Gene regulation by the act of long non-coding RNA transcription

Aleksandra E Kornienko, Philipp M Guenzl, Denise P Barlow and Florian M Pauler*

Abstract

Long non-protein-coding RNAs (lncRNAs) are proposed to be the largest transcript class in the mouse and human transcriptomes. Two important questions are whether all lncRNAs are functional and how they could exert a function. Several lncRNAs have been shown to function through their product, but this is not the only possible mode of action. In this review we focus on a role for the process of lncRNA transcription, independent of the lncRNA product, in regulating protein-coding-gene activity in cis. We discuss examples where lncRNA transcription leads to gene silencing or activation, and describe strategies to determine if the lncRNA product or its transcription causes the regulatory effect.

Keywords: Gene expression regulation, Histone modifications, LincRNA, LncRNA, Silencing, Transcriptional interference

LncRNAs - a new layer of genome regulatory information

It is now well appreciated that less than two percent of the human genome codes for proteins and the majority of the genome gives rise to non-protein-coding RNAs (ncRNAs) [1], which are predicted to play essential roles in a variety of biological processes [2,3]. The focus of this review is long ncRNAs (known as lncRNAs), which constitute the biggest class of ncRNAs with approximately 10,000 lncRNA genes so far annotated in humans [4]. LncRNAs are RNA polymerase II (RNAPII) transcripts that lack an open reading frame and are longer than 200 nucleotides. This size cut-off distinguishes lncRNAs from small RNAs such as microRNAs, piwi-interacting RNAs (piRNAs), small nucleolar RNAs (snoRNAs) and small interfering RNAs (siRNAs) and arises from RNA preparation methods that capture RNA molecules above this size. Although the function of most lncRNAs is unknown, the number of characterized lncRNAs is growing and many publications suggest they play roles in negatively or positively regulating gene expression in development, differentiation and human disease [2,5-10]. LncRNAs may regulate protein-coding (pc) gene expression at both the post-transcriptional and transcriptional level. Posttranscriptional regulation could occur by lncRNAs acting as competing endogenous RNAs to regulate microRNA levels as well as by modulating mRNA stability and translation by homologous base pairing, or as in the example of NEAT1 that is involved in nuclear retention of mRNAs [11]. In this review we focus on the regulation at the transcriptional level.

Modes of transcriptional regulation by lncRNAs

Regulation of transcription is considered to be an interplay of tissue and developmental-specific transcription factors (TFs) and chromatin modifying factors acting on enhancer and promoter sequences to facilitate the assembly of the transcription machinery at gene promoters. With a growing number of lncRNAs implicated in transcriptional gene regulation, this view may need refinement to include networks of tissue and developmental-stage specific lncRNAs that complement known regulators to tightly control gene expression and thereby organism complexity [12,13]. Transcriptional regulation by lncRNAs could work either in cis or in trans, and could negatively or positively control pc gene expression. LncRNAs work in cis when their effects are restricted to the chromosome from which they are transcribed, and work in trans when they affect genes on other chromosomes.

Regulation in trans

Some significant examples of lncRNAs that act in trans are those that can influence the general transcriptional output of a cell by directly affecting RNAPII activity (Figure 1a,b). One example is the 331 nucleotide 7SK...
IncRNA, which represses transcription elongation by preventing the PTEFβ transcription factor from phosphorylating the RNAPII carboxy-terminal domain (CTD) [14] (Figure 1a). Another example is the 178 nucleotide B2 IncRNA, a general repressor of RNAPII activity upon heat shock [15]. The B2 IncRNA acts by binding RNAPII and inhibiting phosphorylation of its CTD by TFIIH, thus disturbing the ability of RNAPII to bind DNA [16,17].

Regulation in trans can also act locus-specifically. While the ability of IncRNAs to act locus-specifically to regulate a set of genes was first demonstrated for imprinted genes where IncRNA expression was shown to silence from one to ten flanking genes in cis [18-20], IncRNAs that lie outside imprinted gene clusters, such as the HOTAIR IncRNA, were later found also to have locus-specific action. HOTAIR is expressed from the HOXC cluster and was shown to repress transcription in trans across 40 kb

![Figure 1](http://www.biomedcentral.com/1741-7007/11/59)

**Figure 1.** Long non-protein-coding RNAs (lncRNAs) act at different levels to regulate protein coding gene expression. lncRNAs can inhibit general protein-coding (pc) gene expression in trans (a) by preventing transcription factor (TF) activity (7SK IncRNA) or (b) by inhibiting RNAPII binding to DNA (B2 IncRNA). Xist IncRNA is transcribed from the X inactivation center (XIC) and inactivates a whole chromosome in cis (c) by recruiting epigenetic modifiers (EM). lncRNAs can regulate specific genes, acting in trans like HOTAIR (d) or in cis like HOTTIP (e) by directly recruiting epigenetic modifiers to certain genomic loci. In both cases the lncRNA binds EMs via a specific sequence or structure and targets them to promoter regions via DNA/RNA interaction elements to affect expression of the respective pc gene. Transcription of a lncRNA through a pc gene promoter or a cis-regulatory element (RE) affects pc gene expression in cis independent of the lncRNA product (f) by mechanisms discussed in the text. Both DNA strands are shown as separate boxes to indicate lncRNA transcription over the pc gene promoter in the antisense orientation. For details see text.
of the HOXD cluster [21]. HOTAIR interacts with Polycomb repressive complex 2 (PRC2) and is required for repressive histone H3 lysine-27 trimethylation (H3K27me3) of the HOXD cluster. Targeting of epigenetic modifiers (EMs) by lncRNAs provided a much sought after model to explain how EMs gain locus specificity (Figure 1d), and has since been suggested as a general mechanism for trans-acting lncRNAs [22,23].

