Chapter 1
Transcriptional and Posttranscriptional Programming by Long Noncoding RNAs

Radha Raman Pandey and Chandrasekhar Kanduri

Abstract Recently, several lines of evidence have suggested that noncoding RNAs, which include both small and long noncoding RNAs (ncRNAs), contribute to a significant portion of the transcriptome in eukaryotic organisms. However, the functional significance of this wide-spread occurrence of ncRNAs, and in particular, the long ncRNAs (lncRNAs), for organismal development and differentiation is unclear. The available evidence from a subset of lncRNAs suggests that certain lncRNAs, and/or the act of their transcription, are involved in important biological functions at the transcriptional and posttranscriptional level. This chapter discusses the epigenetic and nonepigenetic mechanisms by which lncRNAs and/or their transcription are involved in the programming of various biological functions in model systems, from yeast to mammals.

1.1 Introduction

A major portion of the eukaryotic genome is occupied by DNA sequences, whose transcripts do not code for proteins. It has been proposed that the size of the noncoding portion of the genome is linked to the development of complex organisms (Mattick 2004; Taft et al. 2007), as the protein-coding portion of the genome, by and large, has remained constant while the noncoding portion has grown significantly during the evolution of more complex organisms from simpler life-forms (Mattick 2004). This hypothesis indicates that these sequences are not "junk" but perhaps play a major role in the generation of organismal complexity. In the initial attempt to define the mouse transcriptome by sequencing of mouse full-length cDNA clones, it was found that the majority of the nonprotein-coding DNA region is transcribed but produces RNA with little or no protein-coding potential

R.R. Pandey and C. Kanduri (✉)
Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, Dag Hammarskjölds Väg 20, 75185 Uppsala, Sweden
e-mail: Kanduri.Chandrasekhar@genpat.uu.se

(Okazaki et al. 2002; Carninci et al. 2005). Moreover, the development of new highly sensitive and ultra high-throughput techniques such as second generation sequencing in combination with preexisting classical molecular biology techniques such as CAGE (Cap Analysis of Gene Expression) (Shiraki et al. 2003), 5′ and 3′ SAGE (Serial Analysis of Gene Expression) (Velculescu et al. 1995), ASSAGE (Asymmetric Strand-specific Analysis of Gene Expression) (He et al. 2008), and GRO (Global Run On Analysis) (Core et al. 2008) have provided us with a detailed overview of the extent of transcription in eukaryotes (Nagalakshmi et al. 2008). The results were surprising in that most of the eukaryotic genome is transcribed and produces a plethora of noncoding RNA (ncRNA) species during various stages of cellular differentiation (Kapranov et al. 2007a, b and references therein; Birney et al. 2007).

A ncRNA is defined as an RNA species with an open reading frame (ORF) of less than 100 amino acids, whereas protein-coding mRNAs have ORFs greater than 100 amino acids in length. Some of the ncRNAs are constitutively expressed in all cells, for example, ribosomal RNA, transfer RNA, and small nuclear and nucleolar RNA (snRNA, SnoRNA), and are hence known as housekeeping ncRNAs. The functions and mechanisms of action of the housekeeping ncRNAs have been investigated in greater detail in recent years. The ncRNAs, other than housekeeping ncRNAs, are broadly categorized into small ncRNAs (less than 100 nucleotides in length) and lncRNAs, which are longer than 200 nucleotides in length. The small ncRNAs are further divided into subgroups (miRNA, siRNA, piRNA, etc.) depending on their size, biogenesis, mode of action, and the proteins with which they are associated. Small ncRNAs regulate gene expression at the transcriptional level by guiding the repressive chromatin complexes known as RNA-induced transcriptional silencing and RNA-dependent RNA polymerase complexes (RITS-RDRCs) to cognate genes, and at the posttranscriptional level by guiding the effector complexes known as RNA-induced silencing complexes (RISCs) either to cleave the target mRNA or to bring about translational inhibition (Bartel 2004, 2009; Grewal and Jia 2007; Malone and Hannon 2009; Ghildiyal and Zamore 2009).

The lncRNAs are the least characterized of all the ncRNAs whose biological functions are, in any case, poorly investigated. The majority of the lncRNAs are transcribed by RNA polymerase II (RNA pol II) and possess a 5′ methyl cap and polyA tail. Depending on their location, with respect to the mRNA gene, they can be classified as (1) Sense, transcribed from the same strand as the mRNA; (2) Antisense, transcribed from the strand opposite the mRNA; (3) Intronic, the transcription unit of the lncRNA lies within an intron of another gene; and (4) Intergenic, transcribed from a region lying outside mRNA genes. Several thousand lncRNAs are predicted to be present in the eukaryotic genome; however, at present, the most difficult issue is the identification of functional lncRNAs from the vast pool of pervasively transcribed noncoding transcripts.

There is a possibility that a significant number of lncRNAs could arise from experimental artifacts. For example, genome tiling array experiments in different organisms reported thousands of *cis* natural antisense transcripts (*cis* NATs)

(Yamada et al. 2003; Bertone et al. 2004; Carninci et al. 2005; David et al. 2006; Samanta et al. 2006). However, a more recent study could only find less than half of the *cis* NATs in yeast when actinomycin D was included in the cDNA synthesis reaction to prevent false second strand synthesis (Perocchi et al. 2007), indicating that experimental artifacts could have contributed to the number of noncoding transcripts. In addition, many of the intronic lncRNAs could be fragments derived from the splicing of pre-mRNAs. Similarly, a large proportion of the intergenic transcripts could arise from the ripple effect of nearby transcription, which induces changes in nucleosome organization, thus providing an opportunity for the transcription machinery to produce transcripts of no significance from cryptic promoters (Ebisuya et al. 2008).

LncRNAs show a very low level of sequence conservation compared to protein-coding mRNAs. Nevertheless, the base substitution rate or constraint (ratio of the nucleotide substitution rate between functional sequences and neutral sequences) for ncRNAs is 90–95%, which is fairly high when compared with protein-coding sequences but still shows positive selection over the neutral sequences in the genome (Ponjavic et al. 2007), indicating that ncRNAs do possess important biological functions. The observations that lncRNAs display subcellular localization (Mercer et al. 2008), tissue- and cell type-dependent expression, specific expression in response to certain environmental cues (Cawley et al. 2004), and transcriptional regulation by key transcription factors such as p53, c-MYC, SP1 (Cawley et al. 2004), and CREB (Euskirchen et al. 2004) further emphasize that lncRNAs could play critical roles in cellular proliferation, differentiation, and the development of complex organisms.

Recently, several different approaches have been used to identify functional lncRNAs. In one approach, several hundred long intervening ncRNAs (lincRNAs) were identified using active chromatin signatures associated with RNA pol II transcription, i.e., the histone H3 lysine 4 trimethylation and histone H3 lysine 36 trimethylation domains (K4-K36 domains) (Guttman et al. 2009; Khalil et al. 2009). The studies identified 1,586 and 3,289 lincRNAs in different mouse and human cell types, respectively, and predicted that the total number of lincRNAs could be around 4,500. The lincRNAs show significant evolutionary conservation when compared to neutral sequences in the genome and many of them show changes in their expression patterns in response to different environmental stimuli, suggesting that lincRNAs could play critical roles in various biological functions (Guttman et al. 2009; Khalil et al. 2009). In another recent study, around 215 functional lncRNAs were identified based on their chromatin interaction properties (Mondal et al. 2010). The chromatin associated RNAs (CARs) also show significant evolutionary conservation and transcribed from both intronic and intergenic regions. Functional characterization of one of the CARs revealed that they regulate gene expression by regulating chromatin structure. Collectively, the above observations suggest that lncRNAs are an integral component of mammalian genetic programming.

Although the functional roles of lncRNAs are very much in evidence in diverse biological functions across the evolutionary spectrum (Bernstein and Allis 2005;

Mattick and Makunin 2006; Prasanth and Spector 2007; Amaral et al. 2008; Amaral and Mattick 2008; Sunwoo et al. 2009, and references therein), the underlying molecular mechanisms are far from clear. In this chapter, we discuss the epigenetic and nonepigenetic mechanisms by which lncRNAs regulate various biological functions in model systems, from yeast to mammals.

1.2 Pervasive ncRNA Transcription at Gene Regulatory Regions and the Link to Transcription

Several high-throughput approaches have uncovered widespread pervasive transcription across the promoter and terminator regions of annotated genes in yeast, mice, humans, and plants, which produce a complex repertoire of noncoding transcripts. These transcripts include small RNAs [miRNA, piRNA, and siRNA] as well as lncRNAs. A recent study (Kapranov et al. 2007a), aimed at profiling human and mouse transcriptomes from cell lines, used polyA+ RNA, longer than 200 nucleotides (nt), from nuclear and cytoplasmic fractions separately, and total cellular RNA of less than 200 nt in length to hybridize to tiling arrays at 5-nucleotide resolution. The study found three different RNA species: promoter-associated small RNAs (PASRs), terminator-associated small RNAs (TASRs), and promoter-associated long RNAs (PALRs). The PASRs and TASRs ranged in size from between 20 and 200 nt; however, a significant number of PASRs were between 26 and 50 nt long. PASRs were centered around the transcription start site of protein-coding genes in both directions, whereas TASRs were mostly oriented in the antisense direction at the 3′ termini of the host genes. This study further demonstrated that PASRs and TASRs are also present in mouse at the 5′ and 3′ ends of genes, respectively, indicating that these RNAs are highly conserved across the evolutionary spectrum and could have a potential role in gene regulation. PALRs are 100 nt to 1.0 kb long and map to 5′ regulatory regions, like PASRs, which suggests that many PASRs could be derived from PALRs. However, in the majority of cases, the expression of PASRs and TASRs is strongly correlated with the associated gene expression. The genes that were found to be highly enriched for PASRs and TASRs were also highly expressed and vice versa Kapranov et al. 2007a. PASRs are not produced by the Dicer-dependent cleavage mechanism as the PASR profile in mouse ES cells lacking Dicer remained unchanged (Kapranov et al. 2007a).

