nature reviews molecular cell biology

https://doi.org/10.1038/s41580-023-00694-9

Review article

Check for updates

Transcription regulation by long non-coding RNAs: mechanisms and disease relevance

Jorge Ferrer [1,2,3] ✉ & Nadya Dimitrova [4] ✉

Abstract

Long non-coding RNAs (lncRNAs) outnumber protein-coding transcripts, but their functions remain largely unknown. In this Review, we discuss the emerging roles of lncRNAs in the control of gene transcription. Some of the best characterized lncRNAs have essential transcription *cis*-regulatory functions that cannot be easily accomplished by DNA-interacting transcription factors, such as *XIST*, which controls X-chromosome inactivation, or imprinted lncRNAs that direct allele-specific repression. A growing number of lncRNA transcription units, including *CHASERR*, *PVT1* and *HASTER* (also known as *HNF1A-AS1*) act as transcription-stabilizing elements that fine-tune the activity of dosage-sensitive genes that encode transcription factors. Genetic experiments have shown that defects in such transcription stabilizers often cause severe phenotypes. Other lncRNAs, such as *lincRNA-p21* (also known as *Trp53cor1*) and *Maenli* (*Gm29348*) contribute to local activation of gene transcription, whereas distinct lncRNAs influence gene transcription in *trans*. We discuss findings of lncRNAs that elicit a function through either activation of their transcription, transcript elongation and processing or the lncRNA molecule itself. We also discuss emerging evidence of lncRNA involvement in human diseases, and their potential as therapeutic targets.

Sections

[1]Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. [2]Centro de Investigación Biomédica en Red de Diabetes y Enfermedades Metabólicas Asociadas (CIBERDEM), Madrid, Spain. [3]Department of Metabolism, Digestion and Reproduction, Imperial College London, London, UK. [4]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT, USA. ✉e-mail: jorge.ferrer@crg.eu; nadya.dimitrova@yale.edu

Introduction

In the mid-1980s, transcripts that lacked obvious open reading frames were identified in the *Drosophila melanogaster* bithorax locus[1]. At the time, the authors of the study hypothesized that these transcripts might be by-products of enhancer activity, that their transcription could regulate adjacent genes, that they might regulate splicing or translation in *trans*, or that they might encode atypical polypeptides. Others speculated that bithorax non-coding RNAs promote or inhibit the activity of nearby enhancers[2]. Decades later, these conjectures have been proven to be remarkably insightful, as careful analyses of individual long non-coding RNAs (lncRNAs) have largely confirmed each of these models.

lncRNAs are defined as transcripts that are longer than 500 nucleotides and do not encode proteins[3]. In practice, this is a catch-all definition that encompasses different types of transcripts that do not have an obvious protein-coding potential, and some lncRNAs defined in this fashion have turned out to encode micropeptides[4,5]. The number of annotated lncRNAs has been growing steadily, and currently includes more than 20,000 lncRNAs[6], most of which have no known function. A relatively small number of lncRNAs, however, has been linked to a wide range of biological processes through disparate molecular mechanisms[3].

Many of the best characterized lncRNAs have been implicated in the regulation of gene transcription[7,8], often through defined molecular interactions. These studies have raised several crucial questions. Importantly, are there lncRNA types that execute distinct regulatory functions? Considering our current models of gene transcription, which are largely shaped by our understanding of how regulatory protein complexes are recruited to *cis*-acting DNA elements, what is the purpose of lncRNAs that regulate gene transcription? Which functions do lncRNAs carry out that cannot be easily enacted by DNA-interacting proteins, and what are the underlying mechanisms?

In this Review, we discuss the accumulating evidence that points to major modes through which lncRNAs participate in the regulation of gene transcription. For the purpose of this Review, we have considered lncRNA transcription units regardless of whether their regulatory function is conferred by the lncRNA molecules, the process of lncRNA transcription or the lncRNA promoter. Given that experimental perturbations of lncRNAs can theoretically disrupt other established transcription-regulating components, we specifically focused on lncRNA functions that cannot be easily ascribed to conventional enhancers, promoters or regulatory proteins. We emphasize functions and mechanisms that have been shown to operate at more than one lncRNA locus, in particular those that have been supported through orthogonal experimental tools. We discuss lncRNAs that act in *cis* or in *trans* to promote gene transcription, and lncRNA loci that act as *cis*-regulatory elements to stabilize the transcription of genes encoding transcription factors (TFs), or as allele-specific repressors. Finally, we discuss how our understanding of lncRNA function is helping unravel the impact of non-coding genome variation in human diseases, and the potential of lncRNAs as therapeutic targets.