Regulation in cis
In contrast to trans-acting lncRNAs, which act via their RNA product, cis-acting lncRNAs have the possibility to act in two fundamentally different modes. The first mode depends on a lncRNA product. The major example of general cis-regulation is induction of X inactivation by the Xist lncRNA in female mammals. Xist is expressed from one of the two X chromosomes and induces silencing of the whole chromosome [24] (Figure 1c). As an example of locus-specific regulation it has been proposed that enhancer RNAs activate corresponding genes in cis via their product [25]. A well-studied cis-acting lncRNA acting through its product is the human HOTTIP lncRNA that is expressed in the HOXA cluster and activates transcription of the functional transcription of another gene [29,30]. TI has been reported in unicellular and multicellular organisms [30]. Mechanistic details are still largely unclear, but TI could theoretically act at several stages in transcription: by influencing enhancer or promoter activity or by blocking RNAPII elongation, splicing or polyadenylation. All that would be required is that the RNA polymerase (RNAPII) initiated from an ‘interfering’ promoter traverses a ‘sensitive’ DNA regulatory sequence. TI has mainly been reported at overlapped promoters [31-35], but there are also examples where TI acts downstream of the promoter. In mouse, overlapping transcription controls polyadenylation choice of two imprinted genes [36,37]. In Saccharomyces cerevisiae, collisions between elongating antisense RNAPIIs can lead to stalling of both polyosomes that is resolved by ubiquitylation-directed proteolysis, and this has been proposed to be a regulatory mechanism [38]. However, it is unknown if RNAPII collisions occur sufficiently frequently in vivo in yeast or other organisms to offer a means of regulating convergent genes, or if this mechanism could lead to an interfering RNAPII eliminating its sensitive collision partner. Despite these examples, the most common reports of TI concern an overlapped promoter, and in the following sections we describe studies investigating the molecular mechanisms underlying interference at the promoter.

Transcriptional interference acting by promoter nucleosome repositioning
DNA in the nucleus is organized into chromatin with the organizational scaffold consisting of nucleosomes, each with two copies of H3, H4, H2A and H2B histones [39]. Nucleosomes can be densely packed, interfering with protein-DNA interactions, or relaxed, facilitating these interactions [40]. The transcription process, which generates single-stranded DNA as RNAPII progresses along a gene locus, can directly affect nucleosome positioning [41-43] (reviewed in [44,45]). Thus, lncRNA transcription could cause TI by depositing nucleosomes in a manner unfavorable for TF binding on promoters or enhancers. An example of this mechanism is the silencing of the yeast SER3 pc gene by transcriptional overlap by the SRG1 lncRNA (Figure 2a) [46]. SRG1 transcription increases nucleosome density at the overlapped SER3 promoter. Deletion of three transcription elongation factors that are associated with the elongating polymerase and are necessary for nucleosome repositioning (SPT16, SPT6, SPT2) [47-49] abolished the silencing effect without stopping transcription of the overlapping lncRNA SRG1 [50,51], indicating the necessity of chromatin reorganization for silencing. In contrast, deletion of epigenetic modifiers (such as SET1/2
Figure 2. Transcription interference-mediated silencing by chromatin changes. (a) Top: in yeast the absence of SRG1 lncRNA allows transcription machinery assembly at the SER3 protein coding gene promoter. Bottom: SRG1 lncRNA transcription causes dense nucleosome packing over the downstream SER3 pc gene promoter that blocks TF binding and pc gene expression. (b) Top: in yeast the absence of IRT1 lncRNA allows IME1 pc gene expression. Bottom: RNAPII transcribing the IRT1 lncRNA carries EMs that deposit repressive histone modifications at the IME1 promoter (EM1 - methyltransferases). These modifications allow the binding of other EMs that remove active histone modifications (EM2 - deacetylases) and cause a repressive chromatin environment that blocks TF binding leading to silencing. (c) Top: in a healthy human, LUC7L and HBA2 pc genes do not overlap and are both expressed. Bottom: a chromosomal deletion of the LUC7L transcriptional stop signal (red 'stop' box) causes transcription of the LUC7L pc gene through the promoter of the HBA2 pc gene. By an unknown mechanism this aberrant transcription causes DNA methylation and silencing of the HBA2 promoter. For details see Figure 1g and text.
Transcriptional interference acting by promoter histone modifications

Promoter associated nucleosomes carry post-translational histone tail modifications that reflect the activity state of the promoter and also influence accessibility of DNA binding factors involved in transcription [55]. Active gene promoters correlate with H3 and H4 acetylation and with H3K4me3, while inactive promoters do not and, in mammals, they also gain repressive histone marks such as H3K9me3 or H3K27me3. Some histone modifying enzymes have been shown to bind and travel with elongating RNAPII [56,57], so it is possible that lncRNA transcription can induce TI by affecting histone modifications at the promoter of an overlapped target gene. For example, in yeast the SET1/2 methyltransferases, which induce H3K4me2 and H3K36me3 in the body of transcribed genes, bind and travel with elongating RNAPII [58-60]. These modifications in turn recruit the SET3C/RPD3S histone deacetylase complexes to create a chromatin environment repressive for transcription initiation [61-63].