In addition to PASRs and TASRs, another category of highly unstable small and long ncRNAs, located close to promoters in yeast and human cells, have been described. In the budding yeast, *Saccharomyces cerevisiae*, these transcripts were upregulated in a mutant, which lacked components of the exosome machinery, and were therefore christened cryptic unstable transcripts (CUTs) (Xu et al. 2009, Wyers et al. 2005). The exosome is known to act as a surveillance pathway for the removal of unwanted RNA molecules from cells. The 3′ SAGE sequencing of CUTs peaked at

50 nt downstream and 550 nt upstream of known open reading frame (ORF) transcription start sites (TSSs). Since the average size of CUTs is around 250–300 nt, it can be concluded that they mostly originate from intergenic regions (Neil et al. 2009). CUTs are transcribed in both divergent and convergent configurations, but the former contributes to the most abundant class. To date, the functional significance of CUTs in various biological functions is still unclear.

Similar to CUTs in yeast, a subclass of promoter upstream transcripts (PROMPTs) were stabilized when HeLa cells were treated with an siRNA to knockdown hRrp40, a crucial component of the human $3'$–$5'$ exoribonucleolytic exosome (Brower et al. 2001). PROMPTs can originate more than 2.0-kb upstream of the TSS with a peak around -1.0 kb. PROMPTs are transcribed in both the sense and antisense directions with respect to the TSS of the associated gene (Preker et al. 2008). The function of PROMPTs is largely unknown, but they may play a regulatory role since certain ncRNAs, known to exert regulatory functions, are located within PROMPT regions. Interestingly, one of the ncRNAs, *Khps1*, which is transcribed in the antisense direction from the TSS of sphingosine-kinase 1 (*SPHK1*), is stabilized in hRrp40-knockdown cells. The *Khps1* transcript has been linked to the demethylation of the *SPHK1* differentially methylated region (DMR) (Imamura et al. 2004); however, the mechanism by which *Khps1* mediates demethylation is not known. Taking the data from yeast, mouse, and human together, it is clear that the divergent transcription of ncRNAs surrounding the promoter regions of annotated genes is a common and conserved feature of eukaryotic RNA pol II transcription. This is demonstrated further by the broad distribution of RNA pol II near TSSs and by the bimodal distribution of active chromatin markers such as histone H3 lysine 4 trimethylation.

Several models have been proposed for the biogenesis of pervasive transcripts at gene regulatory regions. The TSSs for most of the promoter- and terminator-associated ncRNAs fall within the nucleosome-free region (NFR) of the related genes, suggesting that perhaps they originate from the spurious activity of RNA pol II on naked DNA in the promoter, as well as the terminator regions. Nucleosome positioning is known to suppress cryptic transcription by preventing the random access of RNA polymerase to the DNA. This is clearly demonstrated in yeast containing mutations in the *spt6* gene, where the ability to reassemble nucleosomes is lost in the RNA Pol II-elongated portions of coding regions, resulting in cryptic transcription from the NFRs (Cheung et al. 2008). Moreover, insertion of an enhancer with several LexA or Gal4 binding sites induced an NFR around the site of insertion, irrespective of the genomic location, leading to cryptic transcription from the $3'$ ends of the LexA/Gal4 binding sites (Dobi and Winston 2007). Likewise, a very recent study using chromatin signatures specific to enhancer and promoter found that most of the extragenic RNA Pol II peaks overlapped the enhancer regions, indicating that long noncoding transcription is prevalent in the enhancer regions (De Santa et al. 2010). These examples clearly suggest that nucleosome positioning is critical for preventing aberrant transcription across the genome. Moreover, the majority of promoter- and terminator-associated RNAs are less abundant than protein-coding mRNA and rapidly degraded by nuclear quality control pathways in both yeast and human (Preker et al. 2008; Wyers et al. 2005),

indicating that they might possibly represent the by-products of RNA pol II spurious activity in NFR regions. However, the presence of an independent TSS for PASRs, TASRs, and PALRs, and the fact that they are conserved across the evolutionary spectrum, suggests that they are not by-products of RNA pol II spurious activity in NFR regions. Additionally, in yeast, a mutation in the TATA box of the *TPI1* gene affected expression of the mRNA but not of the sense CUT, further supporting the notion that CUTs originate from the assembly of an independent preinitiation complex (PIC) and substantiating their functional role in gene regulation (Neil et al. 2009).

The key question here is, "what is the role of pervasive transcription?" Since promoter- and terminator-associated transcripts are rapidly degraded, the transcript per se may not be directly involved in the gene regulatory process. Interestingly, the expression of promoter-associated RNAs in human cells (PASRs and PROMPTs), as well as in yeast (CUTs), correlates positively with the expression of sense mRNAs. However, when several synthetic sense and antisense PASRs, surrounding the *c-MYC* and connective tissue growth factor (*CTGF*) promoters, were transfected into HeLa cells (Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project 2009), the mRNA levels of both the *c-MYC* and *CTGF* genes were downregulated, in contrast to data suggesting a genome-wide positive correlation of PASRs with mRNA gene expression. This may explain why PASRs in human and CUTs in yeast are rapidly degraded by the exosome machinery.

Interestingly, a couple of recent investigations have further implicated PASRs in the negative regulation of cognate genes. For example, intergenic spacer regions in ribosomal gene clusters encode lncRNAs, whose promoters lie about 2.0-kb upstream of the rRNA promoters. In addition to the 2.0-kb lncRNAs, the spacer regions also contain 150–200 nt RNAs (pRNAs), which span the *rRNA* promoters, indicating that the pRNAs could be derived from the spacer lncRNAs. The pRNAs have been shown to interact with and recruit the nucleolar remodeling complex (NoRC) to *rRNA* gene promoters, and this, in turn, leads to the recruitment of components of the heterochromatin machinery, including HP1 (Mayer et al. 2006).

Like *rRNA* gene promoters, the *p21* promoter also contains promoter-associated RNAs in both sense and antisense directions. Interestingly, the generation of antisense promoter-associated RNAs, which correlates with the silencing and heterochromatinization of the *p21* sense promoter, is dependent on transcription from the *p21* antisense promoter in an Ago-1-dependent manner. This indicates that antisense pRNAs could be derived from the *p21* antisense RNA and play a critical role in the transcriptional silencing of the *p21* sense promoter (Kim et al. 2006; Morris et al. 2008). Alternatively, the transcription of PASRs and CUTs may be involved in establishing an open chromatin configuration, which would be required for high-level mRNA gene expression, or they could act as rheostats involved in maintaining a specific level of mRNA expression by competing for the same pool of transcription factors. This has been shown at least in the case of one antisense CUT promoter, where a mutation in the promoter of the *TPI1* mRNA gene resulted in several fold higher expression of the antisense CUT (Neil et al. 2009). Although there is no genome-wide study yet available to describe the function of the 5′ and 3′ associated small and long ncRNAs in the regulation of mRNA genes, several

studies covering individual CUTs/PARs highlight that different mechanisms are being used to control mRNA gene expression at various levels.

1.3   Transcriptional Silencing by Noncoding Transcription via Transcriptional Interference

Transcriptional interference (TI) refers to the suppressive effect of one transcriptional event on a second transcriptional event *in cis*. TI occurs when two promoters are convergent or in tandem. The elongating complex from one promoter can affect the transcriptional initiation (by interfering with preinitiation complex assembly), elongation, or termination step of the second promoter, depending on its physical relationship with the first promoter. For example, the first promoter only affects PIC assembly when the second promoter is in tandem, but can affect PIC assembly, transcriptional elongation or termination when the second promoter is transcribed convergently. Although few eukaryotic genes have been shown to be regulated by a transcriptional interference mechanism involving lncRNA transcription, the observation that most protein-coding genes in higher eukaryotes have overlapping transcription from promoters in the upstream intergenic region or from downstream intragenic sense and antisense promoters, suggesting that transcriptional regulation by TI could be a common mechanism for regulating protein-coding genes. Here, we provide the biological contexts in which noncoding transcription regulates protein-coding genes via TI.

In the yeast *S. cerevisiae*, a gene involved in the serine biosynthesis pathway, *SER3*, is transcribed in nutrient-poor media; however, in nutrient-rich media, the *SER3* gene is silenced due to the activation of a noncoding RNA gene promoter *SRG1*, located upstream of the *SER3* gene. In the presence of serine in the nutrient-rich media, a serine-dependent activator, Cha4, along with chromatin remodeling complexes such as SAGA and SWI/SNF, binds to the *SRG1* promoter to activate its transcription (Martens et al. 2004; 2005) across the *SER3* promoter, leading to repression of the *SER3* gene. Promoter competition for basal transcription factors is not involved in *SER3* transcriptional repression, as the incorporation of a transcription termination signal for the *SRG1* transcript, upstream of the *SER3* promoter, resulted in derepression of *SER3*. This indicated that it is not the *SRG1* ncRNA but its transcription across the *SER-3* promoter that is required for its transcriptional repression. More importantly, it has been shown that *SRG1* transcription across the *SER-3* promoter interferes with the binding of transcription factors (Fig. 1.1aI), resulting in *SER3* gene silencing (Martens et al. 2004).

