist* expression, suggesting that transcription of *Ftx* is the positive regulator of *Xist* expression. One possibility is that 3D genome architecture can be changed due to the transcriptional activity of a genomic locus (*Figure 4*). Strikingly, the promoter of *Xist* and *Ftx* are flanked by CTCF-occupied sites. However, deleting the CTCF-binding sites alone at the *Ftx* promoter has no effect on the expression level of *Xist*, arguing that genome folding induced by *Ftx* activity does not involve CTCF-binding.

Another good example is the *Chaserr* lncRNA locus, which lies 16 kb upstream of the *Chd2* protein-coding gene (*Rom et al., 2019*). Although, knock-down of *Chaserr* RNA does cause a slight increase in *Chd2* expression, additional lines of evidence infer that the transcription of the lncRNA gene is likely the most important function of *Chaserr* in regulating *Chd2* (*Figure 4*). In addition, the promoter of *Chaserr* interacts with the *Chd2* promoter in chromosome conformation capture analysis. Upon deletion of the *Chaserr* promoter region, the *Chd2* promoter increasingly interacts with other enhancer elements upstream. In contrast, if the gene body of *Chaserr* is deleted, leaving the promoter intact, these changes in enhancer/*Chd2*-promoter contacts are not observed. A plausible explanation is that the transcription initiation activity rather than the transcription elongation is important for regulation of *Chd2* by *Chaserr*.

Similarly, transcription initiation is important for the *PVT1* lncRNA locus. The *Pvt-1* lncRNA was originally discovered as a genomic translocation that causes the activation of the *Myc* oncogene (*Adams and Cory, 1985*). Initially, it was suggested that miRNAs embedded in the lncRNA transcript of *PVT1* are important for regulation of target genes (*Wang et al., 2019*). It turns out that *PVT1* transcription has an RNA-independent function as well. The *PVT1* locus encodes several transcripts with alternative start sites. The activity of its major TSS serves as a boundary element to shield the *MYC* promoter from over-activation by an enhancer located within the transcriptional unit of *PVT1* (*Cho et al., 2018; Figure 4*). The transcriptional activity is important for this shielding capacity, but not the elongation of the transcription (*Figure 4*). This does not mean that the miRNAs produced by *PVT1* do not serve a function, but it seems the major activity of the *PVT1* lncRNA, and its effect on *MYC* is conveyed by the transcriptional activation of *PVT1*.

In addition to the *Upperhand* lncRNA upstream of *Hand2* (see above), there are *Hand2*-regulating lncRNA loci downstream of *Hand2*. We initially characterized this locus and termed it *Handsdwn*, due its location downstream of *Hand2*. The *Handsdwn* locus is expressed in the same tissues as *Hand2* but is most significantly expressed in the developing heart. We have shown that transcription of *Handsdwn* is important to negatively regulate the expression of *Hand2* (*Figure 4*). Moreover, the HAND2 TF binds two distinct sites around the TSS of *Handsdwn* in the developing E9.5 heart (*Laurent et al., 2017*). This suggests that HAND2 activates its own suppressor region in a negative feedback loop to control its dosage. However, deletion of the TSS region of *Handsdwn*, including only one of the HAND2 occupied sites, does not result in the expected upregulation of *Hand2*



**Figure 4.** Alteration of genome interactions by lncRNA activity. DNA:DNA contacts can change upon transcriptional activity of nearby, *cis* located lncRNA genes.

(George *et al.*, 2019). Multiple, potential TSS regions are present in at the 5' region of *Hand2* and the deletion of one or the major TSS can lead to the appearance of alternate transcripts (Lavalou *et al.*, 2019). Therefore, it is plausible that the second HAND2 occupied site may be sufficient to instruct the transcription of an alternate *Hand2* transcript. Hence, as long as transcriptional activity is present in the *Hand2* region, *Hand2* can be negatively regulated and its expression level adjusted. The dosage of *Hand2* is particularly important as loss of one copy of the *Hand2* gene, as well as the gain of an additional copy of the *Hand2* gene, causes malformations during development (Tamura *et al.*, 2014). In addition to these lncRNA loci flanking the *Hand2* gene, additional putative enhancers are predicted up- and downstream of *Hand2*, underlining the complex regulome of this important gene in development.