Transcription activation in *cis* by lncRNAs

Genes with complex expression patterns are controlled by enhancers and other DNA elements located in their surrounding genomic regions, which often extend to hundreds of thousands of base pairs. These regulatory domains adopt 3D configurations known as topologically associating domains (TADs), which enable enhancers to gain spatial proximity to the genes they regulate, while restricting their interactions with genes from nearby TADs. This spatial framework enables regulatory elements to act in *cis*, that is, to control genes located on the same DNA molecule. Multiple lncRNAs have been shown to function in such *cis*-regulatory domains, sometimes through poorly understood mechanisms.

Many lncRNAs spatially interact with neighbouring mRNA-expressing genes, and the expression of these lncRNA–coding gene pairs correlates across tissues and individuals[9–11]. Genetic experiments have revealed that some of these lncRNAs have a *cis*-activating function[12–20]. Although active enhancers also produce non-coding transcripts called enhancer RNAs (eRNAs)[21,22], *cis*-activating lncRNAs are often distinguished from eRNAs because the promoters of *cis*-activating lncRNAs are flanked by nucleosomes with histone modifications typical of promoters, such as high levels of histone H3 trimethylated at lysine 4 (H3K4me3), instead of conventional enhancer modifications such as H3 methylated at lysine 4 (H3K4me1)[23]. Furthermore, lncRNAs are preferentially spliced and are stable, as opposed to eRNAs, which are generally shorter than 500 nucleotides, unspliced and rapidly degraded[21,22]. Despite these differences, the distinction between putative *cis*-activating lncRNAs and eRNAs can be blurry, as stable multi-exonic transcripts are also formed at regions carrying classic active-enhancer chromatin signatures.

The null hypothesis for candidate *cis*-activating lncRNAs is that they are passively produced from active DNA enhancer regions, and that the lncRNA molecules are thus merely 'transcription noise'. An example is mouse *Lockd*, a spliced lineage-specific lncRNA which, as its name indicates, is expressed 5 kb downstream of *Cdkn1b*, the gene that encodes the cell cycle regulator p27 (ref. 24). Whereas deletion of the entire *Lockd* locus in mouse erythroblasts results in reduced transcription of *Cdkn1b*, premature termination of *Lockd* does not affect *Cdkn1b* levels[25]. This finding was taken to indicate that *Cdkn1b* is positively regulated by a *cis* DNA element, but that the transcription of *Lockd* or the transcribed *Lockd* itself is dispensable for *Cdkn1b* expression[26].

Studies at other *cis*-activating lncRNA loci, however, support the notion that lncRNAs are not simply passive by-products of transcription. In this section, we discuss different mechanisms through which lncRNA transcripts or the process of lncRNA transcription can contribute to *cis*-activating functions[12–19].

lncRNAs that function as scaffolds

Some lncRNAs, exemplified by the lncRNA *HOTTIP*, have been proposed to form a local concentration gradient that provides a scaffold for locus-specific recruitment of regulatory complexes, which in turn regulate the transcription of neighbouring genes[27–29] (Fig. 1a). A related proposed mechanism is the RNA-mediated feedback model[30]. In this model, low levels of nascent RNA initially enhance the formation of transcriptional condensates that promote transcription, whereas transcript elongation elevates local RNA concentration, which dissolves the condensates and reduces transcription (Fig. 1b). The key effector in this process is the charge balance of electrostatic interactions provided by RNA molecules, which is proportional to RNA length and local abundance, whereas the importance of the RNA sequence itself is minor. Notably, the cyclic nature of this process suggests that the RNA concentration is the basis of the burst kinetics of promoters' activity, which cannot be explained by classic promoter–enhancer interaction and function[31].