The inhibition of transcriptional initiation and elongation as means of cell type-specific gene regulation by overlapping antisense ncRNA transcription is beautifully illustrated in the diploid and haploid cells of the budding yeast, *S. cerevisiae*. In nutrient-rich media, *S. cerevisiae* cells divide mitotically to produce more diploid cells, whereas during starvation, the yeast undergoes meiotic division to produce
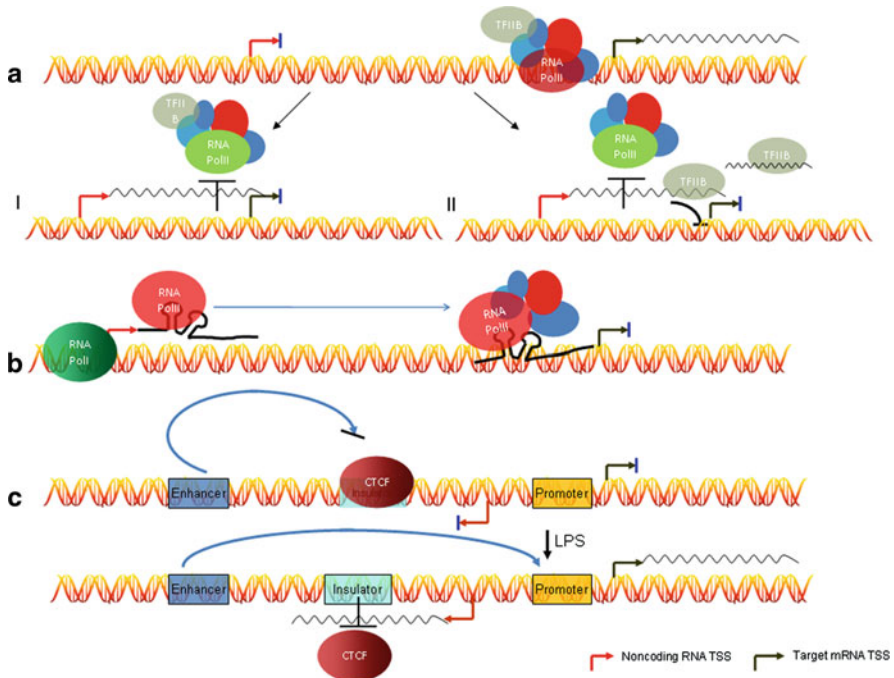
Fig. 1.1 Transcriptional silencing by lncRNAs via transcriptional interference. (**aI**), The transcription of a ncRNA through the promoter region of a target gene causes the occlusion of basal transcription machinery, thus repressing the transcription of the target gene. (**aII**) The lncRNA from DHFR minor promoter binds to TFIIB and titrates away the components of the preinitiation complex (PIC) from the DHFR major promoter. (**b**) The *Alu* and *B2* ncRNAs possess a modular structure, which includes two domains: an RNA pol II binding domain and a transcriptional inhibitory domain, which inhibits the transcription initiation step. The *Alu* and *B2* RNA inhibitory domains do not interfere with the binding of transcription factors to the ncRNAs but inhibit formation of the proper contact between RNA pol II and the DNA promoter elements required for the initiation of transcription. (**c**) In the chicken Lysozyme gene, CTCF target sites maintain silencing of the Lysozyme gene by preventing the communication of the upstream enhancer elements with the downstream Lysozyme promoter. In response to proinflammatory signals such as lipopolysaccharide (LPS), the lncRNA, LINoCR, transcription is activated across the CTCF target sites, resulting in the eviction of CTCF from its target site and activation of downstream Lysozyme promoter

haploid cells. This event is controlled by several genes, including *IME4* (initiator of meiosis). In diploid cells, only *IME4* sense mRNA was detected, whereas in haploid cells, an antisense ncRNA to the *IME4* gene was discovered, indicating that both sense and antisense *IME4* RNAs can affect each other's transcription (Hongay et al. 2006). Moreover, the separation of otherwise overlapping sense and antisense *IME4* transcription units resulted in the loss of the reciprocal effect on transcription, indicating that TI could be the mechanism in common between the sense and antisense transcriptional silencing effects *in cis*.

NcRNA transcription is not always involved in the repression of overlapping genes; sometimes it is engaged in activation of the associated gene by interfering with the binding of repressor complexes such as the chromatin insulator protein CTCF, which is known to function as a transcriptional repressor or an enhancer blocker (Kanduri et al. 2002; Phillips and Corces 2009, and references therein). The lysozyme gene in chicken has three enhancers at 2.7, 3.9, and 6.1 kb upstream of the TSS and is induced in response to proinflammatory signals such as lipopolysaccharide (LPS) in a chicken macrophage cell line. The silencing of the lysozyme gene is maintained by CTCF, whose target site maps to the region between the enhancers and the lysozyme promoter (Fig. 1.1c). The LPS induction of macrophages results in transcription of an ncRNA, LINoCR (LPS induced noncoding RNA). The transcription of LINoCR through CTCF target sites results in expulsion of the CTCF protein due to the positioning of a nucleosome over the CTCF target site (Lefevre et al. 2008). The expulsion of CTCF, and chromatin remodeling by LINoCR transcription, which further inhibits the binding of CTCF to its target site, facilitates enhancer/promoter communication, leading to lysozyme gene activation in response to the LPS proinflammatory signal (Fig. 1.1c).

Intriguingly, the interplay between the transcriptional processes of two intergenic noncoding transcription units in *S. cerevisiae* determines the transcriptional activity of the neighboring *FLO11* protein-coding gene (Bumgarner et al. 2009). *FLO11*, which encodes a cell-wall glycoprotein controlling cell–cell adhesion, has a variegated expression pattern; in some cells the gene is highly expressed, while in the other cells, it is completely repressed. This variegated or binary expression is the result of functional interplay between two *cis*-interfering lncRNAs, upstream of the *FLO11* gene. The 5′ regulatory region of *FLO11* is fairly long (3.4 kb) and harbors binding sites for several transcription factors, such as Sfl1 and Flo8, which overlap the two lncRNAs transcribed from opposite strands (Bumgarner et al. 2009). One of the ncRNAs, *ICR1* (Interfering Crick RNA), is transcribed from the same strand as *FLO11* and runs across the *FLO11* promoter, causing repression of the *FLO11* gene by the promoter occlusion mechanism. The second ncRNA, *PWR1* (Promoting Watson RNA), is transcribed from the complementary strand of *ICR1* and passes through its promoter, causing repression of the *ICR1* ncRNA and indirectly activating *FLO11* transcription. The transcription of *PWR1* is highly regulated. The Flo8 transcription factor specifically activates *PWR1*, resulting in the silencing of *ICR1* and, as a consequence, derepression of the *FLO11* gene. On the other hand, the transcriptional inhibitor, Sfl1, represses the *PWR1* promoter, causing repression of the *FLO11* gene via derepression of the ICR1 promoter, presumably by interfering with the binding of the transcriptional initiation machinery (Bumgarner et al. 2009). This is a very interesting example of how the interplay between two functional intergenic ncRNAs determines the activity of flanking protein-coding mRNA, and it highlights the fact that ncRNA-mediated transcriptional regulatory mechanisms are multilayered and highly complex.

Recent evidence suggests that gene regulation via TI constitutes one of the significant gene regulatory mechanisms in mammals. The functional role of TI in

transcriptional regulation is well characterized in the *DHFR* (di-hydro folate reductase) gene in quiescent cells. *DHFR* has two promoters, one major and one minor. In rapidly growing human cells, *DHFR* mRNA is transcribed from the major promoter to fulfill the high demand for DNA synthesis. In quiescent cells, a high level of *DHFR* gene transcription is not required; therefore, the DHFR gene needs to be silenced. Interestingly, transcriptional silencing of the major promoter is achieved by ncRNA transcription from the 5′ upstream minor promoter. The ncRNA produced from the minor promoter forms a triplex structure at the major promoter and interferes with the formation of the preinitiation complex. Furthermore, the ncRNA from the minor promoter also interacts with TFIIB, thus titrating away the components of the preinitiation complex (Fig. 1.1aII). These results indicate that both the ncRNA and the act of its transcription play a crucial role in the transcriptional repression of the *DHFR* major promoter via dissociation of the preinitiation complex (Blume et al. 2003; Martianov et al. 2007).

1.4 Heritable Epigenetic Gene Inactivation via Noncoding Transcription

Epigenetic gene silencing refers to the heritable mechanisms that mediate gene silencing without any changes in the primary DNA sequence. For example, post-translational histone modifications, such as di- and trimethylation of the histone H3 lysine 9 residue (H3K9me2 and H3K9me3) and trimethylation of the histone H3 lysine 27 residue (H3K27me3), and DNA methylation are often enriched at transcriptionally silenced genes (Kouzarides 2007 and references therein). Recent evidence suggests that transcriptional read-through of a neighboring gene by sense or antisense transcription results in heritable epigenetic gene inactivation, which has been shown to occur mostly in disease conditions. For example, the mismatch repair gene, *MSH2*, is often methylated or deleted in Lynch syndrome patients who are susceptible to colorectal and endometrial cancers. A recent study demonstrated that a deletion at the 3′ end of the *TACSTD1* gene resulted in extension of its transcription into the downstream *MSH2* gene, causing specific methylation and transcriptional inactivation of its promoter (Ligtenberg et al. 2009). However, it is not clear how transcriptional read-through across the *MSH2* promoter leads to its methylation.

A similar mode of action was detected as part of a disease mechanism in patients with an inherited form of alpha-Thalassemia, where transcriptional silencing of the *HBA2* gene was detected due to aberrant antisense transcription across its promoter (Tufarelli et al. 2003). In these patients, deletion of a region between the *HBA2* gene (α2 globin) and the *LUC7L* gene places the truncated *LUC7L* gene very close to the *HBA2* gene, resulting in transcriptional read-through from the *LUC7L* promoter into the normally expressed *HBA2* gene promoter. This transcriptional read-through causes DNA methylation and silencing of the *HBA2* gene.

Furthermore, a transgenic mouse model was used to show that antisense transcription through the *HBA2* promoter CpG island is necessary and sufficient to cause *HBA2* promoter DNA methylation and silencing (Tufarelli et al. 2003). In both instances, transcriptional silencing of protein-coding genes occurred due to aberrant transcriptional read-through, indicating that common mechanisms are used in aberrant and programmed silencing, and the only difference is the direction of transcription: in the former, it is sense, and in the latter, it is antisense.

Transcriptional silencing by aberrant natural antisense transcription across promoters appears to be a common feature in various diseases as it has also been documented in tumor suppressor genes such as *p15* and *p21*. The *p15* gene is a key tumor suppressor gene, and the loss of *p15* expression either by deletion, point mutation, or promoter hypermethylation is associated with a variety of tumors (Nobori et al. 1994). Recently, an ncRNA transcribed antisense to the *p15* gene (*p15AS*) was identified. This antisense RNA was shown to be expressed in leukemia cells at higher levels than in normal cells (Yu et al. 2008). Interestingly, *p15* antisense RNA transcription leads to enrichment of the repressive chromatin mark (H3K9me3) over the *p15* promoter and exon 1. The expression of *p15AS* is also correlated with *p15* promoter DNA hypermethylation. The epigenetic silencing of the *p15* promoter by *p15AS* is Dicer-independent, indicating that it is not mediated by RNA interference. Like the p15 gene, the *p21* gene is also often methylated and silenced in several cancers. Recent investigation has shown that bidirectional transcription of the *p21* gene is critical for its balanced expression. Suppression of steady state levels of the *p21* antisense RNA (*p21AS*) results in activation of the *p21* sense RNA. The repression of *p21* sense RNA by *p21AS* is mediated in an Ago-1-dependent manner via formation of heterochromatin over the *p15* sense promoter (Kim et al. 2006; Morris et al. 2008).