While functions of lncRNAs on the transcript level are becoming increasingly understood, elucidating mechanisms of how such loci, whose function is based on the transcriptional level, exhibit their effect (Table 1) is still in its infancy. While this list is not saturated, the number of lncRNAs that at least partially act by such a mechanism will increase in the future. One very promising model of how they may act are functional microdomains. In such a scenario, these microdomains promote the co-operativity between interacting components such as TFs, co-factors, chromatin regulators, RNA polymerase II, and non-coding RNA, thereby governing basic processes of gene regulation. Such microdomains are favorably formed by super-enhancers that also often generate an RNA, but function on the transcriptional level. Hence, transcriptional activity itself can influence chromatin accessibility, DNA methylation, histone modification, and higher order chromatin structure.

## Outlook

A core question for the near future is to define which of the lncRNA loci are functional on the transcript (RNA) level or on the transcriptional (genome) level and which loci may function on both levels.

The widespread use of the CRISPR toolbox does allow for the generation of targeted genomic modifications to dissect the mode-of-action of a lncRNA locus. With CRISPR/Cas9 deletions even in the mega-base range being feasible, the deletion of the entire transcription unit will allow one to determine if a lncRNA locus is functional at all (*Barutcu et al., 2018; Kraft et al., 2015*). This crude approach eliminates any transcript coming from the locus, also eliminating the possibility that degradation of any residual transcript does cause any effect by, for example, genetic compensation (*El-Brolosy et al., 2019*). Simply put: if the removal of a complete lncRNA locus does not result in even subtle effects on gene expression, this locus can be marked non-functional, at least in the analyzed biological system. Subsequently, the promoter encompassing the TSS can be removed to eliminate any transcriptional initiation of the transcript. A similar result can be achieved to use the CRISPRi (*Ferreira et al., 2018*) system to shut down the locus without removing any parts of the genome. It has to be kept in mind, that removing a TSS might trigger the emergence of new transcripts from secondary TSS-like sites in the vicinity (*Lavalou et al., 2019*). It is therefore important to evaluate this possibility and verify that no 'novel' transcripts arise. To interfere with transcriptional elongation and also study effects in regulatory elements within a gene body, the transcription can be terminated using a strong transcriptional stop signal. To allow for efficient targeting using the CRISPR toolbox, a short and powerful pA signal is preferred (*Ballarino et al., 2018; Lavalou et al., 2019*). In combination with an endogenous CRISPRa (*Konermann et al., 2015*) system, this meddling with the lncRNA does allow for a detailed assessment of its function. In particular, subsequent removal of parts of a lncRNA locus, for example, whole exons or potential regulatory elements within the transcription unit will allow one to define the functional elements on either the RNA or the locus.

The powerful tool of antisense oligonucleotide (ASO)-assisted knock-down of RNA can now help allow for a detailed assessment of RNA vs transcription-based function. Until now, if a lncRNA transcript was inhibited on the RNA level by antisense oligo methods, siRNA or locked nucleic acid (LNA)-based ASO, it was assumed that the RNA, rather than its transcription was important for the resulting phenotype. Initially, ASOs, which employ the endogenous RNaseH enzyme for target RNA degradation (*Grünweller et al., 2003*), were the method of choice, as they can target nuclear and cytosolic RNA similarly well. However, it turns out one must be a bit more cautious with this assumption. Several recent publications demonstrate that ASOs that target the 5' end of an RNA can do this even on nascent RNA that is in the process of being transcribed (*Eaton et al., 2020; Lai et al., 2020; Lee and Mendell, 2020*). This premature cleavage of RNA leads to the recruitment of XRN2 and employs the torpedo mechanism to evict the POL II transcription machinery prematurely. Hence, a 5' directed ASO mimics the loss of transcriptional elongation and may lead to confusion about potential lncRNA mechanistic function. To validate an RNA-based mechanism by ASOs, it is preferred to target the 3' end of the RNA-of-interest. But, more importantly, this mechanism opens the possibility to target lncRNA (or any other locus) whose mechanism requires transcriptional elongation until the endogenous transcriptional termination site, independent of whether the RNA is functional. Carefully designed experiments can increase our understanding of which lncRNA functions are important for gene regulation, which can be beneficial in studying human disease involving dysregulation by such loci. Furthermore, this extends the repertoire of loci that can be targeted for studying the lncRNA genes and their therapeutic use. Now, not only genes that produce a functional RNA can be therapeutically targeted, but also any gene or regulatory locus that generates an RNA and has a gene or genome regulatory function via transcriptional elongation, *per se*, is amenable to ASO targeting.

