SURVEY AND SUMMARY

Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs

Victoria A. Moran[1,2,3], Ranjan J. Perera[4] and Ahmad M. Khalil[1,2,3,5,*]

[1]Center for RNA Molecular Biology, [2]Department of Genetics and Genome Sciences, [3]Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, OH, [4]Sanford-Burnham Medical Research Institute, Orlando, FL and [5]Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA

ABSTRACT

The recent discovery that the human and other mammalian genomes produce thousands of long non-coding RNAs (lncRNAs) raises many fascinating questions. These mRNA-like molecules, which lack significant protein-coding capacity, have been implicated in a wide range of biological functions through diverse and as yet poorly understood molecular mechanisms. Despite some recent insights into how lncRNAs function in such diverse cellular processes as regulation of gene expression and assembly of cellular structures, by and large, the key questions regarding lncRNA mechanisms remain to be answered. In this review, we discuss recent advances in understanding the biology of lncRNAs and propose avenues of investigation that may lead to fundamental new insights into their functions and mechanisms of action. Finally, as numerous lncRNAs are dysregulated in human diseases and disorders, we also discuss potential roles for these molecules in human health.

INTRODUCTION

Recent advances in technologies, such as tiling arrays and RNA deep sequencing (RNA-seq), have made it possible to survey the transcriptomes of many organisms to an unprecedented degree. Several studies utilizing these technologies have unequivocally demonstrated that the genomes of mammals, as well as other organisms, produce thousands of long transcripts that have no significant protein-coding capacity and thus are referred to as long (or large) non-coding RNAs (lncRNAs) (1–6). LncRNAs are strikingly similar to mRNAs: they are RNA polymerase II transcripts that are capped, spliced and polyadenylated, yet do not function as templates for protein synthesis (7).

Although a functional lncRNA known as Xist was discovered and characterized in the early 1990s (8–10), the prevailing view until recently was that such transcripts are rare and only a handful of functional lncRNAs are represented in the genome. However, numerous publications in the past several years have now documented important functions for lncRNAs, affecting many biological processes, including regulation of gene expression, dosage compensation, genomic imprinting, nuclear organization and compartmentalization, and nuclear-cytoplasmic trafficking (7,11–14). It is very likely that additional functions for lncRNAs will be discovered, as only a small percentage of lncRNAs have been studied in detail to date. Furthermore, there are a number of studies that have shown many lncRNAs are dysregulated in various human diseases and disorders, although it is not yet clear if these lncRNAs are causal or symptomatic of the disease state (15,16).

In this review, we will discuss a number of important topics regarding lncRNAs: (i) How many functional lncRNAs are transcribed in mammals?; (ii) Known biological functions of lncRNAs; (iii) How do lncRNAs exert their effects? And finally, (iv) What are the potential roles of lncRNAs in human disease? Although in this review we will focus on mammalian lncRNAs, it is important to point out that lncRNAs are being actively investigated in many other organisms (17,18).

HOW MANY FUNCTIONAL LncRNAs ARE PRESENT IN MAMMALS?

Prior to advances in technologies that made it possible to survey transcriptomes in an unbiased manner and to a much greater depth than previously possible, lncRNAs were discovered and characterized using traditional gene

cloning methods. Initially, these transcripts were thought to code for proteins, but subsequent experimental and bioinformatic data indicated that these transcripts lack long open reading frames (ORFs). The few lncRNAs known prior to the past decade were thought to arise sporadically in the genome (10,19–21). This picture changed radically when, in the early 2000s, the FANTOM consortium examined over 60 000 full-length cDNAs and identified over 11 000 lncRNAs in mouse (22). A significant proportion of these transcripts is overlapping with, and is transcribed in the antisense direction, to protein-coding genes. They are thus referred to as natural antisense transcripts (NATs). Another independent study found that ~40% of protein-coding genes in human cells also express NATs (23). To date, numerous studies demonstrated that NATs regulate their overlapping protein-coding partners *in cis*, either concordantly or discordantly (6,24,25). Furthermore, a number of studies have reported lncRNAs that are expressed solely from the introns of protein-coding genes (26,27).

More recently, it was reported that 'intergenic' regions of the genome, which were previously thought to be gene 'deserts' or 'junk' DNA, also express thousands of long non-coding RNAs, termed large intervening non-coding RNAs (lincRNAs) (1,3–5). Prior to advances in RNA sequencing (RNA-seq) technologies, lincRNAs were discovered by using a chromatin signature of actively transcribed genes (4,5). Essentially, actively transcribed protein-coding genes typically display a specific histone modification pattern: H3K4 trimethylation in the promoter region and H3K36 trimethylation in the body of the gene (28,29). By examining these chromatin marks genome-wide and eliminating those corresponding to protein-coding genes and microRNAs, it was shown that the human and mouse genomes produce over 3300

lincRNAs (4,5). Subsequently, RNA-seq experiments confirmed these observations and revealed an additional 5000 lincRNAs (1,30). It is now estimated that the human genome produces over 8000 lincRNAs, with 4500 of these considered to be high-confidence lincRNAs. These lincRNAs are multi-exonic, capped, polyadenylated and localized to the nucleus, cytoplasm or both. Intriguingly, many of these lincRNAs show tissue-specific expression patterns suggesting potential roles in cell identity (1,4,5).

The exact number of distinct human lncRNAs, which consist of NATs, lincRNAs and intronic lncRNAs (Figure 1), is still a matter of intense debate. However, based on recent publications utilizing the most advanced sequencing platforms and algorithms to assemble transcripts from deep RNA-sequencing reads, the number of total lncRNAs (lincRNAs + NATs + intronic lncRNAs) is in the range of ~20 000 transcripts. Nonetheless, whether the final number will be larger or smaller, it is now clear that the human and other mammalian genomes encode thousands of lncRNAs, yet neither the biological processes in which most of these molecules function, nor their mechanisms of action, have been determined (7).

A crucial outstanding question is whether all lncRNAs are functional. Some have argued that the mere transcription of lncRNAs does not indicate functional significance, especially given that many of these transcripts do not appear to be conserved even between closely related species (1). In this view, establishing the functional significance of any individual lncRNA would require experimental evidence. To date, over 200 lncRNAs have been studied functionally and/or mechanistically. Although the depth of analysis varies considerably, many of these lncRNAs show evidence of functionality, at least *in vitro*. In contrast, only a few lncRNAs have been studied in animal models, with results suggesting that these few
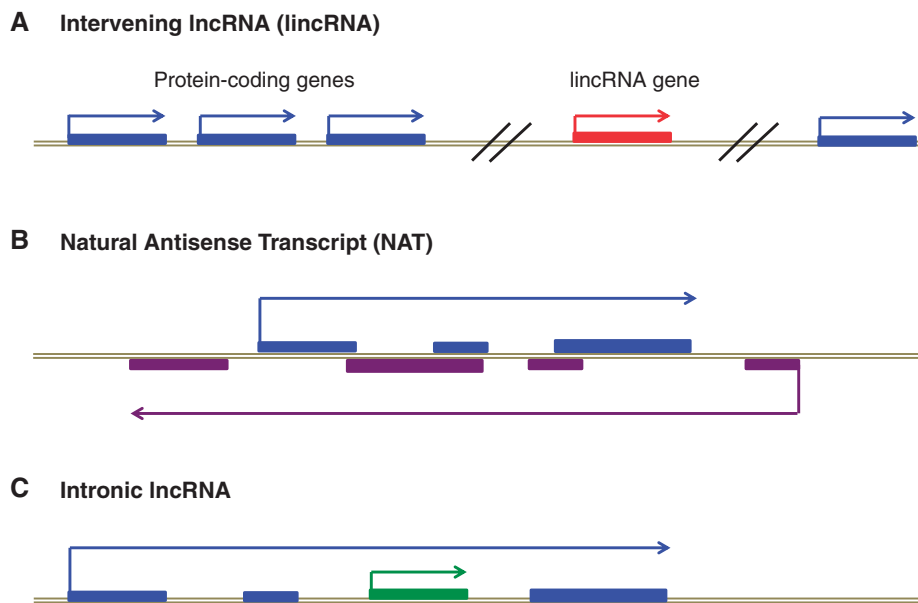