In the above four examples, the sense genes are silenced epigenetically via heterochromatin formation at the promoter due to aberrant transcription in the sense or antisense directions (Fig. 1.2a). Though heterochromatin formation over the silenced promoters is common in all the cases, it is not clear whether common mechanism(s) are involved. It is also not apparent, from the available data, whether the act of transcription, or the RNA itself, mediates transcriptional silencing. Although, in the case of *p15*, the data point towards a functional role for the RNA, it needs to be thoroughly investigated before the act of transcription is ruled out as the mechanism involved in transcriptional silencing.

1.5 LncRNAs Mediate Long-Range Gene Silencing Through the Recruitment of Polycomb Repressor Complexes

In mammals, subsets of genes are expressed from one of the parental alleles, while the other allele is often silenced by repressive epigenetic modifications. This allele-specific silencing is most prevalent in imprinted gene clusters and on the inactive
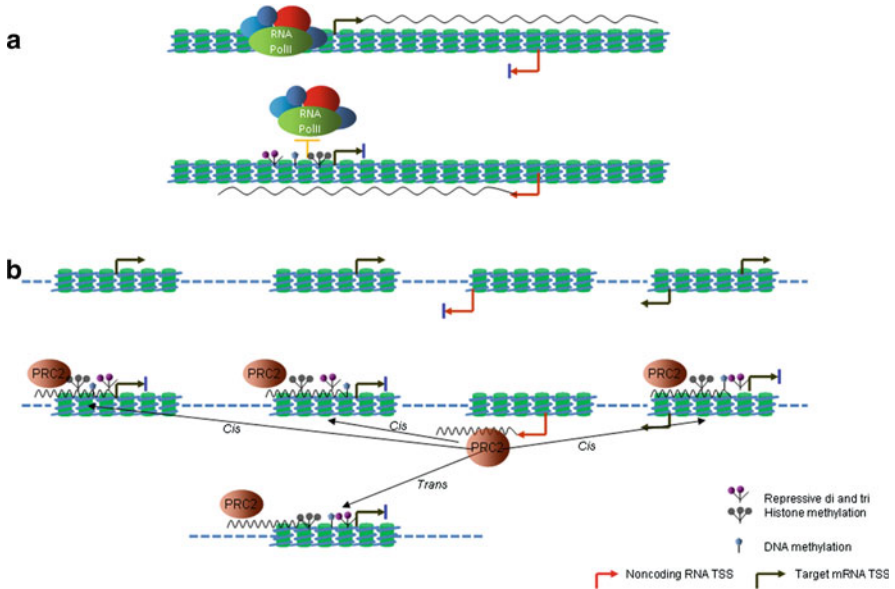
Fig. 1.2 Epigenetic reprograming of individual as well as domain-wide gene regulation by lncRNAs or its transcription. (**a**) An antisense ncRNA transcription across the promoter of the overlapping sense gene causes the formation of repressive chromatin environment via the enrichment of repressive modifications such as H3K27me3, H3K9me2, H3K9me3, and DNA methylation, thus repressing the overlapping sense gene. (**b**) The lncRNA-mediated regulation of gene expression in chromosomal domains via targeting of PRC2 complexes *in cis* or *in trans*

X chromosome in female mammals. In imprinted domains, allele-specific gene silencing occurs in a parent of origin-specific manner. In the case of the X chromosome in female mammals, allele-specific gene silencing also occurs in a parent of origin-specific manner (X-linked genes are silenced only on the paternal chromosome) in preimplantation embryos, whereas it occurs at random later in embryonic development. Interestingly, lncRNAs have been shown to play an important role in the establishment and maintenance of allele-specific gene silencing.

Cells in female mammals have two X chromosomes, whereas males have only one X. In order to equalize the dosage of X-linked gene products between males and females, one of the X chromosomes becomes inactivated during early embryonic development in female mammals (Payer and Lee 2008 and references therein, Chap. 3). The X chromosome inactivation center (XIC), a 500-kb region on the X chromosome, is implicated in X chromosome inactivation (XCI). The XIC harbors several genes for lncRNAs, for example, *Xist* (X inactivation specific transcript), *Tsix* (an antisense transcript to *Xist*), *Xite*, *DXPas34*, and *RepA* among others. *Xist* plays an important role in XCI by directing the heterochromatin machinery along the inactive X chromosome, and the other lncRNAs are involved in the regulation of *Xist* expression, and thus control the counting and choice processes of XCI (Payer and Lee 2008).

Mouse embryonic stem cells (ES cells) have been widely used to study XCI as they faithfully recapitulate the molecular events that serve to establish random XCI in the inner cell mass (ICM) of blastocysts. In ES cells, both X chromosomes are active and *Xist* RNA expression is maintained at very low levels on both chromosomes by pluripotency factors such as Nanog, Oct3/4, and Sox2 (Navarro et al. 2008). Upon differentiation, *Xist* RNA is upregulated on the future inactive X chromosome and spreads along the X chromosome *in cis*, accompanied by accumulation of repressive histone marks (H3k27me3 and H3K9me3), CpG DNA methylation, and deposition of the histone variant macroH2A1, thus establishing a repressive chromatin environment devoid of RNA Pol II. The A region, rich in repeats, at the 5′ end of *Xist* was shown to be critical for the establishment of XCI (Wutz et al. 2002). Deletion of this region compromised the accumulation of repressive histone modifications and silencing of X-linked genes *in cis*, suggesting that this repeat-rich region recruits the repressive histone modification machinery to the X chromosome *in cis*.

Recently, a new lncRNA (*RepA*) of 1.6 kb in length was discovered at the 5′ end of the *Xist* gene, covering the A repeat-rich region of the *Xist* gene (Zhao et al. 2008). *RepA* associates with the PRC2 complex members, EZH2 and SUZ12, before and during XCI. Interestingly, the PRC2 complexes are targeted to chromatin only at the onset of XCI. In light of the identification of a new member in the long list of lncRNAs involved in XCI, it would be interesting to investigate whether the *Xist* and *RepA* RNAs function synergistically in the XCI process or whether they have altogether different functions. However, an earlier study investigating the dynamics of XCI found that the *Xist* RNA forms a repressive compartment in the early phases of ES cell differentiation. The repressive compartment excludes the RNA polymerase II machinery from the genes to be silenced (Chaumeil et al. 2006), and this step is not dependent on the A repeat-rich region of *Xist* as ES cells in which the A region has been deleted still form the repressive compartment. However, the formation of the repressive compartment followed by accumulation of the H3K27me3 marks, and the translocation of X-linked genes into the core of the repressive compartment is dependent on the A repeat-rich region, indicating that the A-repeat plays a critical role in the transcriptional silencing of X-linked genes (Chaumeil et al. 2006). Together, these observations suggest that, at the onset of XCI, *Xist* organizes a repressive chromatin compartment, which includes all the genes to be silenced on the future inactive X chromosome. This is followed by *RepA*-dependent recruitment of the PRC2 complex members to stabilize the repressive compartment by repressive chromatin modifications (Zhao et al. 2008).

Similar to *Xist*/*RepA*-mediated XCI, a subclass of lncRNAs, including *Kcnq1ot1* and *Airn*, mediate transcriptional gene silencing in imprinted chromosomal domains in mouse. The molecular mechanism by which these two lncRNAs mediate gene silencing shows many similarities to the *Xist* RNA-mediated XCI. Both *Kcnq1ot1* and *Airn* are ~100 kb long RNA pol II-encoded ncRNAs, transcribed from the paternal allele of mouse chromosomes 7 and 17, respectively. They are responsible for the silencing of multiple genes spread over several hundred kilobases of the genome (Fitzpatrick et al. 2002; Sleutels et al. 2002;

Thakur et al. 2004; Kanduri et al. 2006). Both the lncRNAs have been shown to coat the chromatin of their target genes (Murakami et al. 2007; Nagano et al. 2008; Mohammad et al. 2008). *Kcnq1ot1* target genes show significant enrichment of the repressive chromatin marks, H3K27me3 and H3K9me3, but not the active chromatin marks H3K9ac and H3K4me3 (Pandey et al. 2008). Similarly, *Airn* ncRNA target genes show enrichment of H3K9me3 (Nagano et al. 2008). The presence of repressive chromatin marks over target genes is correlated with the association of *Kcnq1ot1* with the PRC2 members (EZH2 and SUZ12) and G9a (H3K9 histone methyltransferase) and of *Airn* with G9a (Nagano et al. 2008; Pandey et al. 2008). Collectively, these observations suggest that these lncRNAs interact with heterochromatin proteins and recruit them to the target genes, thus modifying the chromatin structure surrounding the promoters (Fig. 1.2b). Interestingly, both *Kcnq1ot1* and *Airn* have been shown to silence genes by organizing repressive chromatin compartments similar to that seen in case of *Xist* (Redrup et al. 2009). Another striking similarity between *Kcnq1ot1* and *Xist* is that, like *Xist*, *Kcnq1ot1* harbors a 0.9 kb silencing domain (SD) at the 5′ end of the RNA, which is crucial for the epigenetic silencing of its target genes (Wutz et al. 2002; Mohammad et al. 2008). Once the silencing of the target genes is established, it is equally important to maintain silencing through subsequent cell divisions, and it is possible that this is achieved by targeting the silenced gene to the heterochromatin nuclear compartments. Like *Xist*, *Kcnq1ot1* has been shown to maintain transcriptional silencing by recruiting genes to the perinucleolar space, which is enriched with heterochromatin factors such as Ezh2 (Mohammad et al. 2008; Zhang et al. 2007).

