The multidimensional mechanisms of long noncoding RNA function

Francesco P. Marchese¹,², Ivan Raimondi¹,² and Maite Huarte¹,²*

Abstract
A major shift in our understanding of genome regulation has emerged recently. It is now apparent that the majority of cellular transcripts do not code for proteins, and many of them are long noncoding RNAs (lncRNAs). Increasingly, studies suggest that lncRNAs regulate gene expression through diverse mechanisms. We review emerging mechanistic views of lncRNAs in gene regulation in the cell nucleus. We discuss the functional interactions that lncRNAs establish with other molecules as well as the relationship between lncRNA transcription and function. While some of these mechanisms are specific to lncRNAs, others might be shared with other types of genes.

Introduction
The appreciation of the complexity of the human transcriptome has revolutionized our perception of the regulatory potential of RNA. The efforts to generate a comprehensive atlas of the transcripts expressed in cells have revealed an extremely large collection of lncRNAs [1, 2]. The lncRNAs are broadly defined as noncoding RNA molecules longer than 200 nucleotides. Most of them are transcribed by RNA polymerase II, thus sharing similarities with messenger RNAs (mRNAs)—including a 5′ 7-methylguanosine cap and a 3′ poly(A) tail—however, they lack coding capacity. To date, the ENCODE project (GENCODE v26) has conservatively annotated in humans close to 16,000 lncRNA genes that give rise to more than 28,000 distinct transcripts. Moreover, protein-coding genes too can produce transcript variants that lack coding capacity, adding to the vast catalogue of long noncoding transcripts present in the cells.

* Correspondence: maitehuarte@unav.es
¹University of Navarra, Center for Applied Medical Research (CIMA), Pamplona 31008, Spain
²Institute of Health Research of Navarra (IdiSNA), Pamplona 31008, Spain

Despite not being translated into proteins, lncRNAs are functional molecules. Indeed, since the early studies that demonstrated the central role of Xist in the process of X-chromosome inactivation [3, 4], a growing body of evidence has described a myriad of functions for lncRNAs in many cellular processes, such as gene imprinting [5], differentiation and development [6], antiviral response [7], and vernalization in plants [8]. Among the variety of mechanisms reported (Fig. 1; Table 1), many lncRNAs have been shown to interact with chromatin-modifying complexes, to be involved in the conformation of nuclear domains, or in the activity of transcriptional enhancers [9–12]; others have been shown to interfere with the transcriptional machinery or maintain the structure of nuclear speckles [13–15]. Furthermore, some lncRNAs act post-transcriptionally as regulators of splicing, mRNA decay, protein translation, protein stability, or as molecular decoys for microRNAs (reviewed in [16, 17]). What has emerged from all these studies is that lncRNAs are highly heterogeneous and have a substantial functional versatility that relies on their ability as long RNA molecules to conform to different structures and molecular interactions. Moreover, the deregulation of lncRNAs has been related to different human diseases, including cancer and cardiovascular and neurodegenerative diseases [18, 19].

However, despite the rapid growth of the field, intriguing questions remain, such as whether all or just a fraction of the existing lncRNAs have a function, or whether this function can be exclusively ascribed to the RNA product of the lncRNA gene. Here, by reviewing the literature, we highlight and discuss different modes of action of lncRNAs in regulating gene expression in the nucleus. We distinguish between mechanisms that are inherent to the RNA molecule or linked to its gene locus, and recapitulate the current evidence that supports the concept that, indeed, the majority of the lncRNAs might be functionally relevant, although highly heterogeneous in their mode of operation.
Beyond lncRNA transcription: trans-regulatory activities of lncRNAs

A significant body of studies where the lncRNA is specifically depleted without perturbing its gene locus supports the notion that many lncRNAs are active species in regulating gene expression of local or distal genes in different organisms, including yeast, plants and higher eukaryotes [6, 20, 21]. Many of these lncRNAs localize to cellular compartments different from their own locus of transcription, including other nuclear domains, the cytoplasm or even polysomes [22]. These observations suggest that some aspects of lncRNA function are strictly dependent on the inherent properties of the RNA molecules, including their ability to fold into different structures and to conduct molecular interactions with other nucleic acids (i.e. RNA and DNA) and proteins. Furthermore, the long sequences of lncRNAs can contain multiple functional domains that interact with different factors coordinating their activity in time and space. For instance, several lncRNAs act in cooperation with heterogeneous nuclear ribonucleoproteins (hnRNPs) [23–26], a large family of RNA-binding proteins involved in different cellular processes, including alternative splicing, mRNA stability, and transcriptional regulation [27]. In addition, it is worth noting the growing evidence showing that some proteins that lack canonical RNA binding domains are able to bind RNA, expanding the number of potential binding partners for lncRNAs, and so extending our view of their regulatory potential over the cell proteome [28]. Here, we discuss some of the most prominent types of functional interactions reported for lncRNAs in the nucleus.