## Acknowledgements

We would like to thank the members of the lab for critical comments and discussions. Important input came from Arica Beisaw and Sandra Rogala, and we'd like to thank them for their important contribution to this review. Thanks also to Tracie Pennimpede for critical reading of the manuscript. Parts of the figures were created with [BioRender.com](https://BioRender.com).

## Additional information

### Funding

Funder	Grant reference number	Author
Deutsche Forschungsgemeinschaft	SFB/Transregio TRR267	Tamer Ali
Deutsche Forschungsgemeinschaft	Excellence Cluster Cardio-Pulmonary System (Exc147-2)	Phillip Grote

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author ORCIDs

Phillip Grote  <https://orcid.org/0000-0002-9254-1458>

## References

- Adams JM, Cory S. 1985. Myc oncogene activation in B and T lymphoid tumours. *Proc R Soc London Ser B, Biol Sci* **226**:59–72. DOI: <https://doi.org/10.1098/rspb.1985.0079>
- Alexanian M, Maric D, Jenkinson SP, Mina M, Friedman CE, Ting CC, Micheletti R, Plaisance I, Nemir M, Maison D, Kernen J, Pezzuto I, Villeneuve D, Burdet F, Ibberson M, Leib SL, Palpant NJ, Hernandez N, Ounzain S, Pedrazzini T. 2017. A transcribed enhancer dictates mesendoderm specification in pluripotency. *Nature Communications* **8**:1–18. DOI: <https://doi.org/10.1038/s41467-017-01804-w>, PMID: 29180618
- Ali T, Renkawitz R, Bartkuhn M. 2016. Insulators and domains of gene expression. *Current Opinion in Genetics & Development* **37**:17–26. DOI: <https://doi.org/10.1016/j.gde.2015.11.009>, PMID: 26802288
- Anderson KM, Anderson DM, McAnally JR, Shelton JM, Bassel-Duby R, Olson EN. 2016. Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. *Nature* **539**:433–436. DOI: <https://doi.org/10.1038/nature20128>, PMID: 27783597
- Ballarino M, Cipriano A, Tita R, Santini T, Desideri F, Morlando M, Colantoni A, Carrieri C, Nicoletti C, Musarò A, Carroll DO, Bozzoni I. 2018. Deficiency in the nuclear long noncoding RNA *charme* causes myogenic defects and heart remodeling in mice. *The EMBO Journal* **37**:e99697. DOI: <https://doi.org/10.15252/embj.201899697>, PMID: 30177572
- Barutcu AR, Maass PG, Lewandowski JP, Weiner CL, Rinn JL. 2018. A TAD boundary is preserved upon deletion of the CTCF-rich *firre* locus. *Nature Communications* **9**:1444. DOI: <https://doi.org/10.1038/s41467-018-03614-0>, PMID: 29654311
- Batut PJ, Gingeras TR. 2017. Conserved noncoding transcription and core promoter regulatory code in early *Drosophila* development. *eLife* **6**:e29005. DOI: <https://doi.org/10.7554/eLife.29005>, PMID: 29260710
- Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, Imbeault M, Izsvák Z, Levin HL, Macfarlan TS, Mager DL, Feschotte C. 2018. Ten things you should know about transposable elements. *Genome Biology* **19**:199. DOI: <https://doi.org/10.1186/s13059-018-1577-z>, PMID: 30454069
- Brannan CI, Dees EC, Ingram RS, Tilghman SM. 1990. The product of the H19 gene may function as an RNA. *Molecular and Cellular Biology* **10**:28–36. DOI: <https://doi.org/10.1128/MCB.10.1.28>, PMID: 1688465
- Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, Rastan S. 1992. The product of the mouse *xist* gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **71**:515–526. DOI: [https://doi.org/10.1016/0092-8674\(92\)90519-I](https://doi.org/10.1016/0092-8674(92)90519-I), PMID: 1423610
- Brown CJ, Hendrich BD, Rupert JL, Lafrenière RG, Xing Y, Lawrence J, Willard HF. 1992. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **71**:527–542. DOI: [https://doi.org/10.1016/0092-8674\(92\)90520-M](https://doi.org/10.1016/0092-8674(92)90520-M), PMID: 1423611
- Canzio D, Nwakeze CL, Horta A, Rajkumar SM, Coffey EL, Duffy EE, Duffié R, Monahan K, O’Keeffe S, Simon MD, Lomvardas S, Maniatis T. 2019. Antisense lncRNA transcription mediates DNA demethylation to drive stochastic protocadherin  $\alpha$  promoter choice. *Cell* **177**:639–653. DOI: <https://doi.org/10.1016/j.cell.2019.03.008>, PMID: 30955885
- Casanova M, Moscatelli M, Chauvière LÉ, Huret C, Samson J, Liyakat Ali TM, Rossopoff O, Rougeulle C. 2019. A primate-specific retroviral enhancer wires the XACT lncRNA into the core pluripotency network in humans. *Nature Communications* **10**:5652. DOI: <https://doi.org/10.1038/s41467-019-13551-1>, PMID: 31827084
- Chen J, Brunner AD, Cogan JZ, Nuñez JK, Fields AP, Adamson B, Itzhak DN, Li JY, Mann M, Leonetti MD, Weissman JS. 2020. Pervasive functional translation of noncanonical human open reading frames. *Science* **367**:1140–1146. DOI: <https://doi.org/10.1126/science.aay0262>, PMID: 32139545
- Cho SW, Xu J, Sun R, Mumbach MR, Carter AC, Chen YG, Yost KE, Kim J, He J, Nevins SA, Chin S-F, Caldas C, Liu SJ, Horlbeck MA, Lim DA, Weissman JS, Curtis C, Chang HY. 2018. Promoter of lncRNA gene PVT1 is a