Two studies indicate that this is a mechanism used by lncRNAs to induce TI in yeast. In the first study the IME1 pc gene, which induces gametogenesis in diploid S. cerevisiae cells but is repressed in haploid cells, was shown to be silenced by the IRT1 lncRNA that overlaps its promoter [64]. Genetic experiments repositioning the IRT1 lncRNA distant from IME1 on the same chromosome showed that IRT1 transcriptional overlap of the IME1 promoter is necessary for silencing. Interestingly, the instability of the IRT1 lncRNA product and its non-specific cellular localization indicated the lncRNA product is unlikely to play a role in the silencing mechanism. Instead, IRT1 lncRNA transcription through the IME1 promoter reduced recruitment of the essential POG1 transcription factor, increased nucleosome density and induced the SET1/2 mediated cascades of histone modifications, which were shown to be necessary for silencing [64] (Figure 2b). In the second study lncRNA transcription was shown to be causative for silencing of the GAL1 and GAL10 genes, involved in galactose metabolism in S. cerevisiae. GAL10 and GAL1 are divergently transcribed from a bidirectional promoter. The 4 kb lncRNA, called GAL10-ncRNA, initiates in the body of the GAL10 gene, and is transcribed through the GAL10/GAL1 promoter antisense to the GAL10 gene. GAL10-ncRNA transcription induces SET2-mediated establishment of H3K36me3 along its gene body, thereby recruiting RPD3S-dependent deacetylation that resulted in reduced transcription factor binding and repression of the GAL1/GAL10 promoter [65]. Both SET3C and RPD3S are proposed to have a general role in repressing cryptic promoters within gene bodies [61,66] and a genome-wide study implied a role for SET3C in overlapping lncRNA-mediated silencing of a set of pc genes in yeast [66]. This indicates that the mechanism described above might be widely used to control gene expression in yeast. Although similar studies have not been described for the mammalian genome, H3K36me3 marks the body of transcribed genes in mammals, raising the possibility that such TI mechanisms could be conserved [56,57].

Transcriptional interference acting by promoter DNA methylation

In mammalian genomes DNA methylation is generally associated with silent CpG island promoters, but the majority of CpG island promoters remain methylation free independent of their expression status [67-69]. The process of \textit{de novo} methylation depends on the DNMT3A/3B methyltransferases and the catalytically inactive DNMT3L homologue and requires histones lacking H3K4me3, ensuring that active promoters remain methylation-free [70]. Notably, while DNA methylation at the promoter blocks transcription initiation, methylation in the gene body does not. Two important examples in humans based on genetic analyses indicate that DNA methylation can be involved in TI-induced silencing, although the causality between DNA methylation and silencing is still a matter of discussion [67]. One study of a patient with inherited α-thalassemia identified a deletion of the LUC7L 3’ end that allowed aberrant transcription of LUC7L through the downstream HBA2 gene, causing its silencing and the disease phenotype [71] (Figure 2c). Mouse models that mimicked the deleted genomic locus showed that the main cause of silencing was the acquisition of DNA methylation at the HBA2 promoter. Notably, DNA methylation acquisition was not simply the consequence of an inactive promoter, as removal of HBA2 transcription by deleting its TATA box did not induce methylation. The sequence of the

histone methyltransferases and SET3C/RPD3S deacetylases described later) did not affect silencing, showing that nucleosome silencing, but not changes in histone modifications, is responsible for repression. The experiments did not directly exclude a role for the SERG1 lncRNA product, but the silencing can be explained solely by the process of transcription [44,45]. TI by nucleosome repositioning may be a general mechanism in yeast, as the RNAPII elongation and chromatin organization factors responsible for SER3 silencing are also known to be involved in the suppression of transcription initiation from cryptic promoters within the body of actively transcribed genes [52,53]. Since genes controlling RNAPII elongation and chromatin organization are largely conserved, it is possible that lncRNAs could use similar nucleosome repositioning silencing in mammals. This is supported by the example that chromatin reassembly factors are necessary for silencing an HIV provirus when integrated into an actively transcribed host gene in a human cell system [54].
The LUC7L gene and thus the aberrant RNA product was also not essential for HBA2 silencing, as replacing the LUC7L gene body with another protein-coding gene did not remove the repressive effect. In a second example, a subset of Lynch syndrome patients display DNA methylation and inactivation of the mismatch repair MSH2 gene that correlates with aberrant transcription from the flanking EPCAM gene that carries a 3’ deletion [72].

In both these examples, the molecular details of methylation establishment and the mechanism by which the methylation machinery targets the overlapped promoter are yet unknown. However, the data so far show that it is a cis-acting mechanism as only the allele carrying the deletion silences the overlapped protein-coding gene. In addition, although a role for the aberrant RNA product was not excluded, it appears unlikely that mutation-induced transcription of two independent intergenic chromosomal regions in the described diseases produces IncRNA products with similar repressive functions. Interestingly, the silencing of imprinted pc genes by IncRNAs is also often correlated with the gain of DNA methylation on the silent pc gene promoter [73]. In the case of the Igf2r gene, this DNA methylation mark is not necessary for initiation or maintenance of the silent state but seems to play a role in re-enforcing the silent state [35,74].