Figure 1. LncRNAs fall into one of three categories. (**A**) Long intervening non-coding RNAs (lincRNAs) are transcribed from regions far away from protein-coding genes. (**B**) Natural Antisense Transcripts (NATs) are transcribed from the opposite strand of a protein-coding gene. (**C**) Intronic lncRNAs (shown in green) are transcribed from within introns of protein-coding genes.

lncRNAs are not essential to the viability of the organism (31). One intriguing hypothesis is that many lncRNAs are genetically redundant and therefore loss of one lncRNA is easily compensated for by one or more other lncRNAs. It is also critical to point out that mouse knockouts of many protein-coding genes are also viable and display no obvious phenotypes (32).

In summary, it is now clear that the human and other mammalian genomes produce thousands of lncRNAs (33). Due to the complexity of the lncRNA population, it will be many years before we can elucidate the functions of all lncRNAs, particularly as many have multiple alternatively spliced forms. In an effort to expedite progress toward this important goal, some researchers are utilizing bioinformatic strategies such as 'guilt by association' or high throughput approaches such as siRNA/shRNA screens to home in on candidate lncRNAs for functional and mechanistic studies (4,5,34–36). As elaborated in the next section, it is very likely that lncRNAs are involved in a wide range of biological functions as they have been implicated in key biological processes thus far.

KNOWN BIOLOGICAL FUNCTIONS OF LNCRNAS

Although only a very small percentage of all lncRNAs identified to date have been examined experimentally, an emerging paradigm suggests that lncRNAs function in many biological contexts. Thus far, lncRNAs have been implicated in such diverse processes as regulation of gene expression both *in cis* and *in trans*, guidance of chromatin-modifying complexes, X chromosome inactivation (Xi) and genomic imprinting, nuclear compartmentalization, nuclear-cytoplasmic trafficking, RNA splicing and translational control (11–14). In the following sections, we will discuss the evidence supporting the established roles in which lncRNAs function.

Regulation of gene expression by lncRNAs both *in cis* and *in trans*

Regulation of gene expression is a complex process that typically requires many factors and co-factors that either open up chromatin and make DNA accessible to RNA polymerases, or conversely, lead to chromatin condensation and DNA inaccessibility. Recent studies have clearly shown that a number of lncRNAs also contribute to gene regulation by various mechanisms (11).

One of the best-studied lncRNAs to date is Xist, which is responsible for the initiation and spreading of X chromosome inactivation (Xi) in female somatic cells. Xist was discovered in 1991 and, despite having been studied extensively, the exact mechanism of Xist-mediated X chromosome inactivation is yet to be fully elucidated (37,38). Nonetheless, it is widely accepted that Xist is required for the silencing of hundreds of genes on the inactive X chromosome. It is thought that a small repeat region within Xist, which is referred to as RepA, is initially transcribed from both X chromosomes along with Xist's antisense partner Tsix (39). Tsix, also a lncRNA, prevents RepA from binding to either X chromosome. However, post-cellular differentiation, RepA, in association with

the chromatin-modifying complex PRC2 (polycomb repressive complex 2), binds to one of the two X chromosomes at the X inactivation center to initiate X inactivation (39). Subsequently, production of full-length Xist from the X chromosome destined for inactivation, which also binds to PRC2, leads to the spreading of X inactivation from the X inactivation center to the entire X chromosome *in cis* (39). On the active X chromosome, Tsix is believed to prevent Xist transcription and thus maintains an active chromatin state. It is currently not known what prevents Xist from 'escaping' the inactive X and acting on the active X or other chromosomes, *in trans*, or how Xist is prevented from silencing genes that escape X inactivation; as many as 20% of genes escape Xi in human females (40,41). Finally, it is worth noting that, in addition to Xist and Tsix, several other lncRNAs are transcribed from the X inactivation center and these also appear to play roles in the counting and choice in Xi (38,42). Once Xi is established, the inactive X chromosome is condensed into facultative heterochromatin and appears as a condensed round body (Barr Body), usually at the periphery of nuclei (43). The inactive X chromosome, in contrast to the active X and the autosomes, is marked with repressive chromatin marks and DNA methylation at CpG islands, with the exception of regions that escape Xi (41,44,45).

Another epigenetic phenomenon that, similar to Xi, also utilizes lncRNAs to regulate gene expression is known as genomic imprinting (46). The expression of imprinted genes depends on their parental origin, and the level of differential expression of the two alleles of an imprinted gene can vary from one imprinted gene to another. Since imprinted genes play critical roles in mammalian development, their expression must be tightly regulated (47). Intriguingly, many imprinted gene loci express, in addition to mRNAs, a significant number of lncRNAs that appear to play major roles in regulating the expression of neighboring imprinted protein-coding genes *in cis* (48). One such lncRNA is *Air*, which is monoallelically expressed from the paternal allele, associates with the histone methyltransferase G9a and localizes to chromatin to silence three imprinted genes known as Slc22a3, Slc22a2 and Igf2r *in cis* (49). Ablation of Air results in the biallelic expression of Slc22a3 and loss of G9a recruitment to the Slc22a3 promoter, suggesting that Air plays a role in guiding the methyltransferase G9a to chromatin at the Slc22a3 promoter (49). In a similar mechanism to Air, the lncRNA Kcnq1ot1 also regulates the expression of imprinted genes in a lineage-specific manner by directing G9a, PRC2 and Dnmt1 to the Kcnq1 locus *in cis* (50,51). Collectively, these observations and others suggest that lncRNAs are intimately involved in regulating the expression of imprinted genes, potentially in a complex and layered manner.

In contrast to Air and Kcnq1ot1, which regulate gene expression *in cis*, an elegant study from Rinn *et al.* (52) led to the discovery of an intervening lncRNA (lincRNA), HOTAIR, that regulates human HOXD genes expression *in trans*. Subsequent studies demonstrated that HOTAIR regulates gene expression *in trans* on a genome-wide scale by associating with the chromatin-modifying complexes PRC2, LSD1 and CoREST/REST (4,52–54). HOTAIR guides as well as serves as a scaffold for PRC2 and

LSD1/CoREST complexes at their endogenous target genes (54). Subsequently, the LSD1/CoREST complex demethylates histone H3 at lysine 4 (H3K4), while PRC2 methylates histone H3 at lysine 27 (H3K27) leading to the loss of an activating histone mark (i.e. H3K4 dimethylation) and the gain of a repressive histone mark (H3K27 trimethylation) at genes targeted by HOTAIR (54). In contrast to the human HOTAIR, mouse Hotair does not appear to regulate Hoxd genes suggesting that HOTAIR may have distinct functions in human and mouse (31).