- Tumor-Suppressor DNA boundary element. *Cell* **173**:1398–1412. DOI: <https://doi.org/10.1016/j.cell.2018.03.068>
- Chuong EB**, Elde NC, Feschotte C. 2017. Regulatory activities of transposable elements: from conflicts to benefits. *Nature Reviews Genetics* **18**:71–86. DOI: <https://doi.org/10.1038/nrg.2016.139>, PMID: 27867194
- Chureau C**, Prissette M, Bourdet A, Barbe V, Cattolico L, Jones L, Eggen A, Avner P, Duret L. 2002. Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine. *Genome Research* **12**:894–908. DOI: <https://doi.org/10.1101/gr.152902>, PMID: 12045143
- Chureau C**, Chantalat S, Romito A, Galvani A, Duret L, Avner P, Rougeulle C. 2011. Ftx is a non-coding RNA which affects xist expression and chromatin structure within the X-inactivation center region. *Human Molecular Genetics* **20**:705–718. DOI: <https://doi.org/10.1093/hmg/ddq516>, PMID: 21118898
- Deaton AM**, Bird A. 2011. CpG islands and the regulation of transcription. *Genes & Development* **25**:1010–1022. DOI: <https://doi.org/10.1101/gad.2037511>, PMID: 21576262
- 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, Röder M, Kokocinski F, Abdelhamid RF, Alioto T, et al. 2012. Landscape of transcription in human cells. *Nature* **489**:101–108. DOI: <https://doi.org/10.1038/nature11233>, PMID: 22955620
- Eaton JD**, Francis L, Davidson L, West S. 2020. A unified allosteric/torpedo mechanism for transcriptional termination on human protein-coding genes. *Genes & Development* **34**:132–145. DOI: <https://doi.org/10.1101/gad.332833.119>, PMID: 31805520
- Eaton JD**, West S. 2020. Termination of transcription by RNA polymerase II: boom!. *Trends in Genetics* **36**:664–675. DOI: <https://doi.org/10.1016/j.tig.2020.05.008>, PMID: 32527618
- El-Brolosy MA**, Kontarakis Z, Rossi A, Kuenne C, Günther S, Fukuda N, Kikhi K, Boezio GLM, Takacs CM, Lai SL, Fukuda R, Gerri C, Giraldez AJ, Stainier DYR. 2019. Genetic compensation triggered by mutant mRNA degradation. *Nature* **568**:193–197. DOI: <https://doi.org/10.1038/s41586-019-1064-z>, PMID: 30944477
- Engreitz JM**, Haines JE, Perez EM, Munson G, Chen J, Kane M, McDonel PE, Guttman M, Lander ES. 2016. Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**:452–455. DOI: <https://doi.org/10.1038/nature20149>, PMID: 27783602
- Ferreira R**, Skrekas C, Nielsen J, David F. 2018. Multiplexed CRISPR/Cas9 genome editing and gene regulation using Csy4 in *Saccharomyces cerevisiae*. *ACS Synthetic Biology* **7**:10–15. DOI: <https://doi.org/10.1021/acssynbio.7b00259>, PMID: 29161506
- Frankish A**, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, Mudge JM, Sisu C, Wright J, Armstrong J, Barnes I, Berry A, Bignell A, Carbonell Sala S, Chrast J, Cunningham F, Di Domenico T, Donaldson S, Fiddes IT, García Girón C, et al. 2019. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Research* **47**:D766–D773. DOI: <https://doi.org/10.1093/nar/gky955>, PMID: 30357393