Intriguingly, lncRNAs have also been implicated in gene silencing *in trans*. In an elegant study using human primary fibroblast cells, it was shown that transcription of the *HOTAIR* lncRNA from the *HOXC* cluster correlates with the appearance of H3K27me3 marks over the *HOXD* cluster, which resides on another chromosome (Rinn et al. 2007). Depletion of *HOTAIR* using siRNA technology resulted in the loss of H3K27me3 marks over the *HOXD* cluster, indicating a link between *HOTAIR* expression from the *HOXC* locus and the enrichment of H3K27me3 marks over the *HOXD* cluster. Moreover, *HOTAIR* was shown to interact with the PRC2 members, EZH2 and SUZ12, in both in vitro and in vivo experiments. On the basis of the above observations, the authors speculated that *HOTAIR* interacts and guides the PRC2 complex to the *HOXD* cluster to silence the genes by H3K27me3 chromatin modification (Fig. 1.2b) (Rinn et al. 2007). Furthermore, a recent study demonstrated that the overexpression of *HOTAIR* in epithelial cancer cells resulted in genome-wide changes in the PRC2 complex occupancy and enhanced cancer invasiveness and metastasis (Gupta et al. 2010). This link between lncRNA-mediated epigenome reprogramming and cancer is most interesting.

Taken together, a consensus seems to be emerging by which lncRNAs are involved in epigenetic gene silencing. Upon transcription, these lncRNAs form ribonucleoprotein (RNP) complexes with repressive histone modification machinery. This could be achieved either by the interaction of proteins with a linear RNA sequence or by formation of an RNA secondary structure. The latter possibility is perhaps more likely as, even though there are no sequence similarities between the

above-mentioned lncRNAs, they still form RNP complexes with the same proteins. Supporting this idea, a 2-D structure of the *Xist* A region in mouse and human has been shown to be important for binding of the PRC2 complex to *Xist* (Maenner et al. 2010). The RNPs are then directed to the target genes, either *in cis* or *in trans*, by an unknown mechanism, thus resulting in higher order repressive chromatin formation and silencing of the associated genes. This silenced state can be further stabilized and maintained through subsequent cell divisions by targeting the silenced genes to the nucleolar or perinuclear region (Zhang et al. 2007; Mohammad et al. 2008).

Although our knowledge of lncRNA-mediated epigenetic gene silencing has significantly improved in the past few years, several key questions remain to be answered. First, how do lncRNAs maintain their high levels of expression in a repressive chromatin environment? Do they need a repressive chromatin environment for high expression levels, or do they have a different mechanism to combat this problem? For example, the presence of boundary elements flanking the lncRNA promoter and coding sequences, which prevent the spread of heterochromatin formation into the lncRNA gene, or the presence of strong promoter elements, which can overcome the heterochromatinization by recruiting p300/pCAF, or similar transcriptional activators (Pandey et al. 2004), or both. Second, how are RNP complexes targeted to specific genes, whereas other genes residing in between the target genes escape silencing? Since no sequence homology between lncRNAs and their target genes has been reported so far, it is unlikely that targeting is based on sequence similarity.

1.6 LncRNA-Mediated Targeting of Activator Complexes in Epigenetic Gene Activation

Some lncRNAs have been shown to activate genes through targeting activator complexes to gene regulatory regions. This is best exemplified in the case of the *roX* RNA-mediated hyperactivation of the X chromosome in *Drosophila melanogaster* (see Chap. 7). In contrast to mammals, where dosage of X-linked gene products between males and females is achieved via inactivation of one of the two X chromosomes in females, equal dosage of X-linked gene products between male flies with one X chromosome and female flies with two X chromosomes is achieved by hypertranscription of the lone X chromosome in males. The upregulation of X-linked genes is achieved by *roX* RNA-dependent targeting of the dosage compensation complex (DCC) at several loci along the X chromosome. The DCC consists of five proteins, MSL1 (male specific lethal), MSL2, MSL3, MLE (Male-less), MOF (Males absent on the first), and two lncRNAs: *roX1* and *roX2* (RNA on the X). MSL1 and MSL2 are necessary for DCC binding to DNA; MOF is an enzyme that catalyzes the acetylation of lysine 16 on histone H4 (H4K16ac), a modification crucial for the transcriptional upregulation of genes on the X chromosome (Gelbart et al. 2009); MLE is an ATP-dependent RNA/DNA helicase,

required for the incorporation of *roX* RNA into the DCC. The *roX1* and *roX2* ncRNAs are transcribed from the X chromosome and either of them is sufficient for correct localization of the DCC along the X chromosome. Deletion or mutation of both *roX* RNAs resulted in mislocalization of the DCC complex to the chromocenter and the heterochromatin regions of autosomes (Meller and Rattner 2002; Chap. 7).

In flies, hundreds of small GA-rich DNA elements, known as chromatin entry sites (CESs) or high affinity sites (HASs), are present across the X chromosome. The DCC can recognize and bind to CESs in the absence of *roX* lncRNAs; however, gene activation cannot be achieved (Alekseyenko et al. 2006, 2008; Straub et al. 2008), indicating that *roX* lncRNAs are an integral part of the DCC complex. Intriguingly, CESs are enriched only twofold on the X chromosome when compared to autosomes, suggesting that CESs alone are not sufficient for X chromosome recognition by the DCC. Moreover, autosomal transgene copies of *roX* can rescue male embryos carrying deletions of the *roX1/2* RNA genes. In these embryos, the DCC was localized to the X chromosome and also to limited autosomal loci, further suggesting that the mere presence of CESs on autosomes is not sufficient for correct targeting of the DCC to autosomes. The CES provides an entry point for the DCC; however, transcriptional upregulation of genes requires spreading of the DCC from the CES and the H4K16ac modification of chromatin (Gelbart and Kuroda 2009 and references therein). MSL3, another member of the DCC, contains a chromodomain, which has been shown to bind to nucleosomes with the H3K36me3 modification in vitro. The chromodomain of MSL3, along with MLE and MOF, is required for the spreading of the DCC complex (Sural et al. 2008). Although the exact role of the *roX* lncRNAs is not yet clear, it has been suggested that they are vital for the cotranscriptional assembly of the DCC, increasing the affinity of the DCC for the CES and in enhancing the enzymatic activity of MOF in the DCC complex (Gelbart et al. 2009).

LncRNA-mediated transcriptional activation through the recruitment of activator complexes has also been reported at the single gene level. For example, ncRNAs, encoded by polycomb/trithorax elements in the *Bxd* region in *Drosophila*, recruit a member of the trithorax complex, ASH1, to the downstream *Ubx* gene by forming base pair interactions with DNA. ASH1 is a histone methyltransferase containing a SET domain and its ncRNA-dependent recruitment to the *Ubx* gene promoter results in active chromatin formation and transcriptional activation of the *Ubx* gene (Fig. 1.3) (Sanchez-ELsner et al. 2006).

Epigenetic gene activation by lncRNAs is also implicated in the regulation of *Hox* genes during the primitive streak phase of embryoid body (EB) differentiation in mice (Dinger et al. 2008). *Evx1as* and *Hoxb5/6as* lncRNAs show concordant expression with the *Evx1* and *Hox5/6* genes, respectively. The *Evx1as* and *Hoxb5/6as* lncRNAs are enriched in the active chromatin compartment (H3K4me3) and also interact with MLL1 (a histone methyltransferase responsible for H3K4me3 methylation), which suggests that these lncRNAs activate flanking genes through the establishment of active chromatin structures (Fig. 1.3) (Dinger et al. 2008).
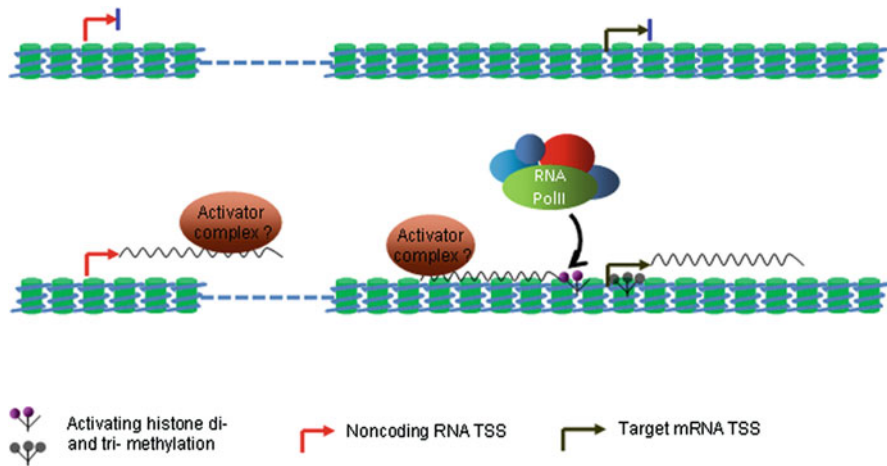
Fig. 1.3 Epigenetic gene activation through the targeting of activator complexes to the gene regulatory regions. Intergenic lncRNAs have been shown to associate with H3K4me3 histone methylatransferases such as ASH1 in Drosophila, and MLL1 in mammals, and target them to the promoters of nearby genes to activate their transcription through establishing active chromatin marks

However, the absolute requirement of *Evx1as* and *Hoxb5/6as* lncRNAs in the gene activation process has not been investigated.

Interestingly, in a recent investigation, a long intergenic ncRNA, *Intergenic 10*, was implicated in the activation of the flanking genes, *FANK1* and *ADAM12*, via the formation of active chromatin structures (Mondal et al. 2010). Downregulation of *Intergenic 10* in human fibroblasts resulted in significant loss of expression and active chromatin marks, such as H3K4me3, from the flanking genes, indicating that this lncRNA specifically activates its flanking genes. Except for *roX* lncRNAs, which act at the RNA level, it is not clear whether the process of transcription, or the ncRNA itself, takes part in the biological events involving lncRNAs described above.