A broad range of TFs were recently shown to harbour RNA-binding domains[32]. The resulting RNA–TF interactions, which also have limited

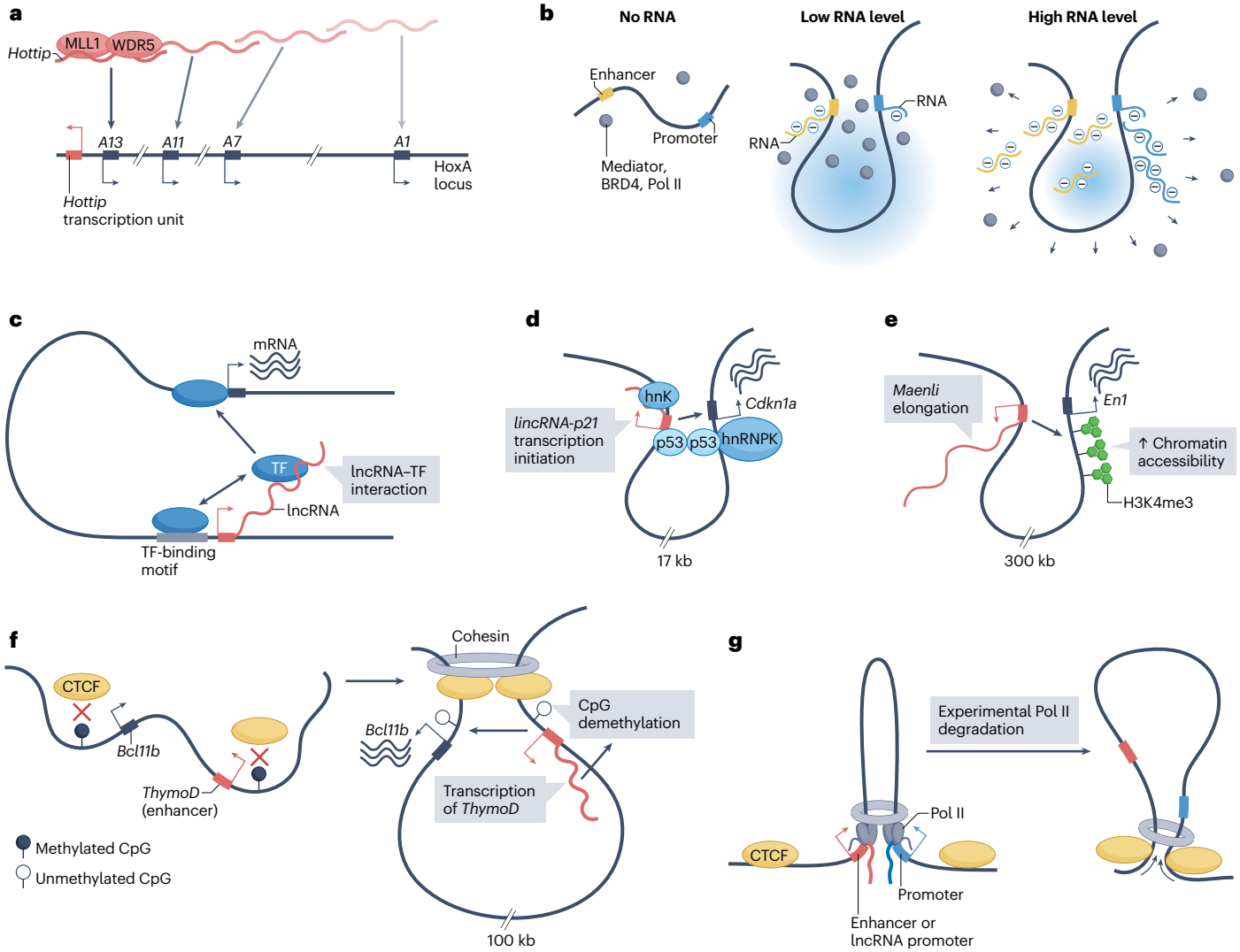


Fig. 1 | Mechanisms of transcription activation in *cis* by long non-coding RNAs. a, The long non-coding RNA (lncRNA) *Hottip* is expressed from the HoxA locus and serves as a scaffold for the local recruitment of the histone methyltransferase complex comprising MLL1 (also known as KMT2A) and WDR5 to HoxA gene (*A1*–*A13*) sites of transcription. Consistent with local activity, RNAi-mediated depletion of *Hottip* preferentially affects *Hottip* proximal, compared with distal, HoxA genes (fading colour gradient). **b**, Local RNA abundance provides feedback on transcription initiation. Left: the Mediator complex, the histone-acetylation reader bromodomain-containing protein 4 (BRD4) and RNA polymerase II (Pol II) are present in low abundance at transcriptionally inactive promoter and enhancer elements. Middle: upon transcription initiation, nascent RNAs produced from promoter and enhancer regions nucleate the formation of a condensate, which increases the local concentration of transcription regulators, thereby causing a burst in transcription. Right: as transcription proceeds, the increase in local RNA abundance beyond a certain threshold generates electrostatic repulsive forces that disperse the transcriptional condensates, thereby ending the transcription burst. **c**, Many transcription factors (TF) have RNA-binding domains, which potentially interact with nascent transcripts, including of lncRNAs. These lncRNA–TF interactions could contribute to the targeting or the strength of association of the TFs to their genomic target sites by taking advantage of their pre-existing 3D proximity. **d**, *lincRNA-p21* and its *cis*-activated target, the neighbouring gene *Cdkn1a* (encoding p21), are in 3D proximity, and are co-regulated by the TF p53. Transcription initiation of *lincRNA-p21* is sufficient to enhance the expression of *Cdkn1a* by creating a scaffold for the recruitment of the *Cdkn1a* transcriptional co-activator heterogeneous nuclear ribonucleoprotein K (hnRNPK). **e**, Transcription elongation of the lncRNA *Maenli* increases local histone H3 trimethylated at lysine 4 (H3K4me3), which is a mark of transcriptionally active chromatin, and promotes the expression of the neighbouring gene *En1*. **f**, Transcription of a lncRNA in the *Bcl11b* locus named thymocyte differentiation factor (*ThymoD*) promotes the demethylation of CTCF-binding sites and, therefore, CTCF recruitment and chromatin reorganization. This process brings the *ThymoD*-associated enhancer region in proximity with their target, the promoter *of Bcl11b*, resulting in transcription activation. **g**, Pol II recruitment to enhancers and promoters blocks chromatin-loop extrusion and stabilizes the loops at a configuration that brings active enhancers in proximity of active promoters. Experimental degradation of Pol II leads to the formation of larger loops, extrusion of which is now limited by CTCF.