- Furlan G**, Gutierrez Hernandez N, Huret C, Galupa R, van Bemmel JG, Romito A, Heard E, Morey C, Rougeulle C. 2018. The ftx noncoding locus controls X chromosome inactivation independently of its RNA products. *Molecular Cell* **70**:462–472. DOI: <https://doi.org/10.1016/j.molcel.2018.03.024>, PMID: 29706539
- George M**, Duan Q, Nagle A, Kathiriya I, Huang Y, Rao K, Halder S, Bruneau B. 2019. Minimal in vivo requirements for developmentally regulated cardiac long intergenic non-coding RNAs. *bioRxiv*. DOI: <https://doi.org/10.1101/793257>
- Grünweller A**, Wyszko E, Bieber B, Jahnel R, Erdmann VA, Kurreck J. 2003. Comparison of different antisense strategies in mammalian cells using locked nucleic acids, 2'-O-methyl RNA, phosphorothioates and small interfering RNA. *Nucleic Acids Research* **31**:3185–3193. DOI: <https://doi.org/10.1093/nar/gkg409>, PMID: 12799446
- Han X**, Zhang J, Liu Y, Fan X, Ai S, Luo Y, Li X, Jin H, Luo S, Zheng H, Yue Y, Chang Z, Yang Z, Tang F, He A, Shen X. 2019. The lncRNA Hand2os1/Uph locus orchestrates heart development through regulation of precise expression of Hand2. *Development* **146**:dev.176198. DOI: <https://doi.org/10.1242/dev.176198>
- Herrmann CJ**, Schmidt R, Kanitz A, Artimo P, Gruber AJ, Zavolan M. 2020. PolyASite 2.0: a consolidated atlas of polyadenylation sites from 3' end sequencing. *Nucleic Acids Research* **48**:D174–D179. DOI: <https://doi.org/10.1093/nar/gkz918>, PMID: 31617559
- Hnisz D**, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. 2017. A phase separation model for transcriptional control. *Cell* **169**:13–23. DOI: <https://doi.org/10.1016/j.cell.2017.02.007>, PMID: 28340338
- Hon CC**, Ramilowski JA, Harshbarger J, Bertin N, Rackham OJ, Gough J, Denisenko E, Schmeier S, Poulsen TM, Severin J, Lizio M, Kawaji H, Kasukawa T, Itoh M, Burroughs AM, Noma S, Djebali S, Alam T, Medvedeva YA, Testa AC, et al. 2017. An atlas of human long non-coding RNAs with accurate 5' ends. *Nature* **543**:199–204. DOI: <https://doi.org/10.1038/nature21374>, PMID: 28241135
- Houseley J**, Rubbi L, Grunstein M, Tollervy D, Vogelauer M. 2008. A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. *Molecular Cell* **32**:685–695. DOI: <https://doi.org/10.1016/j.molcel.2008.09.027>, PMID: 19061643
- Ibrahim DM**, Mundlos S. 2020. The role of 3D chromatin domains in gene regulation: a multi-faceted view on genome organization. *Current Opinion in Genetics & Development* **61**:1–8. DOI: <https://doi.org/10.1016/j.gde.2020.02.015>
- Ingolia NT**, Lareau LF, Weissman JS. 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**:789–802. DOI: <https://doi.org/10.1016/j.cell.2011.10.002>, PMID: 22056041
- Isoda T**, Moore AJ, He Z, Chandra V, Aida M, Denholtz M, Piet van Hamburg J, Fisch KM, Chang AN, Fahl SP, Wiest DL, Murre C. 2017. Non-coding transcription instructs chromatin folding and compartmentalization to