1.7 Transcriptional Regulation of Heat Shock Response by lncRNAs

LncRNAs have been implicated in the global transcriptional upregulation of heat shock responsive genes and in the downregulation of housekeeping genes during the heat shock response. Transcriptional upregulation upon heat shock in mammals is mediated by heat shock factor 1 (HSF1). Under normal growth conditions, HSF1 is associated with hsp90 and other chaperones in an inactive complex, which cannot bind to heat shock elements (HSEs) found in the promoters of heat shock responsive genes. Upon heat shock treatment of cells, HSF1 is released from the inactive complex and forms an HSF1 trimer with the help of eEF1A (eukaryotic elongation factor 1A) and a lncRNA, *HSR1* (Shamovsky and Nudler 2008). The trimeric HSF1

then binds to HSEs to activate heat shock responsive genes. The lncRNA, *HSR1*, is ubiquitously expressed in cells growing under normal conditions. Heat shock causes a conformational change in the *HSR1* structure, which, together with eEF1A, facilitates *HSF1* trimerization and its DNA binding, leading to transcriptional activation of heat shock responsive genes.

Conversely, two other lncRNAs, *Alu* and *B2*, transcribed from *Alu* repeats in human and *SINE B2* (short interspersed elements B2) repeats in mouse, respectively, are known to inhibit transcription from housekeeping genes during the heat shock response (Mariner et al. 2008; Yakovchuk et al. 2009). The *Alu* and *B2* ncRNAs possess a modular structure, which includes two domains: an RNA pol II binding domain and a transcriptional inhibitory domain, both of which are essential for the transcriptional repression of target genes. It has been demonstrated that *Alu* and *B2* RNAs bind to RNA pol II before formation of the preinitiation complex and that the binding of the ncRNA with RNA pol II does not inhibit the association of RNA poll II with general transcription factors (Chap. 6). The *Alu* and *B2* RNA inhibitory domains inhibit formation of the contact between RNA pol II and the DNA promoter elements required for the initiation of transcription, perhaps by changing the structure of the transcription complex (Fig. 1.1b). Intriguingly, *Alu* and *B2* RNAs share no sequence similarity, yet they function via a similar mechanism (Yakovchuk et al. 2009), probably due to the similarity of their secondary structures, indicating that the lack of conservation at the primary sequence level does not necessarily mean lack of function and that secondary structures could harbor critical functional information.

1.8 LncRNAs Regulate Transcription by Modulating Protein Activity

Many transcription factors are localized in the cytoplasm of resting cells. In response to external stimuli, they are transported from the cytoplasm to the nucleus to activate the transcription of an array of genes. This cytoplasmic to nuclear transport is mediated by various different mechanisms generally thought to involve proteins. A genome-wide screen to identify lncRNAs that inhibit the NFAT (Nuclear Factor of Activated T cells) activity in a human cell line identified a noncoding repressor of NFAT (NRON) (Willingham et al. 2005). The NRON inhibits NFAT nuclear import by associating with members of the importin-beta superfamily, which are involved in the nucleocytoplasmic transport of protein cargos (Willingham et al. 2005). Although the exact mechanism of this inhibition in not clear, it suggests the importance of lncRNAs in such processes.

It is intriguing to note that lncRNAs can also modulate gene activation programs globally by regulating the functions of key transcription factors or signaling molecules. One such case is the regulation of the transcriptional activation of several genes by the glucocorticoid receptor (GR) in response to glucocorticoids.

The GR is a cytoplasmic protein, which upon ligand binding, moves into the nucleus and binds to glucocorticoid response elements (GRE) via its DNA binding domain. This results in the recruitment of transcriptional activators and coactivators to the regulatory regions of GR-responsive genes, and ultimately, in the activation of GR-responsive genes.

A noncoding transcript known as growth arrest-specific 5 (*Gas5*) accumulates in growth-arrested cells. Overexpression of *Gas5* inhibits GR binding to GRE elements in a dose-dependent manner, suggesting a direct role for the *Gas5* ncRNA in GR-mediated transcriptional reprogramming. Deletion studies, to pin down the *Gas5* ncRNA region responsible for the inhibition of GR binding to the GRE, revealed a short region forming a hairpin structure with a GRE-like sequence. Mutation in this GRE-like sequence, or in the DNA binding domain of GR, abolished GR binding to *Gas5*. Taken together, these results suggest that the GRE-like structure in the *Gas5* ncRNA titrates out the ligand-bound GR, thus inhibiting the activation of GR-responsive genes (Kino et al. 2010).

Furthermore, ncRNAs can also alter chromatin-bound protein activity by allosterically modifying protein structure. This has been elegantly demonstrated in the case of the cyclin D1 (*CCND1*) gene in response to DNA damage. The transcription of the *CCND1* gene is dependent on histone acetylation of its promoter, mediated by the histone acetyl transferase (HAT) activity of CREB binding protein (CBP). The *CCND1* gene is silenced when cells are exposed to agents that damage DNA, such as ionizing radiation. Upon exposure to ionizing radiation, an RNA binding protein, TLS (translocated in liposarcoma), is recruited to the *CCND1* gene promoter by ncRNAs transcribed from the *CCND1* 5′ regulatory region (Wang et al. 2008). These ncRNAs are not only responsible for TLS recruitment but also allosterically modify the TLS protein such that it inhibits the HAT activity of CBP (Wang et al. 2008). The examples described above further emphasize the complexity of gene regulation in higher organisms and the power of lncRNAs to regulate each and every step of transcriptional regulatory mechanisms.

1.9 LncRNAs Regulates mRNA Splicing, Stability, and Translation

Posttranscriptional control of gene expression is critical for the quick response of cells to changes in external stimuli. Posttranscriptional regulation involves the regulation of mRNA splicing, mRNA localization, and mRNA stability and translation, and evidence from recent investigations suggests that these steps are also regulated by lncRNAs. The epithelial to mesenchymal transition (EMT) is a crucial step in organismal development and involves the downregulation of the E-cadherin gene in mesenchymal cells. E-cadherin is downregulated by ZEB2, a transcriptional repressor (Guaita et al. 2002). Interestingly, the *Zeb2* gene is transcribed in both epithelial and mesenchymal cells, but in epithelial cells, its translation is prevented
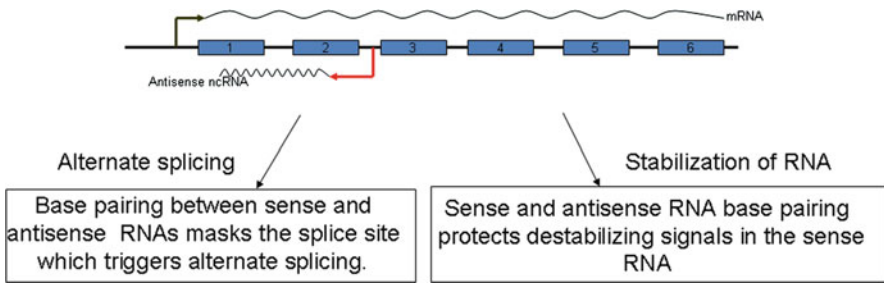
Fig. 1.4 Posttranscriptional gene regulation by an antisense ncRNA. Sense/antisense hybrid formation masks splice junctions or mRNA destabilization signals, leading to alternative splicing or stabilization of the sense transcript

by a splicing event, which removes the IRES (Internal Ribosome Entry Site) containing the 5′ UTR. In the mesenchymal cells, on the other hand, an antisense RNA overlapping the 5′ UTR splice site forms a sense–antisense RNA hybrid, which prevents splicing of the 5′ UTR and the IRES, thereby allowing translation of the *Zeb2* mRNA (Fig. 1.4) (Beltran et al. 2008).

A recent investigation has implicated long antisense ncRNAs in the pathogenesis of Alzheimer's disease. It has been shown that an antisense ncRNA, (*BACE1-AS*), against *β-Secretase*, also known as *BACE1*, is upregulated in Alzheimer's patients, and that *BACE1-AS* upregulation is linked to the stabilization of the *BACE1* mRNA, and thus an increase in its protein level (Fig. 1.4) (Faghihi et al. 2008). This increase in BACE1 results in cell stress through the production of the amyloid β 1–42 peptide, which in turn increases the production of *BACE1-AS* in a feed-forward mechanism (Faghihi et al. 2008). It is not yet clear how *BACE1-AS* increases the stability of *BACE1*. Conversely, an antisense RNA (*aHIF*) originating from the 3′ UTR of the hypoxia-inducible factor 1 alpha (*HIF-1α*) has been proposed to reduce the stability of the *HIF-1α* mRNA (Rossignol et al. 2004). The *HIF-1α* mRNA 3′ UTR has AU-rich elements that are known to act as signals for RNA degradation. In cells expressing low levels of *aHIF*, the AU-rich elements of *HIF-1α* mRNA are not exposed due to complex secondary structure formation; however, when *aHIF* is present at higher levels, it is proposed to form an RNA–RNA hybrid with *HIF-1α* mRNA, thus exposing the AU-rich elements of *HIF-1α* and promoting its degradation (Rossignol et al. 2002).

1.10   Conclusions and Future Perspectives

The last few years have seen an increase in publications describing pervasive transcription in multicellular organisms, which results in the production of a large number of ncRNAs. Among these ncRNAs, the lncRNAs perhaps represent the most complex category of regulatory molecules in the multicellular organisms. So far, no sequence or structural similarity has been reported between those

lncRNAs shown to have a common mode of action. Due to the complexity and diversity of their sequences and their mechanisms of action, progress in the field of lncRNAs has been very slow. Nonetheless, lncRNAs have emerged as key regulators of developmental programs through their control of transcriptional and posttranscriptional gene regulatory pathways. The information from different biological contexts indicates that the functional roles of noncoding transcription and/or the ncRNAs are interpreted in different ways. While the noncoding transcriptional process often interferes with neighboring genes at the transcriptional level via TI mechanisms, on the other hand, ncRNA molecules take part in gene regulation from the single gene level to an entire chromosome via recruitment of chromatin modifying complexes *in cis* or *trans*. Transcription of an ncRNA through the regulatory region of its target gene can inhibit the assembly of transcription factors, or alternatively, the lncRNA can bind directly to key basal transcription factors, thus inhibiting PIC complex formation and leading to gene silencing. Similarly, the act of noncoding transcription or the ncRNA itself can negatively regulate the assembly of repressor complexes at the gene regulatory regions of target genes, thereby leading to transcriptional activation. LncRNAs also affect the target gene transcriptional output by targeting the repressor or activator complexes to the regulatory region of genes, a mechanism that is fairly well established in dosage compensation in mammals and *Drosophila*. At the posttranscriptional level, lncRNAs regulate the splicing, localization, stability, and translation of the target mRNAs by base-pairing with their target RNAs.