In addition to the lncRNAs discussed earlier, several other human and mouse lncRNAs (e.g. Tug1, linc-P21, PANDA, Evf-2 and others) have also been shown to regulate gene expression by guiding their protein partners to specific genomic loci (4,34,36,55,56). Collectively, these studies demonstrate that lncRNAs regulate gene expression *in trans* as well as *in cis*. However, the molecular mechanisms that determine a lncRNA's ability to regulate gene expression *in cis* versus *in trans* are not currently known. It is likely that the sequences and/or the secondary structures of lncRNAs inherently affect their mechanisms of action, an area of lncRNA research that remains in its infancy due to lack of appropriate tools.

A subset of lncRNAs is required for maintaining pluripotency

The ability of a stem cell to give rise to all three germ layers (endoderm, mesoderm and ectoderm) is referred to as pluripotency. Previous studies have identified key transcription and chromatin remodeling factors that are required for maintaining pluripotency (57,58). Intriguingly, some of these transcription factors were found to bind to the promoters of over 100 lincRNAs in mouse ES cells (5). Subsequently, another study found that two lncRNAs, which are regulated by the transcription factors Oct4 and Nanog, to be essential for maintaining pluripotency (59). Furthermore, the authors found that knockdown and overexpression of these two lncRNAs results in dramatic changes in Oct4 and Nanog mRNA levels, pointing to the involvement of a feedback loop in the regulatory mechanism (59).

Recently, the role of lincRNAs in pluripotency was examined in a high throughput manner using shRNAs in mES cells (34). The authors demonstrated that 26 lincRNAs are required for the maintenance of pluripotency, as knockdown of each lincRNA led to either an exit from the pluripotent state or activation of lineage commitment programs. Since some of these lincRNAs interact with chromatin-modifying complexes, they could potentially guide and/or serve as scaffolds for such complexes at specific gene loci that are critical for maintaining the pluripotent state. It is also possible that some lincRNAs function as 'environmental sensors' that alert stem cells to maintain pluripotency or to differentiate depending on changes in the environment. Indeed, a recent study identified a lncRNA termed ANCR (Anti-differentiation non-coding RNA) that is required for maintaining cells in an undifferentiated state in the epidermis (60). Future studies are needed to dissect the exact roles of lncRNAs in maintaining pluripotency and promoting cellular differentiation programs.

Nuclear organization: Paraspeckle formation requires the lncRNA NEAT1

Regulation of gene expression is a complex process and can take place at the transcriptional, co-transcriptional, post-transcriptional and translational levels. Recently it was suggested that nuclear structures known as paraspeckles may regulate gene expression post-transcriptionally by the retention of hyperedited mRNAs in the nucleus (61–64). Paraspeckles are identified cytologically based on the presence of a few protein components, which include PSP1, PSP2 and p54 (63). Although the function of PSP1 is not known, PSP2 is involved in splicing, while P54 is involved in splicing as well as several other processes including transcriptional regulation, DNA unwinding and nuclear retention of hyperedited dsRNA (61–63,65). Typically, each nucleus contains several paraspeckles, which are dynamic structures and require ongoing RNA polymerase II transcription (65). During telophase, when transcription is shut down, and in cells treated with drugs that inhibit RNA polymerase II, paraspeckles disappear (61–63).

Intriguingly, the formation and maintenance of paraspeckles require the lncRNA NEAT1 (nuclear-enriched autosomal transcript 1), which localizes exclusively to paraspeckles (61,65–67). Paraspeckles are dynamic structures that are absent in embryonic stem cells but appear after differentiation coincident with NEAT1 activation (61–63). Moreover, depletion of NEAT1 is sufficient to cause loss of paraspeckles in the nucleus, and overexpressing NEAT1, but not paraspeckle-associated proteins, leads to an increase in the number of paraspeckles suggesting an essential role for NEAT1 in paraspeckle formation and/or integrity (61,62,65). Although the mechanism of NEAT1 mediated paraspeckle formation is not completely clear, it is possible that NEAT1 serves as a scaffold for proteins involved in paraspeckle formation, and therefore loss of NEAT1 prevents the ability of these proteins to co-localize. It would not be surprising if future research uncovers other lncRNAs that are involved in forming or maintaining the integrity of other nuclear and/or cytoplasmic structures.

Regulation of alternative splicing by lncRNAs

Alternative splicing of pre-mRNAs increases the proteomic complexity of cells by resulting, in many cases, in several protein products with non-overlapping functions from a single mRNA. The lncRNA MALAT1 (metastasis-associated lung adenocarcinoma transcript 1), which was initially identified in a screen for genes associated with metastasis, plays a critical role in pre-mRNA alternative splicing (68). MALAT1 localizes to nuclear speckles, which contain several proteins that are known to be involved in alternative splicing (68). Furthermore, depletion of MALAT1 is sufficient to alter the alternative splicing patterns of a subset of mRNAs (68). It appears that MALAT1 forms a molecular scaffold for several proteins present in nuclear speckles, as well as modulates the phosphorylation of SR proteins (68). SR proteins are

serine/arginine-rich proteins which are involved in the regulation and selection of the splice sites in pre-mRNAs. By regulating the phosphorylation of SR proteins, MALAT1 may thus regulate the cellular levels of active SR proteins and subsequently the splicing of many pre-mRNAs (68). Finally, it remains to be determined if other lncRNAs may also play a role in alternative splicing. Performing biochemical assays such as RNA co-immunoprecipitation followed by deep sequencing of RNAs (RIP-seq) on proteins that are known to be involved in alternative splicing may lead to the identification of other lncRNAs with functions similar to MALAT1.

HOW DO lncRNAs EXERT THEIR EFFECTS?

Thus far lncRNAs have been implicated in numerous biological functions and pathways, and their mechanisms of actions are very diverse (Figure 2) (7,11). Here we will discuss several mechanisms by which lncRNAs exert their effects, although it is worth noting that in every case much remains to be learned about the detailed mechanism of action.

lncRNAs 'guide' chromatin-modifying complexes to specific genomic loci *in cis* and *in trans*

Cellular identity in an organism is determined by epigenetic factors that modulate specific gene expression programs (69). These epigenetic factors, such as chromatin-modifying complexes and DNA methyltransferases, activate and repress specific genes by enzymatically modifying chromatin and DNA (70). One of the most puzzling questions in biology is how do these ubiquitous enzymes, which lack DNA binding capacity, recognize their target genes in the various cell types. Emerging evidence suggest that some lncRNAs 'guide' chromatin-modifying complexes as well as other nuclear proteins to specific genomic loci to exert their effects (14,71,72). In essence, some lncRNAs may function as 'GPS' devices to target other cellular components to their sites of action. Below, we will highlight several examples of lncRNAs that have been shown to possess such activity.