- dictate Enhancer-Promoter communication and T cell fate. *Cell* **171**:103–119. DOI: <https://doi.org/10.1016/j.cell.2017.09.001>, PMID: 28938112
- Ji Z, Song R, Regev A, Struhl K. 2015. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *eLife* **4**:e08890. DOI: <https://doi.org/10.7554/eLife.08890>
- Kapusta A, Kronenberg Z, Lynch VJ, Zhuo X, Ramsay L, Bourque G, Yandell M, Feschotte C. 2013. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLOS Genetics* **9**:e1003470. DOI: <https://doi.org/10.1371/journal.pgen.1003470>, PMID: 23637635
- Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H, Nureki O, Zhang F. 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**:583–588. DOI: <https://doi.org/10.1038/nature14136>, PMID: 25494202
- Kraft K, Geuer S, Will AJ, Chan WL, Paliou C, Borschiwer M, Harabula I, Wittler L, Franke M, Ibrahim DM, Kragesteen BK, Spielmann M, Mundlos S, Lupiáñez DG, Andrey G. 2015. Deletions, inversions, duplications: engineering of structural variants using CRISPR/Cas in mice. *Cell Reports* **10**:833–839. DOI: <https://doi.org/10.1016/j.celrep.2015.01.016>
- Krchnáková Z, Thakur PK, Krausová M, Bieberstein N, Haberman N, Müller-McNicoll M, Stanek D. 2019. Splicing of long non-coding RNAs primarily depends on polypyrimidine tract and 5' splice-site sequences due to weak interactions with SR proteins. *Nucleic Acids Research* **47**:911–928. DOI: <https://doi.org/10.1093/nar/gky1147>, PMID: 30445574
- Lagarde J, Uszczyńska-Ratajczak B, Carbonell S, Pérez-Lluch S, Abad A, Davis C, Gingeras TR, Frankish A, Harrow J, Guigo R, Johnson R. 2017. High-throughput annotation of full-length long noncoding RNAs with capture long-read sequencing. *Nature Genetics* **49**:1731–1740. DOI: <https://doi.org/10.1038/ng.3988>
- Lai F, Damle SS, Ling KK, Rigo F. 2020. Directed RNase H cleavage of nascent transcripts causes transcription termination. *Molecular Cell* **77**:1032–1043. DOI: <https://doi.org/10.1016/j.molcel.2019.12.029>
- Lander S, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, et al. 2001. Initial sequencing and analysis of the human genome international Human genome sequencing consortium\* the sanger centre: beijing genomics institute/Human genome center. *Nature* **409**:860–921. DOI: <https://doi.org/10.1038/35057062>
- Latos PA, Pauler FM, Koerner MV, Şenergin HB, Hudson QJ, Stocsits RR, Allhoff W, Stricker SH, Klement RM, Warczok KE, Aumayr K, Pasierbek P, Barlow DP. 2012. Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science* **338**:1469–1472. DOI: <https://doi.org/10.1126/science.1228110>, PMID: 23239737
- Laurent F, Girdziusaitė A, Gamart J, Barozzi I, Osterwalder M, Akiyama JA, Lincoln J, Lopez-Rios J, Visel A, Zuniga A, Zeller R. 2017. HAND2 target gene regulatory networks control atrioventricular canal and cardiac valve development. *Cell Reports* **19**:1602–1613. DOI: <https://doi.org/10.1016/j.celrep.2017.05.004>, PMID: 28538179