lncRNA interactions with chromatin complexes

A large repertoire of lncRNAs are able to interact with chromatin-modifying complexes. Given the central role of these complexes during development and disease, this mode of action has been the subject of great attention, and it has been proposed that lncRNAs place these proteins at specific gene loci to achieve appropriate temporal and spatial gene regulation (reviewed in [12, 29, 30]). The lncRNA-containing complexes can promote either selective repression or activation of genes, according to the nature of the chromatin complex [12, 31–33]. For instance, several lncRNAs have been shown to recruit histone H3K4 methyltransferases to promote activation of gene expression [34–36]. Others bind DNA methyltransferases, such as DNMT1 and DNMT3b, and therefore repress transcription by promoting DNA methylation [37–39]. Finally, many lncRNAs have been shown to interact with the polycomb repressive complex 2 (PRC2), which catalyses generation of the H3K27me3 silencing mark [34, 40–45]. For several reasons (reviewed in [46]), PRC2 is the most studied chromatin complex with respect to the functional role of lncRNAs in epigenetic regulation of gene expression. This
lncRNAs as modulators of proteins and enzyme cofactors

Table 1 Long-noncoding-RNA-mediated nuclear mechanisms (Continued)

| Long noncoding RNA | Function | Reference |
|---------------------|----------|-----------|
| VIM-AS1             | Promotion of transcriptional activation of the VIM gene | [75] |
| COOLAIR             | Negative regulation of FLC expression and flowering | [77] |
| Khps1               | Transcriptional activation of SPHK1 | [82] |
| MEG3                | Transcriptional regulation of genes of the TGF-β pathway | [83] |

LncRNAs in higher-order structures

- Xist: X chromosome inactivation [106]
- Fire: Role in adipogenesis; mediates inter-chromosomal interactions [107, 124]
- NEAT1: Nucleation of paraspeckles [14, 15, 109]
- MALAT1: Formation of nuclear speckles [110]

LncRNAs as modulators of proteins and enzyme cofactors

In most of the known examples, the detailed mechanisms by which lncRNAs enhance or inhibit the activity of proteins is not completely understood, but are probably diverse and not restricted to the control of protein localization on the chromatin. In some cases, lncRNAs can interfere in the interaction between proteins and protein-modifying enzymes, impeding posttranslational modifications and impacting important signalling pathways. A study describing an NF-kappaB (NF-κB)-interacting lncRNA (NKILA) showed that this lncRNA binds NF-κB/IκB in a ternary complex [48]. In this context, using in vitro kinase assays, the addition of NKILA to the reaction was found to inhibit IKK-mediated IκB phosphorylation by directly masking the phosphorylation sites, leading to NF-κB inactivation [48]. Also the lncRNA expressed in dendritic cells (Inc-DC) regulates protein modification. Inc-DC was found to regulate the expression of genes involved in dendritic cell (DC) differentiation [49]. The results suggested a mechanism that involves line-DC interaction with the transcription factor signal transducer and activator of transcription 3.
Such interaction was shown to prevent dephosphorylation of STAT3 at tyrosine Y705 by the tyrosine phosphatase SHP1, and in this way to control the transcriptional programme for differentiation of DCs (Fig. 2a) [49].

Another emerging mode of action of lncRNAs involves their ability to modulate the enzymatic activity of some proteins. Among the early studies identifying and annotating lncRNAs [45, 50, 51], a set of them, named as activating noncoding RNA (ncRNA-a), was found to have enhancer-like functions on the neighbouring protein-coding genes [52]. Two such enhancer-like RNAs (ncRNA-a3 and ncRNA-a7) were found associated with mediator, a transcriptional co-activator complex (Fig. 2b) [53] and increased the occupancy of two different sub-units of the complex on the regulated neighbour genes [53]. Interestingly, the lncRNAs were shown to specifically stimulate the kinase activity of mediator towards the histone H3 [53]. The results obtained with ncRNA-a3 and ncRNA-a7 [53] suggested that the direct interaction between an lncRNA and a protein is responsible for the proper enzymatic activity of the protein partner.

CONCR (cohesion regulator noncoding RNA) is also included in the small number of lncRNAs reported so far to modulate an enzymatic activity [54]. CONCR has been shown to physically interact with DEAD/H box

---

**Fig. 2** Long noncoding RNAs (lncRNAs) modulate protein activities. a The Inc-DC regulates gene expression through a mechanism of action that involves its translocation to the cytoplasm, where it interacts with phosphorylated STAT3 and prevents pY705-STAT3 dephosphorylation by the tyrosine phosphatase SHP1 [49]. b The lncRNAs ncRNA-a3 and ncRNA-a7 functionally and physically associate with the Mediator complex to promote gene expression of their respective target genes. Either ncRNA-a stimulates the kinase activity of the CDK8 subunit of the complex towards the histone H3, catalysing the phosphorylation of the serine 10 (H3S10) [53]. c CONCR functions in sister-chromatid cohesion by binding and promoting the ATPase activity of DDX11 during DNA replication [54]
different modes of action have been suggested with respect to the ability of DRBPs to bind DNA and RNA—for example simultaneous or competitive binding. For instance, the IncRNA GAS5 has been identified as a functional and physical interactor of the glucocorticoid receptor (GR) [58]. GAS5 was found to interact with the activated GR to suppress its binding to glucocorticoid response elements (GREs) and therefore the expression of the glucocorticoid-responsive genes (Fig. 3a) [58]. Interestingly, the GAS5–GR interaction occurs at the DNA-binding domain of the transcription factor, likely through a mimetic GRE that forms in the secondary structure of GAS5 [58]. This evidence not only suggests that a transcription factor is equally able to bind DNA and RNA, but also indicates that a lncRNA can regulate gene expression by acting as a binding competitor for DBPs.