As discussed previously, the lncRNA HOTAIR directs the chromatin-modifying complexes PRC2 and LSD1 to numerous gene loci on a genome-wide scale *in trans* (4,52–54). By contrast, other lncRNAs, such as
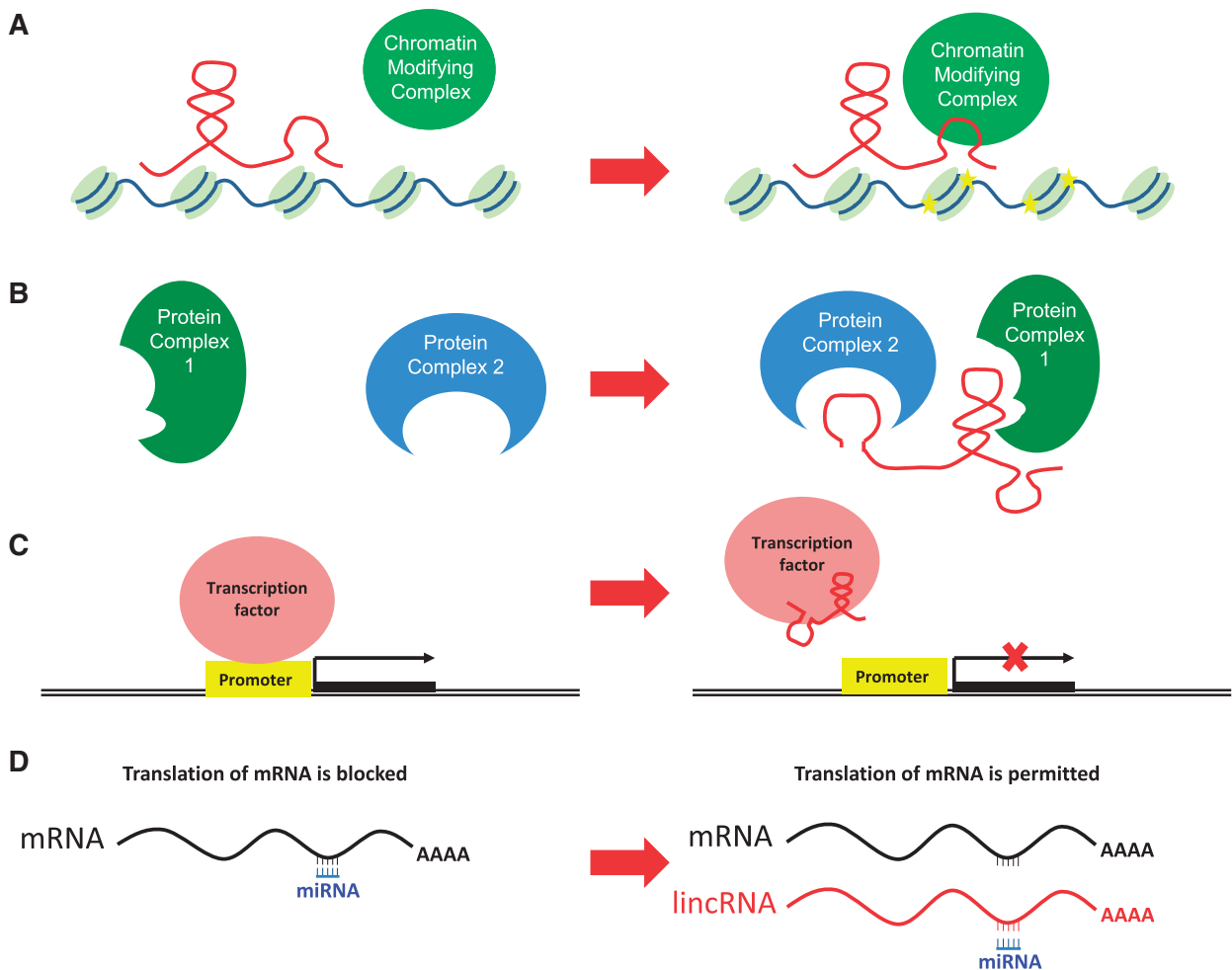

Figure 2. LncRNAs exert their effects by diverse mechanisms. (**A**) lncRNAs can act as guides and tethers for chromatin-modifying complexes, and thus contribute to tissue-specific gene expression. (**B**) lncRNAs can act as molecular scaffolds for protein complexes that lack protein–protein interaction domains. (**C**) lncRNAs can bind to transcription factors and prevent them from binding to their target DNA sequence. (**D**) lncRNAs can interact directly with microRNAs (miRNAs) and prevent them from binding to mRNA, thus regulating protein synthesis.

Air, Kcnq1ot1 and Evf-2, target chromatin-modifying complexes to their target genes *in cis* (39,49,50,56). Currently, the sequence of events that leads to a lncRNA-mediated guidance of a protein complex to chromatin have not been fully elucidated. However, recent evidence suggests that lncRNAs may bind to chromatin first, and then serve as docking stations for chromatin-modifying complexes. In support of this model, a recent study, which utilized a novel technology that allows the genomic occupancy of a lncRNA to be determined, found that HOTAIR localizes to chromatin independent of its protein-binding partner PRC2 (73).

Since numerous lncRNAs are known to bind to chromatin-modifying complexes (4,74), it is conceivable that many of these lncRNAs also function by 'guiding' their protein partners to chromatin. To fully understand this mechanism of action, a number of key questions must be addressed: (i) Are there specific motifs in lncRNAs that are responsible for the targeting mechanism to specific genomic regions? (ii) How do proteins recognize and specifically bind to certain lncRNA(s) but not others? (iii) Do these lncRNAs directly interact with DNA to form lncRNA:DNA hybrids or triplexes or (iv) do DNA binding proteins serve as intermediates between a lncRNA and DNA? These key questions will be critical to answer in order to fully understand the mechanisms through which chromatin-associated lncRNAs function.

lncRNAs serve as structural links in ribonucleoprotein complexes (RNPs)

While it is clear that the majority of RNA does not float around in the cell naked but rather is complexed with protein, the precise molecular composition of most ribonucleoprotein complexes (RNPs) has not been determined. Recent studies have shown that many lncRNAs exist in the cell as part of RNPs (71). Thus far, lncRNAs have been shown to form RNPs with chromatin-modifying complexes, transcription factors, splicing factors, as well as other classes of proteins (71). For example, as discussed earlier, HOTAIR forms a RNP with the chromatin-modifying complexes PRC2 and LSD1/CoREST (54). Mutational analysis of HOTAIR revealed that HOTAIR integrity is required for PRC2 interaction with LSD1/CoREST (54). This suggests that, in addition to HOTAIR's role in guiding PRC2 and LSD1 to chromatin, HOTAIR also serves as a structural bridge between PRC2 and LSD1/CoREST at endogenous target genes. Also, Xist has recently been shown to interact with the transcription factor YY1, which helps tether Xist, and consequently PRC2, to the inactive X chromosome (75). In essence, Xist appears to be forming a molecular bridge between YY1 and PRC2 to repress genes on the inactive X chromosome.

The lncRNAs NEAT1 and MALAT1 form molecular scaffolds for several proteins that are core components of paraspeckles and speckles, respectively (61–63,68). It is not currently known how lncRNAs recognize their protein partners and initiate the formation of these nuclear compartments. It is possible that secondary structures of lncRNAs contribute to the specificity of lncRNAs interactions with their protein partners. Mutation analysis

of structural lncRNAs may lead to the identification of lncRNA 'domains' that are critical for their functions as molecular scaffolds. In summary, lncRNAs appear to provide an extensive infrastructure within the nucleus, and potentially in the cytoplasm, that makes it possible for various proteins to co-localize and coordinate their functions to accomplish a specific biological function (71). Disruption of such lncRNAs may lead to undesired biological consequences (76).

LncRNAs regulate distinct transcriptional programs

A few lncRNAs have been shown to be activated in response to specific stimuli, and subsequently activate specific transcriptional programs that allow the cell to respond to these stimuli. For example, lncRNAs, such as linc-P21, PANDA, Tug1 and others, are transcriptionally activated in response to DNA damage by the direct binding of the tumor-suppressor protein p53 to their promoters (4,5,36,55). Subsequently, these lncRNAs regulate gene expression by distinct pathways. Linc-P21, which represses numerous genes in the p53 pathway, requires the RNA/DNA binding protein hnRNP K and other as yet unidentified factors (36). By contrast, the lncRNA PANDA, which is also activated by p53, requires the downstream effector NF-YA to regulate the expression of pro-apoptotic genes (55), while Tug1 functions via its interaction with the chromatin-modifying complex PRC2 (4).