Transcriptional interference in the absence of chromatin changes at the silenced promoter

In addition to RNAPII acting as a carrier of chromatin modifying enzymes, other Ti models predict that RNAPII from one promoter traversing across another promoter can interfere with its activity without introducing chromatin changes [30,75,76]. An indication that such a mechanism can be used by IncRNAs in mammals comes from a study that used a genetic approach to dissect the silencing function of the imprinted mouse Airn IncRNA [77,78]. Airn is an inefficiently spliced 118 kb IncRNA expressed on paternally inherited chromosomes that overlaps and silences the promoter of the Igf2r pc gene - a dose-sensitive and essential embryonic growth suppressor [18,79] (Figure 3a). To determine if Airn transcription or its IncRNA product were required for silencing, homologous recombination in embryonic stem cells was used to shorten the length of Airn, either before or after the Igf2r promoter, by insertion of a polyadenylation cassette [35]. Notably, only shortened Airn variants that traversed the Igf2r promoter induced silencing. Furthermore, while Igf2r silencing is normally accompanied by DNA methylation, repressive histone marks and chromatin compaction of the silent Igf2r promoter [80,81], Igf2r silencing was not dependent on DNA methylation - in contrast to the silencing of HBA2 by aberrant LUC7L transcription described above. Instead, Airn transcriptional overlap interfered with the accumulation of functional RNAPII on the Igf2r promoter in the presence of open chromatin [35]. Additional support for Igf2r silencing by Airn transcriptional interference is provided by genetic experiments that used an inducible Airn silencing by by Airn transcriptional interference is provided by genetic experiments that used an inducible Airn silencing by by Airn transcriptional interference is provided by genetic experiments that used an inducible Airn silencing by by Airn transcriptional interference is provided by genetic experiments that used an inducible Airn silencing by by Airn transcriptional interference is provided by genetic experiments that used an inducible Airn silencing by by Airn transcriptional interference is provided by genetic experiments that used an inducible Airn silencing by by Airn transcriptional interference is provided by genetic experiments that used an inducible Airn silencing by
model, it was shown that Kcnq1ot1 silences at least one gene by regulating chromatin flexibility and access to enhancers [89]. This is consistent with a two-step model whereby lncRNA transcription initiates silencing of non-overlapped genes by enhancer interference, then repressive histone modifying enzymes maintain that silencing.

**Figure 3. Transcription interference-mediated silencing without chromatin changes.** (a) Top: a wild-type maternal allele does not express Airn lncRNA as its promoter is repressed by a DNA methylation imprint, thus allowing the Igf2r gene to be active. Middle: on the wild-type paternal allele Airn transcription overlaps with and silences the Igf2r pc gene promoter, independent of the Airn lncRNA product. The silent Igf2r promoter is marked by increased nucleosome density and DNA methylation in the absence of active histone modifications. Bottom: increased nucleosome density, loss of active histone marks and DNA methylation are not necessary for Igf2r repression as demonstrated by the FAP allele that moved the Airn promoter close to the Igf2r promoter and silenced Igf2r in the absence of repressive chromatin features. (b) Top: a hypothetical enhancer activates a pc gene by direct long-range DNA interactions. Bottom: transcription of a lncRNA overlapping the enhancer interferes with the DNA interaction and thereby silences the pc gene. For details see Figure 1g, Figure 2d and text.

**lncRNA transcription creating a permissive chromatin environment**
Enhancers are genetic elements that bind transcription factors facilitating transcription machinery assembly at nearby promoters [90,91]. RNAPII transcripts up to 2 kb long are transcribed bi-directionally from some neuronal
enhancers (termed enhancer or eRNAs) [91,92]. Transcription of eRNAs positively correlated with expression of nearby mRNAs and a model was proposed, but not yet experimentally tested, in which their transcription establishes a chromatin landscape that supports enhancer function (Figure 4a). IncRNA transcription, either by opening chromatin or inhibiting repressor protein binding, could similarly result in gene or locus activation. One example of this is the process of V(D)J recombination, which joins elements of the V, D and J multigene family by chromosomal rearrangements to create functional B cell immunoglobulins and T cell receptors [93] (Figure 4b). The V, D and J genes lie next to each other on the same chromosome and antisense intergenic transcription through these genes is detected prior to the recombination process [94]. Genetic experiments have shown that intergenic IncRNA transcription is required for both B and T cell V(D)J recombination [95,96]. Similar correlations between intergenic transcription and gene expression were observed for the mouse β-globin locus [97] where promoter deletion experiments showed that IncRNA transcription was responsible for stable, active and hyper-accessible chromatin [98].

**IncRNA transcription and locus activation**

Other examples indicate that IncRNA transcription activates gene expression by blocking access of repressor complexes to chromatin. In *Drosophila*, intergenic non-coding transcription at the *BITHORAX* complex (BX-C) is implicated in reversing POLYCOMB group (PCG)-mediated gene silencing and is correlated with an active chromatin state [99]. This mode of action was later suggested to be a general mechanism where the act of transcription serves as an epigenetic switch that relieves PCG-mediated gene silencing by recruiting epigenetic modifiers to induce gene expression and generate stable and heritable active chromatin [100]. In line with this hypothesis, intergenic transcription through PCG response elements (PREs) in the BX-C cluster is not only found during embryogenesis but also in late stage larvae, indicating that continuous transcription is required to keep genes active [101]. In mouse and human, a similar role for PRE transcription has been proposed. An analysis of IncRNA transcription in the human *HOXA* cluster revealed a positive correlation between IncRNA transcription and the loss of PCG/chromatin interactions that precedes *HOXA* gene activation [102]. Additionally, IncRNAs have been identified at promoter regions of PCG-regulated genes in mouse cells; while their role is not yet clear, it has been suggested that they either promote or interfere with PCG binding at target genes [103,104].