RNA sequence specificity, were shown to enhance the binding of TFs to chromatin and to promote their ability to activate transcription of episomal reporters[32]. Although the effects and biological contexts in vivo of these types of RNA–TF interactions remain to be elucidated, this model suggests a plausible function for some *cis*-activating lncRNAs (Fig. 1c). The notion that lncRNAs modulate TFs is supported by earlier findings on individual lncRNAs, such as a report that the human lncRNA *A-ROD* recruits TFs upon its release from chromatin to promote the expression of its neighbouring protein-coding gene *DKK1* (ref. 33).

Cis-activation by transcription initiation of lncRNAs

Mouse *lincRNA-p21* (also known as *Trp53cor1*) is a spliced and polyadenylated lncRNA transcribed 17 kb upstream of *Cdkn1a* (encoding the cell cycle inhibitor p21). *lincRNA-p21* illustrates the functional importance of transcription initiation[34]. In this locus, the promoters of *lincRNA-p21* and *Cdkn1a* engage in constitutive 3D chromatin interactions, and both harbour p53 response elements that confer responsiveness to p53 during cellular stress[15] (Fig. 1d). Deletion of the *lincRNA-p21* p53 response element, the *lincRNA-p21* promoter or the entire locus in mice led to reduced p21 levels and function, indicating the presence of a *Cdkn1a cis*-activating element in *lincRNA-p21* (refs. 13,15,35). Indeed, interference with *lincRNA-p21* transcription initiation significantly reduced *Cdkn1a* expression, indicating that the earliest steps of production of nascent *lincRNA-p21* are required for *Cdkn1a cis*-activation[15,36] (Fig. 1d). By contrast, premature transcription termination, abrogation of splicing or ribozyme-mediated degradation of *lincRNA-p21* had no effect on *Cdkn1a*, indicating that the mature transcript is dispensable for p21 upregulation[15]. A similar mechanism was observed for the activation of the developmental gene Eomesodermin (*Eomes*) (a TF) by the lncRNA *Meteor* (mesendoderm transcriptional enhancer organizing region)[16] and for the stimulation of the inflammation regulator *Ptgs2* by *lincRNA-Cox2* (ref. 18). Collectively, these examples indicate that lncRNA transcription initiation and minimal elongation are *cis*-activating processes of a class of lncRNAs.