More recently, while investigating the role of lncRNAs in the p53 pathway, a novel lncRNA, named ‘damage-induced noncoding RNA’ (DINO), was identified in the regulation of the DNA-damage-induced p53 response [59]. DINO binds to and stabilizes p53 (Fig. 3b) [59]. Their interaction was found to involve the C-terminus of p53 and to be maintained while p53 is bound to the responsive elements of its target genes [59]. The tumour suppressor p53 is known to have two distinct nucleic acid binding domains, a DNA-binding core domain and a second nucleic acid-binding domain located at the C-terminus, which has long been known to be able to bind RNA, although earlier studies only obtained evidence of binding in vitro [67]. The results obtained for DINO support the notion that p53 can bind simultaneously DNA and RNA and ascribe to the lncRNA a regulatory role in p53-mediated gene expression.

The emerging evidence concerning the active role of lncRNAs on transcription factors suggests that gene expression benefits from an additional mode of regulation. Indeed, the possibility that a lncRNA can bind and regulate a transcription factor, or any other protein involved in gene expression, confers numerous advantages to the cell. For instance, lncRNAs are known to be highly cell and tissue specific [68, 69], which means that, without changing the transcriptional machinery, cell- and tissue-specific regulation of gene expression could be achieved. Alternatively, lncRNA-dependent regulation could also be considered in terms of cost-effectiveness, as RNAs are energetically less expensive to produce for the cell in comparison with proteins. Also, lncRNAs are more rapidly produced than proteins in response to stimuli—a favourable dynamic that could confer faster cellular responses. Moreover, they can act locally at their site of transcription, whereas proteins need to be translated from their encoding RNAs in the cytoplasm and be returned to the nucleus. Although the lncRNA diversity
remains underappreciated, and their functionality poorly characterized, the evidence obtained to date and the model proposed suggest that IncRNA–DRBP interactions are as relevant as protein–protein interactions in the regulation of gene expression.

Direct interaction between IncRNAs and the DNA: R-loops and triple helixes

While the interaction with transcription factors might confer some IncRNAs with the capacity to recognize specific gene loci, the specificity in gene regulation by IncRNAs has also been linked to their ability as nucleic acids to directly bind to the genomic DNA. These interactions might be key for target recognition. One of the mechanisms whereby IncRNA directly binds DNA involves the formation of structures known as R-loops. These are nucleic acid structures that form usually during transcription by reannealing of the nascent RNA to the DNA template, giving rise to a RNA–DNA hybrid and a displaced single-stranded DNA (ssDNA) [70].

When not properly resolved, R-loops might induce DNA damage and genomic instability [71]. For instance, the telomeric IncRNA TERRA has been shown to form R-loops at short telomeres in yeast, contributing to the activation of the DNA damage response by promoting recruitment of the Rad51 recombinase in a telomere-length-dependent manner [72]. Interestingly, the murine TERRA can also act at distant genomic loci, where it binds and antagonizes the chromatin-remodeler ATRX [73].

In contrast to the above, R-loops can regulate gene expression [74]. For the IncRNAs found to regulate mRNA transcription through the formation of R-loops, the mechanism of regulation is in cis as the R-loop formed by the transcription of the IncRNA regulates the expression of the closest protein-coding gene, transcribed in antisense orientation with respect to the IncRNA. This is, for example, the situation observed for the IncRNA VIM-AS1, which forms an R-loop around the promoter for the gene encoding vimentin (VIM), which, in turn, causes chromatin opening and enhances the binding of
transcriptional activators of the NF-kB pathway [75]. By contrast, an R-loop-dependent transcriptional repression has been observed for the FLOWERING LOCUS C (FLC) in Arabidopsis thaliana. FLC expression and flowering are known to be regulated by multiple pathways [76], as well being as negatively regulated by a lncRNA named COOLAIR [77]. COOLAIR is oriented antisense to the FLC gene, and its transcription, and R-loop formation, represses FLC expression during prolonged periods of low temperature, preventing in this way the plant from flowering [77]. R-loop stabilization (by binding of the protein NODULIN HOMEBOX to the ssDNA of the R-loop) in the promoter region of COOLAIR has by contrast been found to repress COOLAIR transcription, allowing FLC expression [77]. To date, our knowledge of the involvement of lncRNAs in R-loop-mediated gene regulation remains very limited. However, considering the high number of sense–antisense paired genes in the genome [78, 79], as well as the ability of some lncRNAs to form R-loops in trans [80], it is likely that other lncRNAs will be identified to regulate gene expression via R-loops.