The lncRNA Gas5 (growth arrest specific 5), which is highly expressed in cells that have arrested growth, serves as a negative regulator of glucocorticoid receptors (GR), a specific class of nuclear receptors (77). Gas5 interacts directly with the DNA binding domain of GRs, preventing them from binding to their DNA response elements, thereby in effect acting as a molecular decoy (77). The ability of a lncRNA to modulate the effects of a transcription factor can lead, in some cases, to significant changes in gene expression and subsequently profound effects on the cells ability to respond to external stimuli. Further studies are needed to determine the underlying mechanisms of how a lncRNA, once activated, modulates the activity of transcription factor(s) to allow the cells to respond to their environment.

Regulation of microRNAs by lncRNAs

In 2007, a study in *Arabidopsis thaliana* found the non-coding RNA IPS1 to bind to the microRNA miR-399 and block its ability to regulate PHO2 mRNA (78). Recently, evidence came to light suggesting that some mammalian lncRNAs may also regulate gene expression post-transcriptionally by binding to miRNAs, and consequently preventing specific miRNAs from binding to their target mRNAs. In a 2011 published study, it was demonstrated that a lincRNA, linc-MD1, serves as a 'sponge' for two miRNAs, which regulate transcription factors involved in muscle differentiation (79). These findings are very intriguing since they demonstrate that distinct classes of non-coding RNAs cooperate to regulate gene expression. It is conceivable that other lncRNAs can also serve as 'sponges' for miRNAs in a tissue and developmental stage-specific manner.

However, it is not clear how cells regulate the expression levels of lncRNAs and miRNAs, and how lncRNAs receive a signal to bind or not to bind a miRNA. Potentially, novel protein partners of lncRNAs may be critical for regulating lncRNAs interactions with miRNAs in a spatial and temporal manner. There are many key questions, which are yet to be determined regarding this mechanism of lncRNA-mediated regulation, that will require tremendous effort to answer but are key to understanding this mode of action.

WHAT ARE THE POTENTIAL ROLES OF LncRNAs IN HUMAN DISEASE?

In contrast to the extensive evidence that links dysregulation of protein-coding genes to disease etiology, to date only a few lncRNAs have been implicated in human disease (76). However, we are beginning to observe, in some cases, strong associations between lncRNAs and human disease, and it is reasonable to expect that a concrete, mechanistic understanding of these connections will emerge in the coming years. Since both the role of lncRNAs dysregulation in human disease and their molecular mechanisms remain unclear, it is timely to ask important questions such as: (i) How many lncRNAs are differentially regulated in a given human disease compared to healthy counterparts? (ii) What are the molecular and biological functions of lncRNAs that are dysregulated in human disease? (iii) Do disease-specific lncRNAs change their subcellular localization? (iv) How stable are lncRNAs and is their stability altered in various disease states? Thus, an immediate goal of lncRNA research is to determine whether lncRNAs are useful signatures for early disease detection, or can be used as candidate drug targets for disease intervention (80).

LncRNAs have been found to be dysregulated in a wide range of human diseases and disorders, including various types of cancers. This includes breast cancer (53), colorectal cancer (81), prostate cancer (82), hepatocellular carcinoma (83–85), leukemia (86,87), melanoma (88) and possibly others (76,89). On a more mechanistic level, recent studies have revealed the contribution of lncRNAs as proto-oncogenes, e.g. *GAGE6* (90), as tumor suppressor genes, e.g. 'p15 antisense RNA and lincP21' (36,91), as drivers of metastatic transformation, e.g. *HOTAIR* in breast cancer (53), and as regulators of alternative splicing, e.g. MALAT1 (68).

Although the entire mechanisms of action of most lncRNAs that are dysregulated in cancer have not been fully elucidated, a number of studies have provided some insights into the mechanisms of few such lncRNAs. For example, the lncRNA linc-P21 functions as a repressor in the p53 pathway by directing the RNA/DNA binding protein hnRNPK to chromatin (36). The lncRNA, *SPRY4-IT1,* which is up-regulated in human melanomas compared to melanocytes and keratinocytes, affects cell dynamics, including increased rate of wound closure upon ectopic expression. This suggests that the higher expression of *SPRY4-IT1* may have an important role in the molecular etiology of human melanomas (88).

Aside from their roles in cancer, lncRNAs are known to be dysregulated in several other diseases, including heart disease (92,93), Alzheimer's disease (94), psoriasis (95), spinocerebellar ataxia type 8 (96) and fragile X syndrome (97). Finally, we refer the reader to an excellent review article that covers the role of lncRNAs in human disease in depth (76).

CONCLUDING REMARKS

The discovery of thousands of lncRNAs has certainly changed our view of the complexity of mammalian genomes and transcriptomes, as well as many other aspects of biology including transcriptional and post-transcriptional regulation of gene expression. To date, we have only scratched the surface in terms of elucidating the functions and mechanisms of lncRNAs; nonetheless, recent studies suggest that lncRNAs are likely to exert their effects by diverse mechanisms (7,11,13,14,16,71,72,76). In many cases lncRNAs have been shown to work cooperatively with proteins by forming ribonucleoprotein complexes (RNPs) (71). These RNPs appear to depend on their lncRNA constituent(s) for proper localization to specific regions within the cell, as well as to coordinate the interactions between protein complexes that do not have interacting domains (4,52–54,68,98,99). Furthermore, lncRNAs appear to be intimately associated with both chromatin and chromatin-modifying complexes suggesting that lncRNAs are critical for genome organization and regulation of gene expression at the transcriptional levels (11,14,72,73). Lastly, since lncRNAs show strong tissue-specific expression, they are likely to play a major role in cell identity and spatial organization in multicellular organisms (1,4). We envision that lncRNAs work within complex networks involving proteins and other types of ncRNAs (e.g. miRNAs) to achieve a system level of cellular organization.

The studies discussed earlier and others clearly demonstrate that lncRNAs, or at least the ones that have been examined to date, are functional, despite lack of conservation in many cases even among closely related species. This conundrum could be explained by the hypothesis that lncRNAs rely on 'conserved' secondary structures and not their primary sequence to perform their functions. This hypothesis would help explain how many lncRNAs with distinct sequences bind the same protein complex (4,34,71); as well as how lncRNAs have functional orthologs that are not conserved at the sequence level (100). Currently, there are no reliable methods to determine the secondary structures of lncRNAs, and thus, the above hypothesis remains to be experimentally tested. However, an emerging technology that is referred to as parallel analysis of RNA structure (PARS), which is based on deep sequencing of RNA fragments that have been treated with structure-specific enzymes, has shown promise in predicting secondary structure of RNAs in yeast (101). This technology can potentially be modified to perform similar analysis in higher eukaryotes. Finally, it can not be excluded that a small percentage of lncRNAs encodes short peptides (102).

Dysregulation of lncRNAs, which has been observed in numerous human diseases, suggest that lncRNAs can be utilized in medicine as biomarkers and/or drug targets (76). Such therapies would be useful in cases where drugs designed to target proteins have failed, or even in conjunction with available drugs in order to enhance their effects. For example, a previous study found that knocking down a lncRNA enhances the effects of chemotherapeutic drugs *in vitro* (103).