- Lavalou P, Eckert H, Damy L, Constanty F, Majello S, Bitetti A, Graindorge A, Shkumatava A. 2019. Strategies for genetic inactivation of long noncoding RNAs in zebrafish. *RNA* **25**:897–904. DOI: <https://doi.org/10.1261/rna.069484.118>, PMID: 31043511
- Lee JS, Mendell JT. 2020. Antisense-Mediated transcript knockdown triggers premature transcription termination. *Molecular Cell* **77**:1044–1054. DOI: <https://doi.org/10.1016/j.molcel.2019.12.011>, PMID: 31924448
- Li Y, Haarhuis JHI, Sedeño Cacciatore Á, Oldenkamp R, van Ruiten MS, Willems L, Teunissen H, Muir KW, de Wit E, Rowland BD, Panne D. 2020. The structural basis for cohesin-CTCF-anchored loops. *Nature* **578**:472–476. DOI: <https://doi.org/10.1038/s41586-019-1910-z>, PMID: 31905366
- Lubelsky Y, Ulitsky I. 2018. Sequences enriched in alu repeats drive nuclear localization of long RNAs in human cells. *Nature* **555**:107–111. DOI: <https://doi.org/10.1038/nature25757>, PMID: 29466324
- Mattioli K, Volders PJ, Gerhardinger C, Lee JC, Maass PG, Melé M, Rinn JL. 2019. High-throughput functional analysis of lncRNA core promoters elucidates rules governing tissue specificity. *Genome Research* **29**:344–355. DOI: <https://doi.org/10.1101/gr.242222.118>, PMID: 30683753
- McFadden DG, Charité J, Richardson JA, Srivastava D, Firulli AB, Olson EN. 2000. A GATA-dependent right ventricular enhancer controls dHAND transcription in the developing heart. *Development* **127**:5331–5341. PMID: 11076755
- Mercer TR, Gerhardt DJ, Dinger ME, Crawford J, Trapnell C, Jeddelloh JA, Mattick JS, Rinn JL. 2012. Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nature Biotechnology* **30**:99–104. DOI: <https://doi.org/10.1038/nbt.2024>
- Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebbersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG. 2017. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**:930–944. DOI: <https://doi.org/10.1016/j.cell.2017.05.004>, PMID: 28525758
- Paralkar VR, Taborda CC, Huang P, Yao Y, Kossenkov AV, Prasad R, Luan J, Davies JO, Hughes JR, Hardison RC, Blobel GA, Weiss MJ. 2016. Unlinking an lncRNA from its associated Cis element. *Molecular Cell* **62**:104–110. DOI: <https://doi.org/10.1016/j.molcel.2016.02.029>, PMID: 27041223
- Ransohoff JD, Wei Y, Khavari PA. 2018. The functions and unique features of long intergenic non-coding RNA. *Nature Reviews Molecular Cell Biology* **19**:143–157. DOI: <https://doi.org/10.1038/nrm.2017.104>, PMID: 29138516
- Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, Huang X, Shamim MS, Shin J, Turner D, Ye Z, Omer AD, Robinson JT, Schlick T,