Another way of recognizing chromatin DNA by lncRNAs is the formation of RNA–DNA triplexes [81]. Such structures are generally thought to serve as an anchor for the recruitment of chromatin modifiers in proximity to the gene promoters [81]. This was originally reported for noncoding RNAs produced in the promoter of ribosomal RNAs, which, by forming local triple-helix structures, can recruit the DNA methyltransferase DNMT3b and induce silencing of rRNA genes [38]. Similarly, it has been suggested that RNA–DNA triplexes mediate the recruitment of PRC2 and trithorax-group/mixed-lineage leukemia (TrxG/Mll) protein complexes, both in cis (on the Foxf1 gene) and in trans (on the Pitx2 gene), by the murine lncRNA Fendrr [34]. By combining in vitro pull-down and in vivo triplex-capture assays, the lncRNA Khps1 was shown to form a RNA–DNA triplex [82]. Khps1 is an antisense RNA of the proto-oncogene SPHK1 encoding sphingosine kinase 1 and was shown to form the triplex upstream of the transcription start site of SPHK1—in this way activating SPHK1 expression by recruiting the histone acetyltransferase p300/CBP [82]. More recently, other lncRNAs have been reported to recruit chromatin modifiers in an RNA–DNA triplex-dependent manner, at promoters of both neighbouring and distal genes [83–85]. An intriguing possibility is that RNA–DNA triplexes formed by lncRNAs could serve to bypass the need for additional proteins for the specific recruitment of epigenetic factors, although much remains to be understood regarding what drives the lncRNAs to form triplexes and how their formation and resolution dynamics are regulated.

**IncRNA functions linked to their own gene locus**

The activity of several noncoding RNAs is tightly connected to their own site of transcription. In this instance, the IncRNA remains in the proximity of its own locus, where its transcription seems to be closely linked to the local association of factors that determine an active or repressive chromatin state. It is thus challenging to uncouple the function of the IncRNA from other inherent attributes of the gene locus, including DNA sequence, transcriptional state and chromatin conformation. Here, we describe mechanisms of IncRNAs in this context.

**Active enhancers as a source of functional ncRNAs**

Enhancers are short regions of DNA that are bound by transcription factors and augment the expression of genes contacted through chromatin ‘looping’ [86]. They represent a paradigmatic example of noncoding loci that involves regulation by multiple elements: DNA sequence, three-dimensional (3D) chromatin conformation and RNA transcription. Enhancers generally produce unspliced and non-polyadenylated transcripts named enhancer RNAs (eRNAs). Several studies have evaluated the relationship between enhancer transcripts and gene activation (reviewed in [86]), showing that the levels of eRNAs correlate with mRNA synthesis of the neighbouring genes [51, 87]. Also supporting the notion that eRNAs contribute to enhancer function, eRNA levels correlate with the activity of the enhancer [53, 88, 89], and enhancers that produce eRNAs bind transcriptional co-activators to a higher extent [51, 87], have higher DNase hypersensitivity and marks of active chromatin compared with non-transcribed enhancers [86, 90]. While some work has shown that the eRNA is dispensable for the deposition of active histone marks at enhancers [91], several other studies have provided abundant proof that eRNAs are functional as RNA species. For instance, eRNAs produced from enhancers adjacent to E2-upregulated genes are required for the observed ligand-dependent induction of target coding genes, increasing the strength of specific enhancer–promoter looping [89]. Similar observations have been made for eRNAs transcribed from androgen-receptor-regulated enhancers [92], a gonadotropin hormone α-subunit gene enhancer [93] or MYOD1 enhancers, where eRNAs regulate the chromatin accessibility of the transcriptional machinery [94]. Although it is still not well understood how eRNAs boost enhancer activity, a study performed in neurons showed that eRNAs might facilitate the transition of paused RNA polymerase II into productive elongation by acting as a decoy for the negative elongation factor (NELF) complex [95]. This evidence suggests that the eRNA can establish local interactions with proteins that either enhance or inhibit the transcriptional activity of the enhancer, constituting an additional element of the enhancer function.
Cis-regulation by lncRNAs: act of transcription versus RNA product

A scenario where the function of the noncoding RNA is linked to both the RNA product and the act of transcription could well also apply to some lncRNAs. For example, the lncRNA Airn, originally identified in mouse as promoting genomic imprinting of the maternal protein-coding Igf2r gene cluster, where Airn itself is encoded [96], was later shown to act independently of its RNA [97]. Indeed, by shortening the endogenous Airn to different lengths by homologous recombination or repositioning its promoter, Airn-mediated silencing of Igf2r was shown to be caused by transcriptional interference, where the transcriptional overlap of Airn reduced the recruitment of RNAPII to the Igf2r promoter, independently of its lncRNA product (Fig. 4a) [97]. However, transcriptional overlap could not explain the imprinting of the other genes in the Igf2r cluster—Slc22a2 and Slc22a3—suggesting that at least some of Airn silencing properties do reside in its RNA and not just the act of transcription (Fig. 4a) [97, 98].

More recently, using a similar approach to that described for Airn, genome editing of the lncRNA Lockd locus was used to investigate whether the lncRNA functions through its transcript or DNA elements contained in its genomic locus [99]. The study suggested that Cdkn1b is positively regulated by a cis-element at the promoter of Lockd, whereas the lncRNA is dispensable for this function (Fig. 4b) [99]. The authors of the work could not exclude the possibility that the Lockd transcript exerts other functions [99]. However, considering that the transcriptomic analysis, following complete depletion of Lockd, showed Cdkn1b as the only gene affected significantly, this lncRNA might well represent an actual case of there being a functional byproduct of the Lockd cis-element.