In conclusion, the potential roles of lncRNAs in biology and medicine could be tremendous, and will require many years of intensive research before they can be fully deciphered and applied. However, if recent publications are an indication of the progress in this area of scientific research, then this field is certainly moving at a supersonic speed.

REFERENCES

1. Cabili,M.N., Trapnell,C., Goff,L., Koziol,M., Tazon-Vega,B., Regev,A. and Rinn,J.L. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.*, **25**, 1915–1927.
2. Orom,U.A., Derrien,T., Beringer,M., Gumireddy,K., Gardini,A., Bussotti,G., Lai,F., Zytnicki,M., Notredame,C., Huang,Q. *et al.* (2010) Long noncoding RNAs with enhancer-like function in human cells. *Cell*, **143**, 46–58.
3. Ponting,C.P., Oliver,P.L. and Reik,W. (2009) Evolution and functions of long noncoding RNAs. *Cell*, **136**, 629–641.
4. Khalil,A.M., Guttman,M., Huarte,M., Garber,M., Raj,A., Rivea Morales,D., Thomas,K., Presser,A., Bernstein,B.E., van Oudenaarden,A. *et al.* (2009) Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl Acad. Sci. USA*, **106**, 11667–11672.
5. Guttman,M., Amit,I., Garber,M., French,C., Lin,M.F., Feldser,D., Huarte,M., Zuk,O., Carey,B.W., Cassady,J.P. *et al.* (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*, **458**, 223–227.
6. Carninci,P., Kasukawa,T., Katayama,S., Gough,J., Frith,M.C., Maeda,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C. *et al.* (2005) The transcriptional landscape of the mammalian genome. *Science*, **309**, 1559–1563.
7. Mercer,T.R., Dinger,M.E. and Mattick,J.S. (2009) Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.*, **10**, 155–159.
8. Brown,C.J., Hendrich,B.D., Rupert,J.L., Lafreniere,R.G., Xing,Y., Lawrence,J. and Willard,H.F. (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.
9. Brockdorff,N., Ashworth,A., Kay,G.F., McCabe,V.M., Norris,D.P., Cooper,P.J., Swift,S. and 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.
10. Brown,C.J., Ballabio,A., Rupert,J.L., Lafreniere,R.G., Grompe,M., Tonlorenzi,R. and Willard,H.F. (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature*, **349**, 38–44.
11. Wang,K.C. and Chang,H.Y. (2011) Molecular mechanisms of long noncoding RNAs. *Mol. Cell*, **43**, 904–914.
12. Nagano,T. and Fraser,P. (2011) No-nonsense functions for long noncoding RNAs. *Cell*, **145**, 178–181.
13. Clark,M.B. and Mattick,J.S. (2011) Long noncoding RNAs in cell biology. *Semin. Cell Dev. Biol.*, **22**, 366–376.
14. Mattick,J.S., Amaral,P.P., Dinger,M.E., Mercer,T.R. and Mehler,M.F. (2009) RNA regulation of epigenetic processes. *Bioessays*, **31**, 51–59.
15. Wapinski,O. and Chang,H.Y. (2011) Long noncoding RNAs and human disease. *Trends Cell Biol.*, **21**, 354–361.
16. Qureshi,I.A., Mattick,J.S. and Mehler,M.F. (2010) Long non-coding RNAs in nervous system function and disease. *Brain Res.*, **1338**, 20–35.
17. Geisler,S., Lojek,L., Khalil,A.M., Baker,K.E. and Coller,J. (2012) Decapping of long noncoding RNAs regulates inducible genes. *Mol. Cell*, **45**, 279–291.
18. Pauli,A., Valen,E., Lin,M.F., Garber,M., Vastenhouw,N.L., Levin,J.Z., Fan,L., Sandelin,A., Rinn,J.L., Regev,A. *et al.* (2012) Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Res*, **22**, 577–591.
19. Lee,J.T., Davidow,L.S. and Warshawsky,D. (1999) Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat. Genet.*, **21**, 400–404.
20. Brown,C.J., Lafreniere,R.G., Powers,V.E., Sebastio,G., Ballabio,A., Pettigrew,A.L., Ledbetter,D.H., Levy,E., Craig,I.W. and Willard,H.F. (1991) Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature*, **349**, 82–84.
21. Brannan,C.I., Dees,E.C., Ingram,R.S. and Tilghman,S.M. (1990) The product of the H19 gene may function as an RNA. *Mol. Cell. Biol.*, **10**, 28–36.
22. Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H. *et al.* (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature*, **420**, 563–573.
23. He,Y., Vogelstein,B., Velculescu,V.E., Papadopoulos,N. and Kinzler,K.W. (2008) The antisense transcriptomes of human cells. *Science*, **322**, 1855–1857.
24. Katayama,S., Tomaru,Y., Kasukawa,T., Waki,K., Nakanishi,M., Nakamura,M., Nishida,H., Yap,C.C., Suzuki,M., Kawai,J. *et al.* (2005) Antisense transcription in the mammalian transcriptome. *Science*, **309**, 1564–1566.
25. Werner,A., Carlile,M. and Swan,D. (2009) What do natural antisense transcripts regulate? *RNA Biol.*, **6**, 43–48.
26. Kampa,D., Cheng,J., Kapranov,P., Yamanaka,M., Brubaker,S., Cawley,S., Drenkow,J., Piccolboni,A., Bekiranov,S., Helt,G. *et al.* (2004) Novel RNAs identified from an in-depth analysis of the transcriptome of human chromosomes 21 and 22. *Genome Res.*, **14**, 331–342.
27. Rinn,J.L., Euskirchen,G., Bertone,P., Martone,R., Luscombe,N.M., Hartman,S., Harrison,P.M., Nelson,F.K., Miller,P., Gerstein,M. *et al.* (2003) The transcriptional activity of human Chromosome 22. *Genes Dev.*, **17**, 529–540.
28. Mikkelsen,T.S., Ku,M., Jaffe,D.B., Issac,B., Lieberman,E., Giannoukos,G., Alvarez,P., Brockman,W., Kim,T.K., Koche,R.P. *et al.* (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, **448**, 553–560.
29. Barski,A., Cuddapah,S., Cui,K., Roh,T.Y., Schones,D.E., Wang,Z., Wei,G., Chepelev,I. and Zhao,K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823–837.
30. Guttman,M., Garber,M., Levin,J.Z., Donaghey,J., Robinson,J., Adiconis,X., Fan,L., Koziol,M.J., Gnirke,A., Nusbaum,C. *et al.* (2010) Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat. Biotechnol.*, **28**, 503–510.

31. Schorderet,P. and Duboule,D. (2011) Structural and functional differences in the long non-coding RNA hotair in mouse and human. *PLoS Genet.*, **7**, e1002071.

32. Austin,C.P., Battey,J.F., Bradley,A., Bucan,M., Capecchi,M., Collins,F.S., Dove,W.F., Duyk,G., Dymecki,S., Eppig,J.T. *et al.* (2004) The knockout mouse project. *Nat. Genet.*, **36**, 921–924.

33. Amaral,P.P., Dinger,M.E., Mercer,T.R. and Mattick,J.S. (2008) The eukaryotic genome as an RNA machine. *Science*, **319**, 1787–1789.

34. Guttman,M., Donaghey,J., Carey,B.W., Garber,M., Grenier,J.K., Munson,G., Young,G., Lucas,A.B., Ach,R., Bruhn,L. *et al.* (2011) lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature*, **477**, 295–300.