A further example of a lncRNA mediating chromatin opening was described at the *S. cerevisiae PHO5* gene. Transcription of an antisense IncRNA that initiates near the 3’end of *PHOS* and overlaps its gene body and promoter is associated with rapid activation of *PHOS* by enabling nucleosome eviction. Biochemical inhibition of RNA polymerase II elongation as well as genetic disruption of IncRNA elongation demonstrated a direct role in *PHOS* activation [105]. The association of IncRNA transcription with gene activation needs, however, to be considered within the framework that most protein-coding gene promoters in yeast and mammalian cells give rise to a bidirectional antisense IncRNA transcript [106,107]. To date it is unclear if promoter-associated bidirectional IncRNAs represent spurious transcription in the context of open chromatin [108,109] or is required to maintain open chromatin. In the latter case enhanced TF binding ensures accessible chromatin that allows more constant pc gene expression within a cell population [110] (Figure 4c).

### Strategies for distinguishing a role for the lncRNA product from that of its transcription

Following genome-wide lncRNA mapping, functional studies so far have mainly focused on lncRNA products [7,111]. As it becomes clear that IncRNAs can act through their transcription, it is important to identify strategies to determine the function and mode of action of each particular IncRNA. One common starting point to determine IncRNA function has been RNA interference (RNAi)-mediated knockdown, despite long-standing observations that the RNAi machinery in mammalian cells is located in the cytoplasm [112]. While there is evidence that some RNA-induced silencing complex (RISC) components are found in the nucleus, functional complexes are specifically loaded in the cytoplasm, prohibiting the application of RNAi strategies for nuclear localized IncRNAs [113]. In contrast, antisense oligonucleotides (ASO) that work via an RNaseH-dependent pathway will deplete nuclear-localized IncRNAs [114,115]. However, three additional points of caution should be noted. First, non-specific effects arising from nuclear transfection reagents [116] have confused some observations. One critical validation step for knockdown studies would be a rescue experiment in which the IncRNA, modified to be invulnerable to the knockdown, is expressed as a transgene under the same transfection conditions [111]. Second, some results have highlighted major differences when functional studies used post-transcriptional depletion strategies in cell lines in contrast to genetic studies in the organism. Notable examples are *Neat1* [117], *Malat1* [116,118,119] and *Hotair* [120] where studies of mice carrying genetically disrupted alleles of these three IncRNAs failed to reproduce phenotypes deduced from cell lines following RNAi, ASO or over-expression studies. Third, while knockdown experiments may elucidate the function of IncRNAs acting through their product, the function of cis-acting IncRNAs that depend only on transcription will not be disturbed.
Features such as subcellular localization, half-life and steady-state abundance would form a good basis to allow functional tests to be designed. In addition, knowledge of the IncRNA splicing efficiency, conservation of splicing pattern in multiple tissues and species, an estimation of transcript repeat content and, finally, an accurate mapping of IncRNA 5' and 3' ends are essential preliminary steps. We have previously proposed that a subclass of IncRNAs, ‘macro’ IncRNAs, show RNA biology hallmarks such as inefficient splicing, extreme length, high repeat content, lack of conservation and a short half-life. These features are also indicators that the IncRNA product is less important than the act of transcription [121]. Once RNA biology features are known, experiments can be designed to distinguish between a role for the IncRNA product or its transcription.

From the caveats of posttranscriptional knockdown experiments described above, it becomes clear that genetic
strategies are optimal for testing lncRNA function. These strategies include manipulating the endogenous locus to delete the promoter or the whole gene or to shorten its length using inserted polyadenylation signals, as described for several examples above. This may appear a formidable task with the appreciation that lncRNAs in the human genome may outnumber protein-coding genes [4]; however, suitable cell systems already exist. These include the use of haploid cell lines with transcriptional stop signal insertions in most human genes [122], gene targeting by engineered zinc-finger nucleases [123] or CRISPR systems [124] or the use of mouse embryonic stem cells that have efficient rates of homologous targeting [125,126].
These genetic strategies could be applied to determine if the lncRNA is functional and if its function requires the lncRNA product or only depends on the act of transcription (Figure 5). Once these answers are obtained, it will be useful to test whether additional chromatin features are involved. This could include chromatin accessibility assays to address nucleosome density in the regulated gene; and mapping of histone modifications and DNA methylation, and of the presence of RNAPII and other transcription machinery components. These studies have been made easier in the mouse and human genome due to the publicly available ENCODE data [127]. As lncRNA identification becomes easier due to improved sequencing and bioinformatics tools, the number of annotated lncRNAs will likely increase sharply [4,128]. It is therefore a high priority to determine which lncRNAs are functional and which represent spurious transcription [109,129]. To date only a relatively small number of mammalian lncRNAs have clearly been shown to regulate gene expression and most attention has centered on lncRNAs that act through their transcription product [23]. With the recent demonstration that for some mammalian lncRNAs the act of their transcription is sufficient for function [35], it becomes clear that there can be a number of lncRNAs acting in a similar way. If the above described findings and approaches are used as guidelines, many new lncRNAs regulating genes by the act of transcription are likely to be discovered.