It remains to be determined how transcription activation of some lncRNA promoters evokes them to function as enhancers, although several mechanisms have been proposed, including the possibility that nascent lncRNA transcripts tethered to the locus by RNA polymerase II (Pol II) may enhance the recruitment of additional factors. The recruitment includes factors that can contribute to the formation of transcription condensates (Fig. 1b), TFs (Fig. 1c) or transcriptional cofactors, as exemplified by the recruitment of the *Cdkn1a* activator heterogeneous nuclear ribonucleoprotein K (hnRNPK) by *lincRNA-p21* (refs. 13,34) (Fig. 1d).

Cis-activation by lncRNA-transcript elongation and processing

For some lncRNAs, transcript elongation or processing, rather than initiation, is functionally important. *Maenli* (also known as *Gm29348*), a multi-exonic mouse lncRNA transcribed 300 kb upstream of the homeobox TF *En1*, illustrates a transcription-based activation mechanism with important implications for limb development[17]. A series of genetically engineered mouse strains revealed that *Maenli* elongation is required to establish a permissive chromatin environment for *En1* expression in *cis* in limbs[17] (Fig. 1e). Elongation and splicing of the lncRNA *Blustr* were also proposed to regulate the neighbouring Polycomb-group gene *Sfmbt2* (ref. 37). *Blustr* length, splicing and rate of transcription — but not specific sequence elements in the mature *Blustr* transcript — contribute to its *cis*-activating function[37].

Evidence for the role of splicing in promoting *cis*-activation of transcription comes also from genome-scale observations that the production of multi-exonic non-coding transcripts from strong enhancers is evolutionarily constrained[38]. Accordingly, splicing motifs in lncRNAs, rather than exonic sequences or expression levels of lncRNAs, are under strong purifying selection[39,40]. Importantly, individuals who carry nucleotide variants at splice sites of enhancer lncRNAs exhibit changes in local chromatin epigenetic signatures and altered expression of target mRNAs[41].

Cis-activation through 3D genome reconfiguration

lncRNA transcription has been shown to influence 3D chromatin architecture and enhancer–promoter interactions at several loci, including the lncRNAs *PLUT* in human or thymocyte differentiation factor (*ThymoD*) in mouse, as well as the protocadherin-α gene cluster lncRNAs[12,19,42]. The transcription of *ThymoD* lncRNA from an enhancer of the *Bcl11b* gene was linked to the demethylation of CpG dinucleotides at CTCF-binding sequences in the locus, which increases CTCF occupancy[19] (Fig. 1f). The resulting changes in chromatin topology reposition the enhancer region in greater proximity to its target, the *Bcl11b* gene promoter[19] (Fig. 1f). A related mechanism was described for the stochastic selection of alternative protocadherin-α promoters[42]. Each protocadherin-α promoter overlaps with an antisense lncRNA, and transcription activation of individual antisense lncRNAs causes demethylation of CTCF-binding sites and increased CTCF binding, which enables the formation of long-range chromatin loops with a distal cluster of enhancers[42].

In addition to these studies of individual loci, recent work suggests that Pol II recruitment to lncRNAs could have a widespread influence on 3D genome organization. Transcript elongation of lncRNAs was shown to dissociate long-range chromatin contacts (loops) by displacing CTCF and cohesin, without necessarily modifying DNA methylation at the CTCF-binding sites[43]. By contrast, recent genome-scale studies have illustrated how Pol II recruitment and productive elongation from enhancers and promoters can block the extrusion of chromatin loops, thereby increasing contacts between enhancers and promoters[44,45] (Fig. 1g). Although the impact of transcription on 3D genome conformation remains unsettled, these findings raise the possibility that the *cis*-activation function of some transcribed lncRNAs is mediated by the effects of transcription initiation or processing on spatial genome organization.

eRNA contribution to enhancer function

Early studies proposed that eRNAs mediate essential enhancer functions, such as the recruitment of Pol II and loading of the Mediator complex to protein-coding target genes, or the regulation of CTCF-mediated chromatin remodelling to enable enhancer–promoter interactions[46–49]. However, these functions were primarily determined using RNAi and antisense oligonucleotides to deplete eRNAs, which in addition to RNA inhibition have been shown to induce epigenetic changes in the chromatin of target loci[50–52].