35. Loewer,S., Cabili,M.N., Guttman,M., Loh,Y.H., Thomas,K., Park,I.H., Garber,M., Curran,M., Onder,T., Agarwal,S. *et al.* (2010) Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat. Genet.*, **42**, 1113–1117.

36. Huarte,M., Guttman,M., Feldser,D., Garber,M., Koziol,M.J., Kenzelmann-Broz,D., Khalil,A.M., Zuk,O., Amit,I., Rabani,M. *et al.* (2010) A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*, **142**, 409–419.

37. Pontier,D.B. and Gribnau,J. (2011) Xist regulation and function explored. *Hum. Genet.*, **130**, 223–236.

38. Augui,S., Nora,E.P. and Heard,E. (2011) Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat. Rev. Genet.*, **12**, 429–442.

39. Zhao,J., Sun,B.K., Erwin,J.A., Song,J.J. and Lee,J.T. (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science*, **322**, 750–756.

40. Carrel,L. and Willard,H.F. (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature*, **434**, 400–404.

41. Khalil,A.M. and Driscoll,D.J. (2007) Trimethylation of histone H3 lysine 4 is an epigenetic mark at regions escaping mammalian X inactivation. *Epigenetics*, **2**, 114–118.

42. Tian,D., Sun,S. and Lee,J.T. (2010) The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell*, **143**, 390–403.

43. Barr,M.L. and Bertram,E.G. (1949) A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature*, **163**, 676.

44. Panning,B. (2004) X inactivation in mouse ES cells: histone modifications and FISH. *Methods Enzymol.*, **376**, 419–428.

45. Heard,E., Rougeulle,C., Arnaud,D., Avner,P., Allis,C.D. and Spector,D.L. (2001) Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. *Cell*, **107**, 727–738.

46. Reik,W. and Murrell,A. (2000) Genomic imprinting. Silence across the border. *Nature*, **405**, 408–409.

47. Li,Y. and Sasaki,H. (2011) Genomic imprinting in mammals: its life cycle, molecular mechanisms and reprogramming. *Cell Res.*, **21**, 466–473.

48. Mohammad,F., Mondal,T. and Kanduri,C. (2009) Epigenetics of imprinted long noncoding RNAs. *Epigenetics*, **4**, 277–286.

49. Nagano,T., Mitchell,J.A., Sanz,L.A., Pauler,F.M., Ferguson-Smith,A.C., Feil,R. and Fraser,P. (2008) The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science*, **322**, 1717–1720.

50. Pandey,R.R., Mondal,T., Mohammad,F., Enroth,S., Redrup,L., Komorowski,J., Nagano,T., Mancini-Dinardo,D. and Kanduri,C. (2008) Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell*, **32**, 232–246.

51. Mohammad,F., Mondal,T., Guseva,N., Pandey,G.K. and Kanduri,C. (2010) Kcnq1ot1 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1. *Development*, **137**, 2493–2499.

52. Rinn,J.L., Kertesz,M., Wang,J.K., Squazzo,S.L., Xu,X., Brugmann,S.A., Goodnough,L.H., Helms,J.A., Farnham,P.J., Segal,E. *et al.* (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, **129**, 1311–1323.

53. Gupta,R.A., Shah,N., Wang,K.C., Kim,J., Horlings,H.M., Wong,D.J., Tsai,M.C., Hung,T., Argani,P., Rinn,J.L. *et al.* (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, **464**, 1071–1076.

54. Tsai,M.C., Manor,O., Wan,Y., Mosammaparast,N., Wang,J.K., Lan,F., Shi,Y., Segal,E. and Chang,H.Y. (2010) Long noncoding RNA as modular scaffold of histone modification complexes. *Science*, **329**, 689–693.

55. Hung,T., Wang,Y., Lin,M.F., Koegel,A.K., Kotake,Y., Grant,G.D., Horlings,H.M., Shah,N., Umbricht,C., Wang,P. *et al.* (2011) Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat. Genet.*, **43**, 621–629.

56. Feng,J., Bi,C., Clark,B.S., Mady,R., Shah,P. and Kohtz,J.D. (2006) The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev.*, **20**, 1470–1484.

57. Loh,Y.H., Wu,Q., Chew,J.L., Vega,V.B., Zhang,W., Chen,X., Bourque,G., George,J., Leong,B., Liu,J. *et al.* (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.*, **38**, 431–440.

58. Bernstein,B.E., Mikkelsen,T.S., Xie,X., Kamal,M., Huebert,D.J., Cuff,J., Fry,B., Meissner,A., Wernig,M., Plath,K. *et al.* (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, **125**, 315–326.

59. Sheik Mohamed,J., Gaughwin,P.M., Lim,B., Robson,P. and Lipovich,L. (2010) Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *RNA*, **16**, 324–337.

60. Kretz,M., Webster,D.E., Flockhart,R.J., Lee,C.S., Zehnder,A., Lopez-Pajares,V., Qu,K., Zheng,G.X., Chow,J., Kim,G.E. *et al.* (2012) Suppression of progenitor differentiation requires the long noncoding RNA ANCR. *Genes Dev*, **26**, 338–343.

61. Clemson,C.M., Hutchinson,J.N., Sara,S.A., Ensminger,A.W., Fox,A.H., Chess,A. and Lawrence,J.B. (2009) An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell*, **33**, 717–726.

62. Chen,L.L. and Carmichael,G.G. (2009) Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol. Cell*, **35**, 467–478.

63. Bond,C.S. and Fox,A.H. (2009) Paraspeckles: nuclear bodies built on long noncoding RNA. *J. Cell Biol.*, **186**, 637–644.

64. Prasanth,K.V., Prasanth,S.G., Xuan,Z., Hearn,S., Freier,S.M., Bennett,C.F., Zhang,M.Q. and Spector,D.L. (2005) Regulating gene expression through RNA nuclear retention. *Cell*, **123**, 249–263.

65. Sunwoo,H., Dinger,M.E., Wilusz,J.E., Amaral,P.P., Mattick,J.S. and Spector,D.L. (2009) MEN varepsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res.*, **19**, 347–359.

66. Mao,Y.S., Sunwoo,H., Zhang,B. and Spector,D.L. (2011) Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat. Cell Biol.*, **13**, 95–101.

67. Shevtsov,S.P. and Dundr,M. (2011) Nucleation of nuclear bodies by RNA. *Nat. Cell Biol.*, **13**, 167–173.

68. Tripathi,V., Ellis,J.D., Shen,Z., Song,D.Y., Pan,Q., Watt,A.T., Freier,S.M., Bennett,C.F., Sharma,A., Bubulya,P.A. *et al.* (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell*, **39**, 925–938.

69. Bernstein,B.E., Meissner,A. and Lander,E.S. (2007) The mammalian epigenome. *Cell*, **128**, 669–681.

70. Kouzarides,T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.

71. Khalil,A.M. and Rinn,J.L. (2011) RNA-protein interactions in human health and disease. *Semin. Cell Dev. Biol.*, **22**, 359–365.

72. Koziol,M.J. and Rinn,J.L. (2010) RNA traffic control of chromatin complexes. *Curr. Opin. Genet. Dev.*, **20**, 142–148.

73. Chu,C., Qu,K., Zhong,F.L., Artandi,S.E. and Chang,H.Y. (2011) Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell*, **44**, 667–678.