Similarly, to gain insight into the relationship between lncRNAs and regulation of gene expression in cis, Engreitz and colleagues approached the question on a larger scale, evaluating the effects of the genetic manipulation of 12 lncRNA and six mRNA loci on the expression of nearby genes [100]. By combining CRISPR–Cas9-mediated genome editing (including heterozygous deletions of the promoters, insertion of polyadenylation signals downstream of the transcription start sites, and deletions of exonic, intronic or splicing sites) with a variety of measurements (i.e. RNA-seq, GRO-seq, ChIP-seq), the investigators observed that functional interactions

![Fig. 4 Gene regulation linked to long noncoding RNA (lncRNA) transcription.](https://example.com/fig4.png)

Fig. 4 Gene regulation linked to long noncoding RNA (lncRNA) transcription. a Airn transcription reduces the recruitment of RNAPII to the Igf2r promoter, a mechanism known as ‘transcriptional interference’, causing the silencing of the maternal Igf2r gene. Transcriptional overlap does not occur at the other imprinted genes in the Igf2r cluster—Slc22a2 and Slc22a3—where the silencing properties of Airn might reside in its RNA [97, 98]. b Cdkn1b expression is positively regulated by a cis-element of the genomic locus of Lockd, whereas the lncRNA is dispensable for this function [99]
between neighbouring genes are frequent, both for lncRNAs and protein-coding loci, and that such cross-talk relies on different manners of function [100]. These include transcriptional- or co-transcriptional-related processes, such as promoters that act as proximal enhancers, where the RNA has no function per se, or the process of splicing that has cis-regulatory functions, partially dependent on the nascent transcript [100]. Although the number of loci evaluated in the study is still very limited considering the thousands of lncRNA–mRNA gene pairs present in the genome [101], none of the lncRNA loci included in the study seemed to require the lncRNA itself for the investigated cis-regulatory function [100]. It is interesting to note that similar results were achieved for the protein-coding genes analyzed, which points to the notion that genes, independently of their coding or noncoding status, can exert cis-regulatory activities that are independent of the RNA produced. However, it should be considered that, as this study only evaluated the effect of the lncRNAs over their neighbouring genes, no conclusion can be extrapolated regarding the functions of the mature RNA products beyond the regulation of local genes.

**InCRNAS and higher-order structures**

The distinction between cis- and trans-regulatory activity of lncRNAs is not clear when taking into consideration the 3D organization of nuclear compartments. The nuclear conformation can explain co-activation or co-repression of gene loci dependent on the special proximity and the local concentration of the involved factors, including lncRNAs. The genome is in fact a packaged 3D structure that forms higher-order chromatin structures, such as intra- and inter-chromosomal loops and nuclear compartmentalization [102]. This is nowadays known to be the result of specific and regulated interactions between DNA, proteins and RNAs, and lncRNAs are increasingly being recognized as important organizers of this architecture [9, 103, 104]. Moreover, emerging evidence supports the notion that nuclear compartmentalization could be based on liquid-phase separation dependent on the biophysical properties of the molecules that constitute the distinct chromatin domains [105]. In this context, the cooperative association of highly abundant lncRNAs with other RNAs and proteins might be significant for the formation and dynamics of nuclear compartments.

Among the lncRNAs involved in shaping the 3D structure of the genome, some seem to have a role in the compartmentalization of the DNA in a 3D proximity-guided mechanism, such as Xist, which scaffolds a number of proteins that in turn tether the lncRNA to genomic DNA [106], the lncRNA Firre that, through the binding of hnRNP-U, acts as a platform for trans-chromosomal interactions (Fig. 5a) [107], and the more recently identified trait-relevant (TR)-lincRNAs suggested to regulate proximal TR-protein-coding gene expression by modulating local chromosomal architecture [108]. By contrast, other lncRNAs have been found to form specific nuclear structures, such as the lncRNAs NEAT1 and MALAT1 involved in the nucleation of paraspeckles [14, 15, 109] and nuclear speckles [110], respectively.

The role of lncRNAs in nuclear organization is well exemplified by the lncRNA NEAT1, which, as mentioned above, drives the formation of nuclear bodies known as paraspeckles. Paraspeckles are subnuclear structures characterized by high local concentrations of specific proteins and RNAs, implicated in the regulation of gene expression by sequestering mRNAs and proteins involved in nuclear processes, including transcription [111, 112]. NEAT1 is an abundant mono-exonic lncRNA that, following processing at its 3′ end, produces a polyadenylated 3.7-kb NEAT1_1 isoform and a non-canonically processed 23-kb NEAT1_2 isoform [111]. Several studies have identified NEAT1_2 as an essential component for the formation of paraspeckles and have shown that the nucleation process begins during the biogenesis of the lncRNA and that the lncRNA acts as a scaffold for the binding of paraspeckle proteins (Fig. 5b) [14, 15, 109, 113, 114]. Under physiological conditions, the expression of NEAT1_2, and thereby paraspeckle formation, occurs in specific cell types involved in pregnancy and lactation [115, 116]. However, the appearance of paraspeckles is known to occur ubiquitously under certain stress conditions [111]. In line with this, NEAT1 has been identified recently as a direct transcriptional target of p53, and NEAT1-dependent formation of paraspeckles has been shown to contribute to the tumour-suppressor function of p53 [117, 118]. In conclusion, as paraspeckle formation requires both the NEAT1 RNA and its ongoing transcription, and NEAT1 expression is rapidly regulated in response to stress, the resulting nuclear compartmentalization represents a functional and dynamic manner for controlling gene expression and cellular responses.