More recent work, based on varied perturbation tools, has supported the notion that eRNAs could elicit similar functions to those described for *cis*-activating lncRNAs. For example, eRNAs, similar to other lncRNAs, have been reported to contribute to the formation of local condensates, which lead to enhanced gene expression through increased formation and stabilization of transcription co-activator complexes[53,54]. Furthermore, the suggested regulatory impact of RNA–TF interactions theoretically applies to both eRNAs and lncRNAs[32]. One study has specifically focused on eRNAs containing Alu repetitive

sequences and found that they can promote enhancer–promoter pairing by forming RNA duplexes with promoter-associated RNAs[55].

The spatial effects of Pol II recruitment could be relevant to the *cis*-activating function of enhancers and *cis*-activating lncRNA loci, by blocking loop extrusion and promoting proximity to target genes[44,45] (Fig. 1g). Pol II recruitment to enhancers has also been proposed to increase enhancer–promoter contacts by tethering both elements at transcriptionally active hubs or condensates[56]. These findings suggest that enhancer transcription is important for normal enhancer–promoter interactions.

Although the analysis of *cis*-activating transcription units has largely focused on lncRNA promoters and eRNAs, active promoters of protein-coding genes can also have long-range enhancer activity[57]. The conditions in which transcription regulation functions are enacted by eRNA, lncRNA or mRNA promoters are poorly understood.

lncRNAs as local rheostats of transcription-factor genes

In contrast to lncRNAs that contribute to the activation of transcription by enhancers, other lncRNAs carry out specialized rheostat-like functions in *cis*.

Many lncRNA genes are located in the vicinity of TF genes[9,58,59], and in numerous studies the knockout of a lncRNA has led to moderately increased expression of a TF gene in the same TAD (Table 1). Further analyses of some of these individual lncRNAs indicate that they are fundamentally different from transcription silencers, which repress their target genes[60,61], or from enhancers that confer cell type-specific gene activation. Instead, these lncRNAs regulate already-active genes and tune their transcription to ensure appropriate gene product concentrations (Fig. 2a). We refer to such lncRNAs as gene expression stabilizers, in analogy to voltage stabilizers, which are devices used to protect electronic equipment from excessively high voltage levels.

One of the earliest examples of such a stabilizer is the mouse lncRNA *Halr1* (also known as *Haunt* or *linc-HOXA1*), which is located ~40 kb from the homeobox TF gene *Hoxa1*. Three studies showed that *Halr1* depletion or mutations in its promoter led to increased expression *of Hoxa1* and other genes of the Hoxa cluster in pluripotent stem cells exposed to retinoic acid[62–64] (Fig. 2b), indicating that *Halr1* guards against inappropriately elevated *Hoxa* gene expression.

Similarly, two mouse lncRNA transcription units curtail the expression of *Hand2*, which encodes a cardiac basic helix–loop–helix TF (Fig. 2c). One is the lncRNA *Hand2os1* (also known as *Upperhand*, *lncHand2* or *HAND2-AS1*), which is transcribed divergently of *Hand2* (refs. 65,66). Deletion of two *Hand2os1* distal exons or of a much broader *Hand2os1* sequence in mice causes increased cardiac expression of *Hand2* mRNA, which leads to augmented expression of HAND2-dependent genes in specific cardiac cell subpopulations[66]. The other lncRNA, *Hdnr* (also known as *Handsdown*), is located downstream of *Hand2*. Insertion of a polyadenylation signal to prematurely