Acknowledgements
We thank Quanah Hudson and Federico Santoro for comments on the manuscript. The authors are partly supported by the Austrian Science Fund FWF SFB-F43 and FWF W1207-BO9. PG is recipient of a DOC Fellowship of the Austrian Academy of Sciences.

Published: 30 May 2013

References
1. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Willems BA, Zalesi C, Rozowsky J, Roder M, Kokocinski F, Abdelhamid RF, Alotto T, Antoshechkin I, Baer MT, Bell K, Chen Y, Chladek J, Curado M, et al. Landscape of transcription in human cells. Nature 2012, 489:101–108.
2. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. Genes Dev 2009, 23:1494–1504.
3. Pauli A, Rinn JL, Schier AF. Non-coding RNAs as regulators of embryogenesis. Nat Rev Genet 2011, 12:136–149.
4. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Neklous DG, Lageman J, Veevaert L, Ruan X, Ruan Y, Lassmann T, Campani P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shekharik L, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigó R. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 2012, 22:1775–1789.
5. Taft RJ, Pang KC, Mercer TR, Dinger M, Mattick JS. Non-coding RNAs: regulators of disease. J Pathol 2010, 220:126–139.
6. Huarte M, Rinn JL. Large non-coding RNAs: missing links in cancer? Hum Mol Genet 2010, 19:R152–R161.
7. Guttman M, Donahghey J, Carey BW, Garber M, Grenier JK, Munson G, Young C, Lucas AB, Ach R, Bruhn L, Yang X, Amst I, Meissner A, Regev A, Rinn JL, Root DE, Landor ES. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature 2011, 477:295–300.
8. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Botsios P, Kong B, Li R, West RB, van de Vijler MJ, Sukumar S, Chang HY. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 2010, 464:1071–1076.
9. Premsner Jr, Iyer MK, Balbin OA, Dhanasekaran SA, Qao C, Brenner KC, Laxman B, Asangani IA, Gassio CS, Kominsky HD, Cao X, Jinx X, Wang X, Siddiqui JI, Wei JT, Robinson D, Iyer HK, Palanisamy N, Maher CA, Chinman AM. Transcriptionte sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lncRNA implicated in disease progression. Nat Biotechnol 2011, 29:742–749.
10. Yap KL, Li S, Munoz-Gabello AM, Ragusi S, Zeng L, Mijtaba S, Gil J, Walsh MJ, Zhou MM. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Mol Cell 2010, 38:662–674.
11. Yoon JH, Abdelmohsen K, Gorospe M. Posttranscriptional gene regulation by long noncoding RNA. J Mol Biol 2012, pii:00022-2836(12)00896-0.
12. Mattick JS. Deconstructing the dogma: a new view of the evolution and genetic programming of complex organisms. Annu N Y Acad Sci 2009, 1178:29–46.
13. Mattick JS, Taft RJ, Faulkner GJ. A global view of genomic information—moving beyond the gene and the master regulator. Trends Genet 2010, 26:21–28.
14. Peterlin BM, Brodie JE, Price DH. 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. Wiley Interdiscip Rev RNA 2012, 3:92–103.
15. Espinoza CA, Allen TA, Hieb AR, Kugel JF, Goodrich JA. B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. Nat Struct Mol Biol 2004, 11:822–829.
16. Espinoza CA, Goodrich JA, Kugel JF. Characterization of the structure, function, and mechanism of B2 RNA, an ncRNA repressor of RNA polymerase II transcription. RNA 2007, 13:583–596.
17. Yakovichuk P, Goodrich JA, Kugel JF. B2 RNA represses TFIIH phosphorylation of RNA polymerase II. Transcription 2011, 2:45–49.
18. Sleuteels F, Zwart R, Barlow DP. The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature 2002, 415:810–813.
19. Mancini-Onnado D, Steele SJ, Levonse JM, Ingram RS, Tiglihammer SM. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. Genes Dev 2006, 20:1268–1282.
20. Williamson CM, Ball ST, Dawson C, Mehta S, Beechey CV, Fray M, Teboul L, Dear TN, Kelsey G, Peters J. Uncoupling antisense-mediated silencing and DNA methylation in the imprinted Gnas cluster. PLoS Genet 2011, 7:e1001347.
21. Rinn JL, Kiresuk M, Wang JK, Szaqqlo SL, Xu X, Brugmann SA, Goodnough LH, Helens JA, Farnham PJ, Segel E, Chang HY. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 2007, 129:1311–1323.
22. Ng SY, Johnson R, Stanton LW. Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. EMBO J 2011, 31:522–533.
23. Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. Nat Reviews 2012, 482:339–346.
24. Wutz A. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nat Rev Genet 2011, 12:542–553.
25. Örom UA, Derrien T, Berenger M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytynicki M, Notredame C, Huang Q, Guigo R, Sheikhatr M. Long non-coding RNAs with enhancer-like function in human cells. Cell 2010, 143:46–58.
26. Wang KC, Yang YW, Liu B, Sanyal A, Cores-Zimmerman R, Chen Y, Lajoie BR, Protacio A, Flynn RA, Gupta RA, Wysocka J, Lei M, Dekker J, Helms JA, Chang HY. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 2011, 472:120–124.
27. Magistri M, Faghihi MA, St Laurent G, 3rd, Wahskowitz C. Regulation of chromatin structure by long noncoding RNAs: focus on natural antisense transcripts. Trends Genet 2012, 28:389–396.
28. Heo JB, Sung S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science 2011, 331:67–70.
29. Shearwin RE, Callen BP, Egan JB. Transcriptional interference—a crash course. Trends Genet 2005, 21:339–345.
Nucleosome positioning: how is it transcription from within coding regions throughout the Saccharomyces cerevisiae gene cluster.