**Concluding remarks**

Compiling evidence supports the involvement of IncRNAs in the correct execution of gene expression programs, which can be ascribed to three different levels of gene activity: (i) the underlying genomic sequence of the locus, which contains elements able to bind regulatory proteins such as transcription factors; (ii) the act of transcription that can either act as positive feedback or cause transcriptional interference; and (iii) the RNA product itself. The combination of these three dimensions of gene function together with the topological localization in the nucleus is therefore what mediates the effect on gene regulation. However, what makes functional IncRNAs unique is their
ability to establish molecular interactions with proteins and nucleic acids to temporally and spatially modulate their activities and/or localization. The examples reviewed here illustrate this mechanistic versatility of lncRNAs—flexibility arising also owing to their evolutionary plasticity. However, only the identification of the RNA sequences and structural elements that confer lncRNAs with these capabilities, as well as the determination of the biochemical and biophysical properties of the lncRNA-containing complexes, will lend further insight into the mechanisms that lncRNAs employ for gene regulation. As our comprehension of lncRNA mechanisms progresses, this will not only expand our view of transcriptional regulation, but also of other important biological processes centred on the chromatin, such as the DNA damage response, DNA repair and DNA replication. Considering the many and diverse functions of lncRNAs, it is therefore not surprising that their alterations contribute to the development and maintenance of many different human diseases. A better understanding of the mechanisms underlying the functions of lncRNAs will help us to understand the pathophysiology of human diseases and to design novel therapeutic strategies and will also benefit fundamental research.

Abbreviations

CONCR: Cohesion regulator noncoding RNA; DBP: DNA-binding protein; DDX11: DEAD/H box protein 11; DINO: Damage-induced noncoding RNA; DNMT: DNA methyltransferase; DRBP: DNA- and RNA-binding protein; eRNA: Enhancer RNA; FLC: FLOWERING LOCUS C; GRE: Glucocorticoid response element; lncRNA: Long noncoding RNA; MALAT1: Metastasis-associated lung adenocarcinoma transcript 1; ncRNA-a: Activating noncoding RNA; NEAT1: Nuclear enriched abundant transcript 1; NKILA: NF-kappaB-interacting lncRNA; PRC2: Polycomb repressive complex 2; RNA-binding protein; SOX2: Sex-determining region Y-box 2; STAT3: Signal transducer and activator of transcription 3; YY1: Yin yang 1

Acknowledgements

We thank the members of the Huarte laboratory for critical reading of the manuscript, and we apologize to colleagues whose work was not discussed or cited owing to space constraints.
References
1. Hon CC, Ramilowski JA, Harshbarger J, Bertin N, Guttman M, et al. An atlas of human long non-coding RNAs with accurate 5' ends. Nature. 2017;543:199–204.
2. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Green PH, Schneider R, Kiledjian M, Bilbao JR, Ghosh S. A long non-coding RNA associated with susceptibility to celiac disease. Science. 2016;352:91–5.
3. Brown CI, Hendrich BD, Rupert JL, Lafreniere RG, Xing Y, Lawrence J, Willard HF. 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. 1992;71:527–42.
4. Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, Rastan S. 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. 1992;71:515–26.
5. Kanduri C. Long noncoding RNAs: Lessons from genomic imprinting. Annu Rev Genet. 2014;48:433–50.
6. Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet. 2014;15:72–81.
7. Marchese FP, Huarte M. Long non-coding RNAs and chromatin modifiers: spatial amplifiers that control nuclear structure and gene expression. Nat Rev Mol Cell Biol. 2016;17:756–70.
8. Rinn JL, Chang HY. Genome regulation by long non-coding RNAs. Annu Rev Cell Dev Biol. 2010;26:409–35.
9. Heo JB, Yang YW, Liu B, Sanjal A, Corces-Zimmerman R, Chen Y, Lajec BR, Protacio A, Flynn RA, Gupta RA, et al. A long non-coding RNA maintains active chromatin to coordinate homeotic gene expression. Nature. 2011;472:120–4.
10. Cano IA, Soriano P. 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. 1992;71:527–42.
11. Carninci P, Kasukawa T, Katayama S, Gough J, Frat焕发 C, Maeda N, Noy A, Hayashizaki Y. The transcriptional map of the human genome. Nature. 2005;430:80–8.
12. Rinn JL. lncRNAs: linking RNA to chromatin. Cold Spring Harb Perspect Biol. 2014;6:a018614.
13. Guttman M, Rinn JL. lncRNAs: linking RNA to chromatin. Cold Spring Harb Perspect Biol. 2015;9:53–66.
14. Chen LL. Linking long noncoding RNA localization and function. Trends Biochem Sci. 2016;41:761–72.
15. Huarte M, Guttman M, Feldser D, Garber M, Kozlowski M, Kenzelmann-Broz D, Khalil AM, Zu K, Amit I, Rabani M, et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell. 2010;142:609–19.
16. Castellanos-Rubio A, Fernandez-Jimenez N, Kratchmarov R, Luo X, Bhagat G, Green PH, Schneider R, Kiledjian M, Bilbao JR, Ghosh S. A long non-coding RNA associated with susceptibility to celiac disease. Science. 2016;352:91–5.
17. Huarte M. The emerging role of lncRNAs in cancer. Nat Med. 2015;21:1253–61.
18. Bonasio R, Shekharra R. Regulation of transcription by long noncoding RNAs. Annu Rev Genet. 2014;48:833–55.
19. Ullity J. Evolution to the rescue. Using comparative genomics to understand long non-coding RNAs. Nat Rev Genet. 2016;17:601–14.
20. Wang Z, Zhang XJ, Ji YX, Zhang P, Deng KQ, Gong J, Chen I, Wang H, et al. The long non-coding RNA Chaer defines an epigenetic checkpoint in cardiac hypertrophy. Nat Med. 2016;22:1131–9.
21. Chen LL. Linking long noncoding RNA localization and function. Trends Biochem Sci. 2016;41:761–72.
22. Huarte M, Guttman M, Feldser D, Garber M, Kozlowski M, Kenzelmann-Broz D, Khalil AM, Zu K, Amit I, Rabani M, et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell. 2010;142:609–19.
23. Castellanos-Rubio A, Fernandez-Jimenez N, Kratchmarov R, Luo X, Bhagat G, Green PH, Schneider R, Kiledjian M, Bilbao JR, Ghosh S. A long non-coding RNA associated with susceptibility to celiac disease. Science. 2016;352:91–5.
24. Huarte M, Guttman M, Feldser D, Garber M, Kozlowski M, Kenzelmann-Broz D, Khalil AM, Zu K, Amit I, Rabani M, et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell. 2010;142:609–19.
25. Castellanos-Rubio A, Fernandez-Jimenez N, Kratchmarov R, Luo X, Bhagat G, Green PH, Schneider R, Kiledjian M, Bilbao JR, Ghosh S. A long non-coding RNA associated with susceptibility to celiac disease. Science. 2016;352:91–5.
26. Huarte M. The emerging role of lncRNAs in cancer. Nat Med. 2015;21:1253–61.
27. Bonasio R, Shekharra R. Regulation of transcription by long noncoding RNAs. Annu Rev Genet. 2014;48:833–55.
28. Ullity J. Evolution to the rescue. Using comparative genomics to understand long non-coding RNAs. Nat Rev Genet. 2016;17:601–14.
45. Khalil AM, Guttmann M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A. 2005;102:11667–72.