74. Zhao,J., Ohsumi,T.K., Kung,J.T., Ogawa,Y., Grau,D.J., Sarma,K., Song,J.J., Kingston,R.E., Borowsky,M. and Lee,J.T. (2010) Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol. Cell*, **40**, 939–953.

75. Jeon,Y. and Lee,J.T. (2011) YY1 tethers Xist RNA to the inactive X nucleation center. *Cell*, **146**, 119–133.

76. Taft,R.J., Pang,K.C., Mercer,T.R., Dinger,M. and Mattick,J.S. (2010) Non-coding RNAs: regulators of disease. *J. Pathol.*, **220**, 126–139.

77. Kino,T., Hurt,D.E., Ichijo,T., Nader,N. and Chrousos,G.P. (2010) Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci. Signal.*, **3**, ra8.

78. Franco-Zorrilla,J.M., Valli,A., Todesco,M., Mateos,I., Puga,M.I., Rubio-Somoza,I., Leyva,A., Weigel,D., Garcia,J.A. and Paz-Ares,J. (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet.*, **39**, 1033–1037.

79. Cesana,M., Cacchiarelli,D., Legnini,I., Santini,T., Sthandier,O., Chinappi,M., Tramontano,A. and Bozzoni,I. (2011) A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*, **147**, 358–369.

80. Mattick,J.S. (2009) The genetic signatures of noncoding RNAs. *PLoS Genet.*, **5**, e1000459.

81. Kogo,R., Shimamura,T., Mimori,K., Kawahara,K., Imoto,S., Sudo,T., Tanaka,F., Shibata,K., Suzuki,A., Komune,S. *et al.* (2011) Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res.*, **71**, 6320–6326.

82. Chung,S., Nakagawa,H., Uemura,M., Piao,L., Ashikawa,K., Hosono,N., Takata,R., Akamatsu,S., Kawaguchi,T., Morizono,T. *et al.* (2011) Association of a novel long non-coding RNA in 8q24 with prostate cancer susceptibility. *Cancer Sci.*, **102**, 245–252.

83. Yang,Z., Zhou,L., Wu,L.M., Lai,M.C., Xie,H.Y., Zhang,F. and Zheng,S.S. (2011) Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. *Ann. Surg. Oncol.*, **18**, 1243–1250.

84. Yang,F., Zhang,L., Huo,X.S., Yuan,J.H., Xu,D., Yuan,S.X., Zhu,N., Zhou,W.P., Yang,G.S., Wang,Y.Z. *et al.* (2011) Long noncoding RNA high expression in hepatocellular carcinoma facilitates tumor growth through enhancer of zeste homolog 2 in humans. *Hepatology*, **54**, 1679–1689.

85. Lai,M.C., Yang,Z., Zhou,L., Zhu,Q.Q., Xie,H.Y., Zhang,F., Wu,L.M., Chen,L.M. and Zheng,S.S. (2011) Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation. *Med. Oncol.*, June 16 (doi:10.1007/s12032-011-0004-z; epub ahead of print).

86. Calin,G.A., Pekarsky,Y. and Croce,C.M. (2007) The role of microRNA and other non-coding RNA in the pathogenesis of chronic lymphocytic leukemia. *Best Pract. Res. Clin. Haematol.*, **20**, 425–437.

87. Calin,G.A., Liu,C.G., Ferracin,M., Hyslop,T., Spizzo,R., Sevignani,C., Fabbri,M., Cimmino,A., Lee,E.J., Wojcik,S.E. *et al.* (2007) Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. *Cancer Cell*, **12**, 215–229.

88. Khaitan,D., Dinger,M.E., Mazar,J., Crawford,J., Smith,M.A., Mattick,J.S. and Perera,R.J. (2011) The melanoma-upregulated long noncoding RNA SPRY4-IT1 modulates apoptosis and invasion. *Cancer Res.*, **71**, 3852–3862.

89. Huarte,M. and Rinn,J.L. (2010) Large non-coding RNAs: missing links in cancer? *Hum. Mol. Genet.*, **19**, R152–R161.

90. Li,L., Feng,T., Lian,Y., Zhang,G., Garen,A. and Song,X. (2009) Role of human noncoding RNAs in the control of tumorigenesis. *Proc. Natl Acad. Sci. USA*, **106**, 12956–12961.

91. Yu,W., Gius,D., Onyango,P., Muldoon-Jacobs,K., Karp,J., Feinberg,A.P. and Cui,H. (2008) Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature*, **451**, 202–206.

92. Ishii,N., Ozaki,K., Sato,H., Mizuno,H., Saito,S., Takahashi,A., Miyamoto,Y., Ikegawa,S., Kamatani,N., Hori,M. *et al.* (2006) Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. *J. Hum. Genet.*, **51**, 1087–1099.

93. Pasmant,E., Laurendeau,I., Heron,D., Vidaud,M., Vidaud,D. and Bieche,I. (2007) Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res.*, **67**, 3963–3969.

94. Faghihi,M.A., Modarresi,F., Khalil,A.M., Wood,D.E., Sahagan,B.G., Morgan,T.E., Finch,C.E., St Laurent,G. 3rd, Kenny,P.J. and Wahlestedt,C. (2008) Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat. Med.*, **14**, 723–730.

95. Sonkoly,E., Bata-Csorgo,Z., Pivarcsi,A., Polyanka,H., Kenderessy-Szabo,A., Molnar,G., Szentpali,K., Bari,L., Megyeri,K., Mandi,Y. *et al.* (2005) Identification and characterization of a novel, psoriasis susceptibility-related noncoding RNA gene, PRINS. *J. Biol. Chem.*, **280**, 24159–24167.

96. Daughters,R.S., Tuttle,D.L., Gao,W., Ikeda,Y., Moseley,M.L., Ebner,T.J., Swanson,M.S. and Ranum,L.P. (2009) RNA gain-of-function in spinocerebellar ataxia type 8. *PLoS Genet.*, **5**, e1000600.

97. Khalil,A.M., Faghihi,M.A., Modarresi,F., Brothers,S.P. and Wahlestedt,C. (2008) A novel RNA transcript with antiapoptotic function is silenced in fragile X syndrome. *PLoS One*, **3**, e1486.

98. Wang,K.C., Yang,Y.W., Liu,B., Sanyal,A., Corces-Zimmerman,R., Chen,Y., Lajoie,B.R., Protacio,A., Flynn,R.A., Gupta,R.A. *et al.* (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*, **472**, 120–124.

99. Yang,L., Lin,C., Liu,W., Zhang,J., Ohgi,K.A., Grinstein,J.D., Dorrestein,P.C. and Rosenfeld,M.G. (2011) ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. *Cell*, **147**, 773–788.

100. Ulitsky,I., Shkumatava,A., Jan,C.H., Sive,H. and Bartel,D.P. (2011) Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell*, **147**, 1537–1550.

101. Kertesz,M., Wan,Y., Mazor,E., Rinn,J.L., Nutter,R.C., Chang,H.Y. and Segal,E. (2010) Genome-wide measurement of RNA secondary structure in yeast. *Nature*, **467**, 103–107.

102. Ingolia,N.T., Lareau,L.F. and Weissman,J.S. (2011) Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*, **147**, 789–802.

103. Zhu,Y., Yu,M., Li,Z., Kong,C., Bi,J., Li,J., Gao,Z. and Li,Z. (2011) ncRAN, a newly identified long noncoding RNA, enhances human bladder tumor growth, invasion, and survival. *Urology*, **77**, 510 e511–e515.