Transcription regulation by the noncoding RNA SRG1 requires Spt2-Abrd-A regulation by the iab-8 noncoding RNA.

Transcription of bxd noncoding RNAs promoted by trithorax represses Ubx in cis by transcriptional interference.

Gene Cell 2008, 12:1191–1146.

J Virol 2011, 85:5187–3202.

Cell Res 2011, 21:381–395.

EMBO Rep 2009, 10:1213–1219.

Crit Rev Biochem Mol Biol 2012, 47:464–482.

Proc Natl Acad Sci U S A 2009, 106:18321–18326.

Mol Cell 2003, 11:709–719.

Nat Cell Biol 2006, 8:2314–2329.

EMBO J 2012, 31:1099–1109.

Cell Biol 2003, 241:2250–2260.

Nat Rev Genet 2003, 4:175–186.

Cell 2008, 135:1221–1231.

Science 2004, 304:1221–1226.

Cell 2002, 110:1123–1133.

Nat Cell Biol 2004, 6:733–773.

Proc Natl Acad Sci U S A 2006, 103:9559–68.

Cell 2003, 115:889–900.

Cell 2003, 112:157–168.

Cell 2002, 110:233–248.

Cell 2004, 118:435–446.

EMBO J 2011, 30:5106–5117.

Cell 2003, 112:337–349.

Cell 2003, 114:323–334.

Nat Cell Biol 2006, 8:237–247.

Cell 2007, 131:577–589.

Nat Rev Genet 2003, 4:503–516.

Cell 2003, 115:1027–1038.

Nat Rev Genet 2003, 4:455–458.

Nat Rev Genet 2003, 4:319–320.

Cell 2003, 115:481–493.

Cell 2003, 115:121–134.

Curr Opin Genet Dev 2003, 13:319–325.

Cell 2003, 115:955–967.

Hum Mutat 2002, 20:557–567.

Cell 2003, 115:51–65.

Cell 2003, 115:1241–1253.

Nat Rev Genet 2003, 4:25–35.

Science 2003, 301:271–274.

Sci Signal 2010, 3:ra50.

Cell 2007, 128:1221–1230.

EMBO J 2006, 25:3726–3734.

Nature 2004, 430:571–574.

Cell 2002, 111:59–69.

Cell 2002, 111:907–917.

J Biol Chem 2008, 283:26979–26989.

Hum Mol Genet 2003, 12:2353–2364.

Proc Natl Acad Sci U S A 2007, 104:3863–3868.

Cell 2003, 113:961–972.

EMBO J 2002, 21:3513–3523.

EMBO J 2003, 22:2197–2207.

Hum Mol Genet 2003, 12:1471–1479.

Cell 2003, 115:281–293.

Nat Cell Biol 2004, 6:73–77.

Cell 2003, 115:1363–1374.

Cell 2003, 115:1265–1278.

Hum Mol Genet 2003, 12:529–539.

Cell 2003, 115:759–770.

Cell 2003, 115:1469–1472.

Cell 2003, 115:1125–1135.

Cell 2003, 115:975–991.

Cell 2003, 115:129–143.

Cell 2003, 115:1233–1244.

Cell 2003, 115:1209–1221.

Cell 2003, 115:1181–1200.

Cell 2003, 115:1209–1221.

Cell 2002, 111:597–608.

Cell 2003, 115:1241–1253.

Cell 2002, 111:597–608.

Cell 2003, 115:1241–1253.

Cell 2003, 115:1241–1253.

Cell 2003, 115:1241–1253.

Cell 2003, 115:1241–1253.

Cell 2003, 115:1241–1253.

Cell 2003, 115:1241–1253.

Cell 2003, 115:1241–1253.