46. Davidovich C, Ceci TR. The recruitment of chromatin modifiers by long noncoding RNAs: lessons from PRC2. RNA. 2015;21:2007–22.

47. Porto-Seguro S, Badaró M, Bouchard CL, Carreira de Almeida TS, De Santa F, Cepulka AR, et al. p53 RNA interactions: new clues in an old mystery. Cell. 2015;161:20.

48. Caplin MN, Dunagin MC, McClanahan PD, Biaesch A, Padoyan-Merhar O, Riley KJ, Maher 3rd LJ. p53 RNA interactions: new clues in an old mystery. RNA. 2007;13:1825–33.

49. Caplin MN, Trapnell C, Goff L, Kozioł M, Tazon-Vega B, Regev A, Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 2011;25:1915–27.

50. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. Nat Rev Genet. 2015;16:583–97.

51. Sorlier J, Cimprich KA. Breaking bad: R-loops and genome integrity. Trends Cell Biol. 2015;25:514–22.

52. Graf M, Bonetti D, Lockhart D, Serhal K, Kellner V, Maicher A, Jolivet P, Teixeira MT, Luke B. Telomere length determines TERRA and R-Loop regulation through the cell cycle. Cell. 2017;170:72–85. e14.

53. Chu HP, Cifuentes-Rojas C, Kesner B, Aebly E, Lee HG, Wei C, Oh HJ, Boukhal M, Haas W, Lee JT. TERRA RNA antagonizes ATRX and protects telomeres. Cell. 2017;170:86–101. e116.

54. Scouoirti-Stathaki K, Proudfoot NJ. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev. 2014;28:1384–96.

55. Boque-Sastre R, Soler M, Oliveira-Mateos C, Portela A, Moutinho C, Sayols Y, Villanueva A, Esteller M, Gull S. Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. Proc Natl Acad Sci U S A. 2015;112:5785–90.

56. Kim DH, Sung S. Environmentally coordinated epigenetic silencing of FLc by protein and long noncoding RNA components. Curr Opin Plant Biol. 2012;15:51–6.

57. Sun Q, Cisbro T, Scouoirti-Stathaki K, Proudfoot NJ, Dean C. R-loop stabilization represses antisense transcription at the Arabidopsis FLC locus. Science. 2013;342:61–21.

58. Yein R, Dahary D, Sorek R, Levanon EY, Goldstein O, Shoshan A, Diber A, Biton S, Tamir Y, Khosravi R, et al. Widespread occurrence of antisense transcription in the human genome. Nat Biotechnol. 2003;21:379–86.

59. Chen J, Sun M, Kent WJ, Huang X, Xie H, Wang W, Zhou G, Shi RZ, Rowley JD. Over 20% of human transcripts might form sense-antisense pairs. Nucleic Acids Res. 2004;32:4812–20.

60. Coutier SC, Wang S, Ma WK, Al Husini N, Dhooondia Z, Anansri A, Pascu CE, Tran EJ. Regulated formation of ncRNA-DNA hybrids enables faster transcriptional induction and environmental adaptation. Mol Cell. 2016;61:393–404.