- Bernstein BE, Casellas R, et al. 2017. Cohesin loss eliminates all loop domains. *Cell* **171**:305–320. DOI: <https://doi.org/10.1016/j.cell.2017.09.026>, PMID: 28985562
- Ritter N, Ali T, Kopitchinski N, Schuster P, Beisaw A, Hendrix DA, Schulz MH, Müller-McNicoll M, Dimmeler S, Grote P. 2019. The lncRNA locus handsdown regulates cardiac gene programs and is essential for early mouse development. *Developmental Cell* **50**:644–657. DOI: <https://doi.org/10.1016/j.devcel.2019.07.013>, PMID: 31422919
- Rom A, Melamed L, Gil N, Goldrich MJ, Kadir R, Golan M, Biton I, Perry RB, Ulitsky I. 2019. Regulation of CHD2 expression by the chaserr long noncoding RNA gene is essential for viability. *Nature Communications* **10**:5092. DOI: <https://doi.org/10.1038/s41467-019-13075-8>, PMID: 31704914
- Santoro F, Mayer D, Klement RM, Warczuk KE, Stukalov A, Barlow DP, Pauler FM. 2013. Imprinted Igf2r silencing depends on continuous airn lncRNA expression and is not restricted to a developmental window. *Development* **140**:1184–1195. DOI: <https://doi.org/10.1242/dev.088849>, PMID: 23444351
- Schlackow M, Nojima T, Gomes T, Dhir A, Carmo-Fonseca M, Proudfoot NJ. 2017. Distinctive patterns of transcription and RNA processing for human lincRNAs. *Molecular Cell* **65**:25–38. DOI: <https://doi.org/10.1016/j.molcel.2016.11.029>, PMID: 28017589
- Szabo Q, Bantignies F, Cavalli G. 2019. Principles of genome folding into topologically associating domains. *Science Advances* **5**:eaaw1668. DOI: <https://doi.org/10.1126/sciadv.aaw1668>, PMID: 30989119
- Talbert PB, Meers MP, Henikoff S. 2019. Old cogs, new tricks: the evolution of gene expression in a chromatin context. *Nature Reviews Genetics* **20**:283–297. DOI: <https://doi.org/10.1038/s41576-019-0105-7>, PMID: 30886348
- Tamura M, Amano T, Shiroishi T. 2014. The Hand2 gene dosage effect in developmental defects and human congenital disorders. *Current Topics in Developmental Biology* **110**:8. DOI: <https://doi.org/10.1016/B978-0-12-405943-6.00003-8>
- van Heesch S, van Iterson M, Jacobi J, Boymans S, Essers PB, de Bruijn E, Hao W, MacInnes AW, Cuppen E, Simonis M. 2014. Extensive localization of long noncoding RNAs to the cytosol and mono- and polyribosomal complexes. *Genome Biology* **15**:1–12. DOI: <https://doi.org/10.1186/gb-2014-15-1-r6>
- van Heesch S, Witte F, Schneider-Lunitz V, Schulz JF, Adami E, Faber AB, Kirchner M, Maatz H, Blachut S, Sandmann CL, Kanda M, Worth CL, Schafer S, Calviello L, Merriott R, Patone G, Hummel O, Wyler E, Obermayer B, Mücke MB, et al. 2019. The translational landscape of the human heart. *Cell* **178**:242–260. DOI: <https://doi.org/10.1016/j.cell.2019.05.010>, PMID: 31155234
- van Steensel B, Furlong EEM. 2019. The role of transcription in shaping the spatial organization of the genome. *Nature Reviews Molecular Cell Biology* **20**:327–337. DOI: <https://doi.org/10.1038/s41580-019-0114-6>, PMID: 30886333
- Wang W, Zhou R, Wu Y, Liu Y, Su W, Xiong W, Zeng Z. 2019. PVT1 promotes Cancer progression via MicroRNAs. *Frontiers in Oncology* **9**:609. DOI: <https://doi.org/10.3389/fonc.2019.00609>, PMID: 31380270
- Yin Y, Yan P, Lu J, Song G, Zhu Y, Li Z, Zhao Y, Shen B, Huang X, Zhu H, Orkin SH, Shen X. 2015. Opposing roles for the lncRNA haunt and its genomic locus in regulating HOXA gene activation during embryonic stem cell differentiation. *Cell Stem Cell* **16**:504–516. DOI: <https://doi.org/10.1016/j.stem.2015.03.007>