Cell 2003, 115:1241–1253.
77. Barlow DP: Genomic imprinting: a mammalian epigenetic discovery model. Annu Rev Genet 2011, 45:379–403.
78. Koemer MV, Paulaer FM, Huang R, Barlow DP: The function of non-coding RNAs in genomic imprinting. Development 2009, 136:1771–1783.
79. Wang ZQ, Fung MR, Barlow DP, Wagner EF: Regulation of embryonic growth and lysosomal targeting by the imprinted Igf2/Mirp gene. Nature 1994, 372:464–467.
80. Paulaer FM, Stricker SH, Warczok KE, Barlow DP: Long-range DNAs I hypersensitivity mapping reveals the imprinted Igf2r and Air promoters share cis-regulatory elements. Genome Res 2005, 15:1379–1387.
81. Stoger R, Kubicka P, Liu CG, Kafri T, Razin A, Cedar H, Barlow DP: Maternal-specific methylation of the imprinted mouse Igf2 locus identifies the expressed locus as carrying the imprinting signal. Cell 1993, 73:561–71.
82. Racaneli AC, Turner FB, Xie LY, Taylor SM, Moran RG: A mouse gene that coordinates epigenetic controls and transcriptional interference to achieve tissue-specific expression. Mol Cell Biol 2008, 28:836–848.
83. Palmer AC, Ahlgren-Berg A, Egan JB, Dodd IB, Shearwin KE: Potent transcriptional interference by pausing of RNA polymerases over a downstream promoter. Mol Cell 2009, 34:545–555.
84. Adelman K, Li JT: Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat Rev Genet 2012, 13:720–731.
85. Levine M: Paused RNA polymerase II as a developmental checkpoint. Cell 2011, 145:502–511.
86. Nagano T, Mitchell JA, Sanz LA, Paulaer FM, Ferguson-Smith AC, Feil R, Fraser SD, Peters AH: The Air noncoding RNA epigenetically silences transcription by interacting with chromatin. Science 2008, 322:1717–1720.
87. Mager J, Montgomery ND, de Villena FP, Magnuson T: The function of non-coding transcripts are modulators of noise. Genome Biol Evol 2011, 3:330–337.
88. Kudryk I, Shimamura A, Jan CH, Sive H, Bartel DP: Conserved function of IncRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell 2011, 147:1537–1550.
89. Zeng Y, Cullen BR: RNA interference in human cells is restricted to the cytoplasm. RNA 2002, 8:855–860.
90. Ohl T, Muetze J, Svoboda P, Schwille P: Intracellular localization and routing of mRNAs and RNA pathway components. Curr Top Med Chem 2012, 12:79–88.
91. Ideue T, Hino K, Kitao S, Yokoi T, Hirose T: Efficient oligonucleotide-mediated degradation of nuclear noncoding RNAs in mammalian cultured cells. RNA 2009, 15:1578–1587.
92. Tse MT: Antisense therapeutics: Nuclear RNA more susceptible to knockdown. Nat Rev Drug Discov 2012, 11:674.
93. Zhang B, Arun G, Yao YS, Lazar Z, Huang G, Bhattachajee G, Xiao Y, Booth CJ, Wu J, Zhang C, Spector DL: The IncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. Cell Rep 2012, 2:1111–12.
94. Nakagawa S, Naganuma T, Shi Qi, Hirose T: Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. Cell 2011, 145:311–329.
95. Elfmann M, Gutschner T, Hämmerle M, Günther S, Caudron-Herger M, Groß M, Schirmacher P, Rippe K, Zörmig M, Dederichs S: Loss of the abundant non-coding RNA MALAT1 is compatible with life and development. RNA Biol 2012, 9:1076–1087.
96. Nakagawa S, Ip Jr, Shi Qi, Tripeth V, Zong X, Hirose T, Prasanth KV: Malat1 is not an essential component of nuclear speckles in mice. RNA 2012, 18:1487–1499.
97. Schorderet P, Duboule D: Structural and functional differences in the long non-coding RNA hotair in mouse and human. PLoS Genet 2011, 7:e1002071.
98. Guend PM, Barlow DP: Macro IncRNAs: A new layer of cis-regulatory information in the mammalian genome. RNA Biol 2012, 9:731–741.
99. Carette JE, Guinaudais CP, Wuerthth, I, Boinen VA, Varadoranjan M, Sun C, Bell G, Yuan B, Mueller MK, Nilman SM, Ploegh HL, Brummelkamp TR: Global gene disruption in human cells to assign genes to phenotypes by deep sequencing. Nat Biotechnol 2011, 29:542–546.
100. Wirt SE, Porteus MH: Development of nucleoside-mediated site-specific genome modification. Curr Opin Immunol 2012, 24:599–616.
101. Mall P, Yang L, Esvitt KM, Aach J, Guett M, Dicarlo JE, Nonville JE, Church GM: RNA-Guided Human Genome Engineering via Cas9. Science 2013, 339:823–826.
102. Latos PA, Stricker SH, Steenpass L, Paulaer FM, Huang R, Senergin BH, Regha K, Koerner MV, Warczok KE, Unger C, Barlow DP: An in vitro ESC imprinting model shows that imprinted expression of the Igf2r gene arises from an allele-specific expression bias. Development 2009, 136:457–468.
126. Kohama C, Kato H, Numata K, Hirose M, Takemasa T, Ogura A, Kiyosawa H: ES cell differentiation system recapitulates the establishment of imprinted gene expression in a cell-type-specific manner. *Hum Mol Genet* 2012, 21:1391–1401.

127. Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, Wong MC, Maddren M, Fang R, Heitner SG, Lee BT, Barber GP, Harte RA, Diekhans M, Long JC, Wilder SP, Zweig AS, Karolchik D, Kuhn RM, Haussler D, Kent WJ: ENCODE Data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res* 2013, 41:D56–63.

128. Loven J, Orlando DA, Sigova AA, Lin CY, Rahl PB, Burge CB, Levens DL, Lee TI, Young RA: Revisiting global gene expression analysis. *Cell* 2012, 151:476–482.

129. Clark MB, Mattick JS: Long noncoding RNAs in cell biology. *Semin Cell Dev Biol* 2011, 22:366–376.

doi:10.1186/1741-7007-11-59

Cite this article as: Kornienko AE et al: Gene regulation by the act of long non-coding RNA transcription. *BMC Biology* 2013 11:59.