Although limited numbers of functional lncRNAs have been identified so far, the immense regulatory potential of lncRNAs in various developmental programs in multicellular organisms is already evident, emphasizing that a genome-wide characterization of functional lncRNAs is needed. Once the catalog of lncRNAs has been refined using biochemical and bioinformatic tools, genome-wide RNA interference (RNAi) screens, combined with powerful imaging techniques, such as those used in the identification of cell cycle regulatory proteins (Neumann et al. 2010; Walter et al. 2010), can be applied to characterize the roles of lncRNAs in different biological processes.

References

Affymetrix/Cold Spring Harbor Laboratory ENCODE Transcriptome Project (2009) Post-transcriptional processing generates a diversity of 5′-modified long and short RNAs. Nature 457:1028–1032

Alekseyenko AA, Larschan E, Lai WR, Park PJ, Kuroda MI (2006) High-resolution ChIP-chip analysis reveals that the *Drosophila* MSL complex selectively identifies active genes on the male X chromosome. Genes Dev 20:848–857

Alekseyenko AA, Peng S, Larschan E, Gorchakov AA, Lee OK, Kharchenko P, McGrath SD, Wang CI, Mardis ER, Park PJ, Kuroda MI (2008) A sequence motif within chromatin entry sites directs MSL establishment on the *Drosophila* X chromosome. Cell 134:599–609

Amaral PP, Mattick JS (2008) Noncoding RNA in development. Mamm Genome 19:454–492

Amaral PP, Dinger ME, Mercer TR, Mattick JS (2008) The eukaryotic genome as an RNA machine. Science 319:1787–1789

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297

Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136:215–233

Beltran M, Puig I, Peña C, García JM, Alvarez AB, Peña R, Bonilla F, de Herreros AG (2008) A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. Genes Dev 22:756–769

Bernstein E, Allis CD (2005) RNA meets chromatin. Genes Dev 19:1635–1655

Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, Rinn JL, Tongprasit W, Samanta M, Weissman S, Gerstein M, Snyder M (2004) Global identification of human transcribed sequences with genome tiling arrays. Science 306:2242–2246

Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET et al (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447:799–816

Blume SW, Meng Z, Shrestha K, Snyder RC, Emanuel PD (2003) The 5′-untranslated RNA of the human dhfr minor transcript alters transcription pre-initiation complex assembly at the major (core) promoter. J Cell Biochem 88:165–180

Brouwer R, Allmang C, Raijmakers R, van Aarssen Y, Egberts WV, Petfalski E, van Venrooij WJ, Tollervey D, Pruijn GJ (2001) Three novel components of the human exosome. J Biol Chem. 2;276(9):6177–84

Bumgarner SL, Dowell RD, Grisafi P, Gifford DK, Fink GR (2009) Toggle involving cis-interfering noncoding RNAs controls variegated gene expression in yeast. Proc Natl Acad Sci U S A 106:18321–18326

Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C, Kodzius R, Shimokawa K et al (2005) The transcriptional landscape of the mammalian genome. Science 309:1559–1563

Cawley S, Bekiranov S, Ng HH, Kapranov P, Sekinger EA, Kampa D, Piccolboni A, Sementchenko V, Cheng J, Williams AJ, Wheeler R, Wong B, Drenkow J, Yamanaka M, Patel S, Brubaker S, Tammana H, Helt G, Struhl K, Gingeras TR (2004) Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. Cell 116:499–509

Chaumeil J, Le Baccon P, Wutz A, Heard E (2006) A novel role for *Xist* RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes Dev 20:2223–2237

Cheung V, Chua G, Batada NN, Landry CR, Michnick SW, Hughes TR, Winston F (2008) Chromatin- and transcription-related factors repress transcription from within coding regions throughout the *Saccharomyces cerevisiae* genome. PLoS Biol 6:e277

Core LJ, Waterfall JJ, Lis JT (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 322:1845–1848

David L, Huber W, Granovskaia M, Toedling J, Palm CJ, Bofkin L, Jones T, Davis RW, Steinmetz LM (2006) A high-resolution map of transcription in the yeast genome. Proc Natl Acad Sci U S A 103:5320–5325

De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL, Natoli G (2010) A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol 8:e1000384

Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, Askarian-Amiri ME, Ru K, Soldà G, Simons C, Sunkin SM, Crowe ML, Grimmond SM, Perkins AC, Mattick JS (2008) Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res 18:1433–1445

Dobi KC, Winston F (2007) Analysis of transcriptional activation at a distance in *Saccharomyces cerevisiae*. Mol Cell Biol 27:5575–5586

Ebisuya M, Yamamoto T, Nakajima M, Nishida E (2008) Ripples from neighbouring transcription. Nat Cell Biol 10:1106–1113

Euskirchen G, Royce TE, Bertone P, Martone R, Rinn JL, Nelson FK, Sayward F, Luscombe NM, Miller P, Gerstein M, Weissman S, Snyder M (2004) CREB binds to multiple loci on human chromosome 22. Mol Cell Biol 24:3804–3814

Faghihi MA, Modarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE, Finch CE, St Laurent G 3rd, Kenny PJ, 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

Fitzpatrick GV, Soloway PD, Higgins MJ (2002) Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. Nat Genet 32:426–431

Gelbart ME, Kuroda MI (2009) *Drosophila* dosage compensation: a complex voyage to the X chromosome. Development 136:1399–1410

Gelbart ME, Larschan E, Peng S, Park PJ, Kuroda MI (2009) *Drosophila* MSL complex globally acetylates H4K16 on the male X chromosome for dosage compensation. Nat Struct Mol Biol 16:825–832

Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. Nat Rev Genet 10:94–108

Grewal SI, Jia S (2007) Heterochromatin revisited. Nat Rev Genet 8:35–46

Guaita S, Puig I, Franci C, Garrido M, Dominguez D, Batlle E, Sancho E, Dedhar S, De Herreros AG, Baulida J (2002) Snail induction of epithelial to mesenchymal transition in tumor cells is accompanied by MUC1 repression and ZEB1 expression. J Biol Chem 277:39209–39216

Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, Sukumar S, Chang HY (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 464:071–076

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458:223–227

He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW (2008) The antisense transcriptomes of human cells. Science 322:1855–1857

Hongay CF, Grisafi PL, Galitski T, Fink GR (2006) Antisense transcription controls cell fate in *Saccharomyces cerevisiae*. Cell 127:735–745

Imamura T, Yamamoto S, Ohgane J, Hattori N, Tanaka S, Shiota K (2004) Non-coding RNA directed DNA demethylation of Sphk1 CpG island. Biochem Biophys Res Commun 322:593–600

Kanduri M, Kanduri C, Mariano P, Vostrov AA, Quitschke W, Lobanenkov V, Ohlsson R (2002) Multiple nucleosome positioning sites regulate the CTCF-mediated insulator function of the H19 imprinting control region. Mol Cell Biol 22:3339–3344

Kanduri C, Thakur N, Pandey RR (2006) The length of the transcript encoded from the Kcnq1ot1 antisense promoter determines the degree of silencing. EMBO J 25:2096–2106

Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermüller J, Hofacker IL, Bell I, Cheung E, Drenkow J, Dumais E, Patel S, Helt G, Ganesh M, Ghosh S, Piccolboni A, Sementchenko V, Tammana H, Gingeras TR (2007a) RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science 316:1484–1488

Kapranov P, Willingham AT, Gingeras TR (2007b) Genome-wide transcription and the implications for genomic organization. Nat Rev Genet 8:413–423

Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES, Rinn JL (2009) Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A 106:11667–11672

Kim DH, Villeneuve LM, Morris KV, Rossi JJ (2006) Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. Nat Struct Mol Biol 13:793–797

Kino T, Hurt DE, Ichijo T, Nader N, Chrousos GP (2010) Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Sci Signal 3:ra8

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

Lefevre P, Witham J, Lacroix CE, Cockerill PN, Bonifer C (2008) The LPS-induced transcriptional upregulation of the chicken lysozyme locus involves CTCF eviction and noncoding RNA transcription. Mol Cell 32(1):129–139

Ligtenberg MJ, Kuiper RP, Chan TL, Goossens M, Hebeda KM, Voorendt M, Lee TY, Bodmer D, Hoenselaar E, Hendriks-Cornelissen SJ, Tsui WY, Kong CK, Brunner HG, van Kessel AG, Yuen ST, van Krieken JH, Leung SY, Hoogerbrugge N (2009) Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3′ exons of TACSTD1. Nat Genet 41:112–117

Maenner S, Blaud M, Fouillen L, Savoye A, Marchand V, Dubois A, Sanglier-Cianférani S, Van Dorsselaer A, Clerc P, Avner P, Visvikis A, Branlant C (2010) 2-D structure of the A region of Xist RNA and its implication for PRC2 association. PLoS Biol 8:e1000276

Malone CD, Hannon GJ (2009) Small RNAs as guardians of the genome. Cell 136:656–668

Mariner PD, Walters RD, Espinoza CA, Drullinger LF, Wagner SD, Kugel JF, Goodrich JA (2008) Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. Mol Cell 29:499–509

Martens JA, Laprade L, Winston F (2004) Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. Nature 429:571–574

Martens JA, Py Wu, Winston F (2005) Regulation of an intergenic transcript controls adjacent gene transcription in Saccharomyces cerevisiae. Genes Dev 19:2695–2704

Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A (2007) Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. Nature 445:666–670

Mattick JS (2004) RNA regulation: a new genetics? Nat Rev Genet 5:316–323

Mattick JS, Makunin IV (2006) Non-coding RNA. Hum Mol Genet 15:R17

Mayer C, Schmitz KM, Li J, Grummt I, Santoro R (2006) Intergenic transcripts regulate the epigenetic state of rRNA genes. Mol Cell 22:351–361

Meller VH, Rattner BP (2002) The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. EMBO J 21:1084–1091

Mercer TR, Dinger ME, Sunkin SM, Mehler MF, Mattick JS (2008) Specific expression of long noncoding RNAs in the mouse brain. Proc Natl Acad Sci U S A 105:716–721

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

Mondal T, Rasmussen M, Pandey GK, Isaksson A, Kanduri C (2010) Characterization of the RNA content of chromatin. Genome Res 20(7):899–907