61. Li Y, Syed J, Sugiyama H. RNA-DNA triplex formation by long noncoding RNAs. Cell Chem Biol. 2016;23:1325–33.

62. Postepka-Igliksa A, Gajewska A, Gasi-Piotrowski L, Schmitt N, Dold A, Gisberg D, Grummt I. LncRNA R1ps1 regulates expression of the proto-oncogene SPK1 via triplex changes in chromatin modification. Mol Cell. 2015;60:2626–36.

63. Santos-Pereira JM, De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Regev A, Rinn JL. Discovery of active enhancers through DNA methylation. Cell. 2012;151:510–21.

64. Mondal T, Subhash S, Raj V, Enzo S, Reinius B, Mitra S. Highly conserved large non-coding RNAs in mammals. Nature. 2009;458:223–30.

65. Chu HP, Cifuentes-Rojas C, Kesner B, Aebly E, Lee HG, Wei C, Oh HJ, Boukhal M, Haas W, Lee JT. TERRA RNA antagonizes ATRX and protects telomeres. Cell. 2017;170:86–101. e116.

66. Ng SY, Johnson R, Stanton LW. Human long non-coding RNAs promote phosphorylation and suppress breast cancer metastasis. Cancer Cell. 2015;27:370–81.

67. Wang P, Yue X, Yan H, Lin L, Wu C, Xu S, Jiang Z, Xu J, Wu Q, Cao X. The STAT3-binding long noncoding RNA Inc-CC controls human dendritic cell differentiation. Science. 2014;344:310–3.
92. Hsieh CL, Fei T, Chen Y, Li T, Gao Y, Wang X, Sun T, Sweeney CJ, Lee GS, Chen S, et al. Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. Proc Natl Acad Sci U S A. 2014;111:7319–24.

93. Phuell L, Rudnicky S, Yosefzon Y, Melamed P. RNA transcribed from a distal enhancer is required for activating the chromatin promoter of the gonadotropin alpha-subunit gene. Proc Natl Acad Sci U S A. 2015;112:4369–74.

94. Moussavi K, Zare H, Delforo S, Groenvek L, Gutierrez-Cruz G, Derfoul A, Hager GL, Sartoriello V. eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. Mol Cell. 2013;53:1606–17.

95. Screutels F, Joo JY, Liu X, Watts JK, Martinez C, Rim TK. Enhancer RNA facilitates NELF release from immediate early genes. Mol Cell. 2014;56:29–42.

96. Screutels F, Zwart R, Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature. 2002;418:70–2.

97. Letos PA, Pauler FM, Koerner MV, Senergin HB, Hudson QJ, Stocsits RR, Allhoff W, Stricker SH, Klemens RM, Warczok KE, et al. Aim transcriptional overlap, but not its IncRNA products, induces imprinted Igf2r silencing. Science. 2012;338:1469–72.

98. Sasaki YT, Ideue T, Sano M, Mituyama T, Hirose T. MENepsilon/beta noncoding RNA Firre. Nat Struct Mol Biol. 2014;21:198–204.

99. Sasaki YT, Idogawa M, Ohashi T, Sasaki Y, Nakase H, Tokino T. Long non-coding RNA NEAT1 is a transcriptional target of p53 and modulates p53-induced transactivation and tumor-suppressor function. Int J Cancer. 2017;140:2785–91.

100. Sasaki YT, Ideue T, Sano M, Mituyama T, Hirose T. Paraspeckle formation during the biogenesis of long non-coding RNA Airn transcriptional assembly of a nuclear body by noncoding RNAs. Proc Natl Acad Sci U S A. 2013;110:33862–7.

101. Sasaki YT, Ideue T, Sano M, Mituyama T, Hirose T. Splicing and transcriptional assembly of a nuclear body by noncoding RNAs. Nat Cell Biol. 2013;15:570–83.

102. Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mijatova S, Gill J, Walsh MJ, Zhou MM. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Mol Cell. 2010;38:662–74.

103. Rapicavoli NA, Ku Q, Zhang J, Mikhail M, Laberge RM, Chang HY. A mammalian pseudogene IncRNA at the interface of inflammation and anti-inflammatory therapies. Elife. 2013;2:e00762.

104. Leucci E, Vendramin R, Spinazzi M, Laurette P, Piers M, Wouters J, Radaelli E, Eyckerman S, Leonelli C, Vanderheyden K, et al. Melanoma addiction to the long non-coding RNA S4MMSN. Nature. 2016;153:818–22.

105. Trimmerchi T, Bilal E, Nitzscharistos P, Fabbri G, Dalla-Favera R, Tsirigos A, Affinis I. Genome-wide mapping and characterization of Notch-regulated long non-coding RNAs in acute leukemia. Cell. 2014;158:599–606.

106. Sun L, Goft LA, Trapnell C, Alexander R, Lo KA, Haci sulieyman E, Sauvageau M, Tazon-Vega B, Kelley DR, Hendrickson DG, et al. Long non-coding RNAs regulate adipogenesis. Proc Natl Acad Sci U S A. 2013;110:33837–92.

107. Hacisuleyman E, Goff LA, Trapnell C, Alexander R, Lo KA, Haci sulieyman E, Sauvageau M, Tazon-Vega B, Kelley DR, Hendrickson DG, et al. Long non-coding RNAs regulate adipogenesis. Proc Natl Acad Sci U S A. 2013;110:33837–92.