Table 1 | Long non-coding RNAs whose loss results in increased expression of a neighbouring transcription-factor gene

lncRNA	TF gene	Effect in *cis* or in *trans*	lncRNA transcription required	Modulation of enhancer–promoter contacts	TF–lncRNA feedback	Response signalling	Mouse phenotype of lncRNA loss	Dosage-sensitive phenotype of the human TF gene	Refs.[a]
Halr1	*Hoxa1* *Hoxa2*	*Cis*	Yes	Yes	Negative feedback	Retinoic acid	-	Haploinsufficiency: ear defects	62–64
HASTER	*HNF1A*	*Cis*	No	Yes	Negative feedback	Unknown	Diabetes	Haploinsufficiency: diabetes	70
CHASERR	*CHD2*	*Cis*	Yes	Yes	Negative feedback	Unknown	Lethal	Haploinsufficiency: neurodevelopmental	71
Hand2os1	*Hand2*	Unknown	Unknown	Unknown	Unknown	Unknown	Cardiac abnormalities	Haploinsufficiency: cardiac	65,66
Hdnr	*Hand2*	Unknown	Yes	Unknown	Possible[b]	Unknown	Cardiac abnormalities	Haploinsufficiency: cardiac	67
Flicr	*Foxp3*	*Cis*	Yes	Unknown	Unknown	IL-2	Reduced type 1 diabetes	No (X-linked autoimmune disease)	69
PVT1	*MYC*	*Cis*	Yes	Yes	Possible[b]	Stress, p53	Loss of function: tumour suppression	Gain of function: oncogenic	14,74
METEOR	*EOMES*	Unknown	No	Yes	Unknown	Unknown	-	-	16
Evf2	*Dlx6*	*Cis*	No	Unknown	Unknown	Unknown	Neurodevelopmental	Haploinsufficiency: split-hand and foot malformation	72
Playrr	*Pitx2*	Unknown	Yes	Unknown	Negative cross-regulation	Unknown	Cardiac arrhythmias	Haploinsufficiency: Rieger syndrome	75
NXTAR	*AR*	Unknown	Yes	Unknown	Negative cross-regulation	Androgen	-	Gain of function : oncogenic	78
FENDRR	*FOXF1*	Possibly *trans*	Yes	Unknown	Negative feedback	Unknown	Vascular malformations (AVCD-MPV)	Haploinsufficiency: vascular malformations (AVCD-MPV)	68

AVCD-MPV, alveolar capillary dysplasia and misalignment of pulmonary veins.; [a]References of long non-coding RNA (lncRNA) mutation studies showing increased expression of a proximal transcription regulator gene. [b]Evidence for regulation of a lncRNA gene by the transcription factor (TF), but no direct evidence for a feedback mechanism.

Glossary

CpG islands
Genomic regions of 500 nucleotides or longer with >50% CpG dinucleotide repeat content. CpG islands are associated with the transcription start sites of most housekeeping genes and as many as 40% of tissue-specific genes; they are bound by regulatory proteins.

CTCF
A zinc-finger transcription factor (TF), also known as CCCTC-binding factor, that binds specific DNA sequences and participates in the formation of chromatin loops that influence gene transcription by defining the boundaries of topologically associated domains (TADs) and bringing enhancers into proximity with promoters.

DNA–DNA–RNA triplex
A structure in which single-stranded RNA invades the major groove of double-stranded DNA and binds by forming Hoogsteen hydrogen bonds. DNA–DNA–lncRNA triplexes can be identified by pull-downs with a triplex-specific antibody.

Enhancer RNAs
(eRNAs). Non-coding RNAs that are bidirectionally transcribed from enhancer regions, and are typically ≤500 nucleotides and unstable (half-life ≤ 2 min).

Enhancers
Genomic regions that are recognized by transcription factors (TFs) and activate and increase the transcription of genes in *cis*, sometimes from considerable distances. Active enhancers are flanked by nucleosomes that carry post-translational histone modifications such as histone H3 acetylated at lysine 27 (H3K27ac) and H3 methylated at lysine 4 (H3K4me1).

Expression quantitative trait loci
(eQTL). Genetic loci in which different alleles of a DNA variant influence expression levels of coding or non-coding transcripts.

Focal deletions
Cancer-associated genomic deletions smaller than 5 Mb that affect both alleles.

Genome-wide association studies
(GWAS). Studies that survey DNA variants genome-wide to identify those showing association with a disease or trait. GWAS have been used to discover susceptibility variants for prevalent polygenic diseases. A large fraction of significant associations are found in non-coding genomic regions, indicating that they are mediated by genetic variants that influence regulatory functions.

lncRNAs that act in *cis*
Long non-coding RNAs (lncRNAs) that act on the same chromosome from which they are transcribed, including the regulation of a neighbouring gene, of multiple genes or of the entire chromosome.