Morris KV, Santoso S, Turner AM, Pastori C, Hawkins PG (2008) Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. PLoS Genet 4: e1000258

Murakami K, Oshimura M, Kugoh H (2007) Suggestive evidence for chromosomal localization of non-coding RNA from imprinted LIT1. J Hum Genet 52:926–933

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. Science 320:1344–1349

Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, Fraser P (2008) The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science 322:1717–1720

Navarro P, Chambers I, Karwacki-Neisius V, Chureau C, Morey C, Rougeulle C, Avner P (2008) Molecular coupling of Xist regulation and pluripotency. Science 321:693–695

Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A (2009) Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature 457:1038–1042

Neumann B, Walter T, Hériché JK, Bulkescher J, Erfle H, Conrad C, Rogers P, Poser I, Held M, Liebel U, Cetin C, Sieckmann F, Pau G, Kabbe R, Wünsche A, Satagopam V, Schmitz MH, Chapuis C, Gerlich DW, Schneider R, Eils R, Huber W, Peters JM, Hyman AA, Durbin R, Pepperkok R, Ellenberg J (2010) Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. Nature 464:721–727

Nobori T, Miura K, Wu DJ, Lois A, Takabayashi K, Carson DA (1994) Deletions of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers. Nature 368:753–756

Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, Nikaido I, Osato N, Saito R, Suzuki H, Yamanaka I, Kiyosawa H, Yagi K, Tomaru Y, Hasegawa Y, Nogami A, Schönbach C, Gojobori T, Baldarelli R, Hill DP, Bult C, Hume DA, Quackenbush J, Schriml LM, Kanapin A, Matsuda H, Batalov S, Beisel KW, Blake JA, Bradt D, Brusic V, Chothia C, Corbani LE, Cousins S, Dalla E, Dragani TA, Fletcher CF, Forrest A, Frazer KS, Gaasterland T, Gariboldi M, Gissi C, Godzik A, Gough J, Grimmond S, Gustincich S, Hirokawa N, Jackson IJ, Jarvis ED, Kanai A, Kawaji H, Kawasawa Y, Kedzierski RM, King BL, Konagaya A, Kurochkin IV, Lee Y, Lenhard B, Lyons PA, Maglott DR, Maltais L, Marchionni L, McKenzie L, Miki H, Nagashima T, Numata K, Okido T, Pavan WJ, Pertea G, Pesole G, Petrovsky N, Pillai R, Pontius JU, Qi D, Ramachandran S, Ravasi T, Reed JC, Reed DJ, Reid J, Ring BZ, Ringwald M, Sandelin A, Schneider C, Semple CA, Setou M, Shimada K, Sultana R, Takenaka Y, Taylor MS, Teasdale RD, Tomita M, Verardo R, Wagner L, Wahlestedt C, Wang Y, Watanabe Y, Wells C, Wilming LG, Wynshaw-Boris A, Yanagisawa M, Yang I, Yang L, Yuan Z, Zavolan M, Zhu Y, Zimmer A, Carninci P, Hayatsu N, Hirozane-Kishikawa T, Konno H, Nakamura M, Sakazume N, Sato K, Shiraki T, Waki K, Kawai J, Aizawa K, Arakawa T, Fukuda S, Hara A, Hashizume W, Imotani K, Ishii Y, Itoh M, Kagawa I, Miyazaki A, Sakai K, Sasaki D, Shibata K, Shinagawa A, Yasunishi A, Yoshino M, Waterston R, Lander ES, Rogers J, Birney E, Hayashizaki Y, FANTOM Consortium, RIKEN Genome Exploration Research Group Phase I & II Team (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature 420:563–573

Pandey RR, Ceribelli M, Singh PB, Ericsson J, Mantovani R, Kanduri C (2004) NF-Y regulates the antisense promoter, bidirectional silencing, and differential epigenetic marks of he Kcnq1 imprinting control region. J Biol Chem 279:52685–52693

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

Payer B, Lee JT (2008) X chromosome dosage compensation: how mammals keep the balance. Annu Rev Genet 42:733–772

Perocchi F, Xu Z, Clauder-Münster S, Steinmetz LM (2007) Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. Nucleic Acids Res 35:e128

Phillips JE, Corces VG (2009) CTCF: master weaver of the genome. Cell 137:1194–1211

Ponjavic J, Ponting CP, Lunter G (2007) Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. Genome Res 17:556–565

Prasanth KV, Spector DL (2007) Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. Genes Dev 21:11–42

Preker P, Nielsen J, Kammler S, Lykke-Andersen S, Christensen MS, Mapendano CK, Schierup MH, Jensen TH (2008) RNA exosome depletion reveals transcription upstream of active human promoters. Science 322:1851–1854

Redrup L, Branco MR, Perdeaux ER, Krueger C, Lewis A, Santos F, Nagano T, Cobb BS, Fraser P, Reik W (2009) The long noncoding RNA Kcnq1ot1 organises a lineage-specific nuclear domain for epigenetic gene silencing. Development 136:25–30

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129(7):1311–1323

Rossignol F, Vache C, Clottes E (2002) Natural antisense transcripts of hypoxia-inducible factor 1alpha are detected in different normal and tumour human tissues. Gene 299:135–140

Rossignol F, de Laplanche E, Mounier R, Bonnefont J, Cayre A, Godinot C, Simonnet H, Clottes E (2004) Natural antisense transcripts of HIF-1alpha are conserved in rodents. Gene 339:121–130

Samanta MP, Tongprasit W, Sethi H, Chin CS, Stolc V (2006) Global identification of noncoding RNAs in *Saccharomyces cerevisiae* by modulating an essential RNA processing pathway. Proc Natl Acad Sci U S A 103:4192–4197

Sanchez-Elsner T, Gou D, Kremmer E, Sauer F (2006) Noncoding RNAs of trithorax response elements recruit *Drosophila* Ash1 to Ultrabithorax. Science 311:1118–1123

Shamovsky I, Nudler E (2008) New insights into the mechanism of heat shock response activation. Cell Mol Life Sci 65:855–861

Shiraki T, Kondo S, Katayama S, Waki K, Kasukawa T, Kawaji H, Kodzius R, Watahiki A, Nakamura M, Arakawa T, Fukuda S, Sasaki D, Podhajska A, Harbers M, Kawai J, Carninci P, Hayashizaki Y (2003) Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proc Natl Acad Sci U S A 100:15776–15781

Sleutels F, Zwart R, Barlow DP (2002) The non-coding *Air* RNA is required for silencing autosomal imprinted genes. Nature 415:810–813

Straub T, Grimaud C, Gilfillan GD, Mitterweger A, Becker PB (2008) The chromosomal high-affinity binding sites for the *Drosophila* dosage compensation complex. PLoS Genet 4: e1000302

Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Spector DL (2009) MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. Genome Res 19:347–359

Sural TH, Peng S, Li B, Workman JL, Park PJ, Kuroda MI (2008) The MSL3 chromodomain directs a key targeting step for dosage compensation of the *Drosophila melanogaster* X chromosome. Nat Struct Mol Biol 15:1318–1325

Taft RJ, Pheasant M, Mattick JS (2007) The relationship between non-protein-coding DNA and eukaryotic complexity. Bioessays 29:288–299

Thakur N, Tiwari VK, Thomassin H, Pandey RR, Kanduri M, Göndör A, Grange T, Ohlsson R, Kanduri C (2004) An antisense RNA regulates the bidirectional silencing property of the Kcnq1 imprinting control region. Mol Cell Biol 24:7855–7862

Tufarelli C, Stanley JA, Garrick D, Sharpe JA, Ayyub H, Wood WG, Higgs DR (2003) Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. Nat Genet 34:157–165

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. Science 270:484–487

Walter T, Shattuck DW, Baldock R, Bastin ME, Carpenter AE, Duce S, Ellenberg J, Fraser A, Hamilton N, Pieper S, Ragan MA, Schneider JE, Tomancak P, Hériché JK (2010) Visualization of image data from cells to organisms. Nat Methods 7:S26–S41

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

Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, Hogenesch JB, Schultz PG (2005) A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. Science 309:1570–1573

Woo CJ, Kingston RE (2007) *HOTAIR* lifts noncoding RNAs to new levels. Cell 129:1257–1259

Wutz A, Rasmussen TP, Jaenisch R (2002) Chromosomal silencing and localization are mediated by different domains of *Xist* RNA. Nat Genet 30:167–174

Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Régnault B, Devaux F, Namane A, Séraphin B, Libri D, Jacquier A (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell 121:725–737

Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, Camblong J, Guffanti E, Stutz F, Huber W, Steinmetz LM (2009) Bidirectional promoters generate pervasive transcription in yeast. Nature 457(7232):1033–1037

Yakovchuk P, Goodrich JA, Kugel JF (2009) B2 RNA and Alu RNA repress transcription by disrupting contacts between RNA polymerase II and promoter DNA within assembled complexes. Proc Natl Acad Sci U S A 106:5569–5574

Yamada K, Lim J, Dale JM, Chen H, Shinn P, Palm CJ, Southwick AM, Wu HC, Kim C, Nguyen M, Pham P, Cheuk R, Karlin-Newmann G, Liu SX, Lam B, Sakano H, Wu T, Yu G, Miranda M, Quach HL, Tripp M, Chang CH, Lee JM, Toriumi M, Chan MM, Tang CC, Onodera CS, Deng JM, Akiyama K, Ansari Y, Arakawa T, Banh J, Banno F, Bowser L, Brooks S, Carninci P, Chao Q, Choy N, Enju A, Goldsmith AD, Gurjal M, Hansen NF, Hayashizaki Y, Johnson-Hopson C, Hsuan VW, Iida K, Karnes M, Khan S, Koesema E, Ishida J, Jiang PX, Jones T, Kawai J, Kamiya A, Meyers C, Nakajima M, Narusaka M, Seki M, Sakurai T, Satou M, Tamse R, Vaysberg M, Wallender EK, Wong C, Yamamura Y, Yuan S, Shinozaki K, Davis RW, Theologis A, Ecker JR (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. Science 302:842–846

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

Zhang LF, Huynh KD, Lee JT (2007) Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing. Cell 129:693–706

Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 322:750–756