Silencers
Genomic regions that are bound by repressive transcription factors (TFs) and decrease the transcription of genes in *cis*.

Splicing quantitative trait loci
Genetic loci in which different alleles influence RNA splicing patterns.

Topologically associated domains
(TADs). Genomic regions defined by having a higher frequency of long-range chromatin contacts, such as between genes and their regulatory elements, than the frequency of contacts with elements outside the region.

Transcriptional condensates
Chromatin-associated, dynamic nuclear assemblies comprising a heterogeneous mix of RNAs, transcription factors (TFs) and co-regulators that modulate transcriptional output.

terminate *Hdnr* transcription also increases *Hand2* levels, whereas its CRISPR-based induction reduces *Hand2* expression[67].

The lncRNA *FENDRR* and the TF gene *FOXF1* share an ~1.4 kb promoter region in the mouse and human genomes (Fig. 2d). The replacement of mouse *Fendrr* exon 1 with a polyadenylation cassette that interrupts *Fendrr* transcription resulted in increased *Foxf1* mRNA levels in the caudal end of embryonic day 8.5 and 9.5 mouse embryos, and ectopic expression of *Foxf1* mRNA in the heart of embryonic day 12.5 embryos[68]. Mutations in several other lncRNA genes similarly elicit moderately increased expression of adjacent transcription regulator genes, including *Flicr*[69], *HASTER* (also known as *HNF1A-AS1*)[70], *CHASERR*[71], *Evf2* (also known as *Dlx6os1*)[72], *Hotair*[73], *Pvt1*[14,74] and *Playrr*[75] (Table 1 and Fig. 2e–g).

Fine-tuning the expression of transcription-factor genes through feedback

The dampening effect of lncRNAs on adjacent TF genes could theoretically represent a steady-state tonic inhibition, but some lncRNAs have been shown to act as robust feedback circuits that maintain stable concentrations of regulatory proteins within a narrow concentration range. An example of such a rheostat-like mechanism is the human lncRNA named *HASTER*. *HASTER* transcription starts in the first intron of *HNF1A*, which encodes a homeodomain TF, and proceeds in antisense orientation to *HNF1A* (ref. 70) (Fig. 2f). High HNF1A protein concentrations result in the direct activation of the *HASTER* promoter, and this leads to inhibition of *HNF1A* transcription[70]. Consequently, deletion of the *HASTER* promoter in mice or human stem cell-derived hepatocytes boosted *HNF1A* expression[70]. This indicates that *HNF1A* and *HASTER* form a classic negative feedback loop that prevents HNF1A overexpression (Fig. 2f).

This type of feedback likely exists at other loci. The lncRNA *Halr1*, which limits retinoic acid-induced transcription of *Hoxa1*, is positively regulated by HOXA1 (ref. 63) (Fig. 2b). Depletion studies have shown that FOXF1 is a positive regulator of the lncRNA *FENDRR*[76], which in turn inhibits *FOXF1* (ref. 68) (Fig. 2d), and that HAND2 binds at two sites in the *Hdnr* locus, which could conceivably modulate the inhibitory function of *Hdnr*[77].

Other types of feedback appear to operate in some lncRNA–TF gene pairs. A recent study reported in a preprint points to a negative cross-regulatory feedback loop formed by the mouse lncRNA *Playrr* and its adjacent homeodomain TF gene *Pitx2*[75] (Table 1). These two genes are expressed in mutually exclusive heart domains, and loss of a *Playrr* splice site leads to increased and inappropriate *Pitx2* expression patterns, causing cardiac arrythmia[75]. The human lncRNA *NXTAR* also forms negative cross-regulatory feedback with the *AR* gene, which in turn inhibits *NXTAR*[78].

These examples raise the question of why TF genes require feedback. Negative feedback systems dampen fluctuations and can provide a cell type-specific range for TF concentrations. This is important because TF concentrations are crucial determinants of genomic-site binding choices[79]. Lineage-selective TFs are often expressed at different levels across cell types, and typically bind to different cell type-specific genomic sites[80,81]. Tight regulation of TF concentrations may be particularly relevant for pioneer TFs, which bind DNA sequences in inaccessible chromatin. HNF1A, in particular, has typical pioneer factor

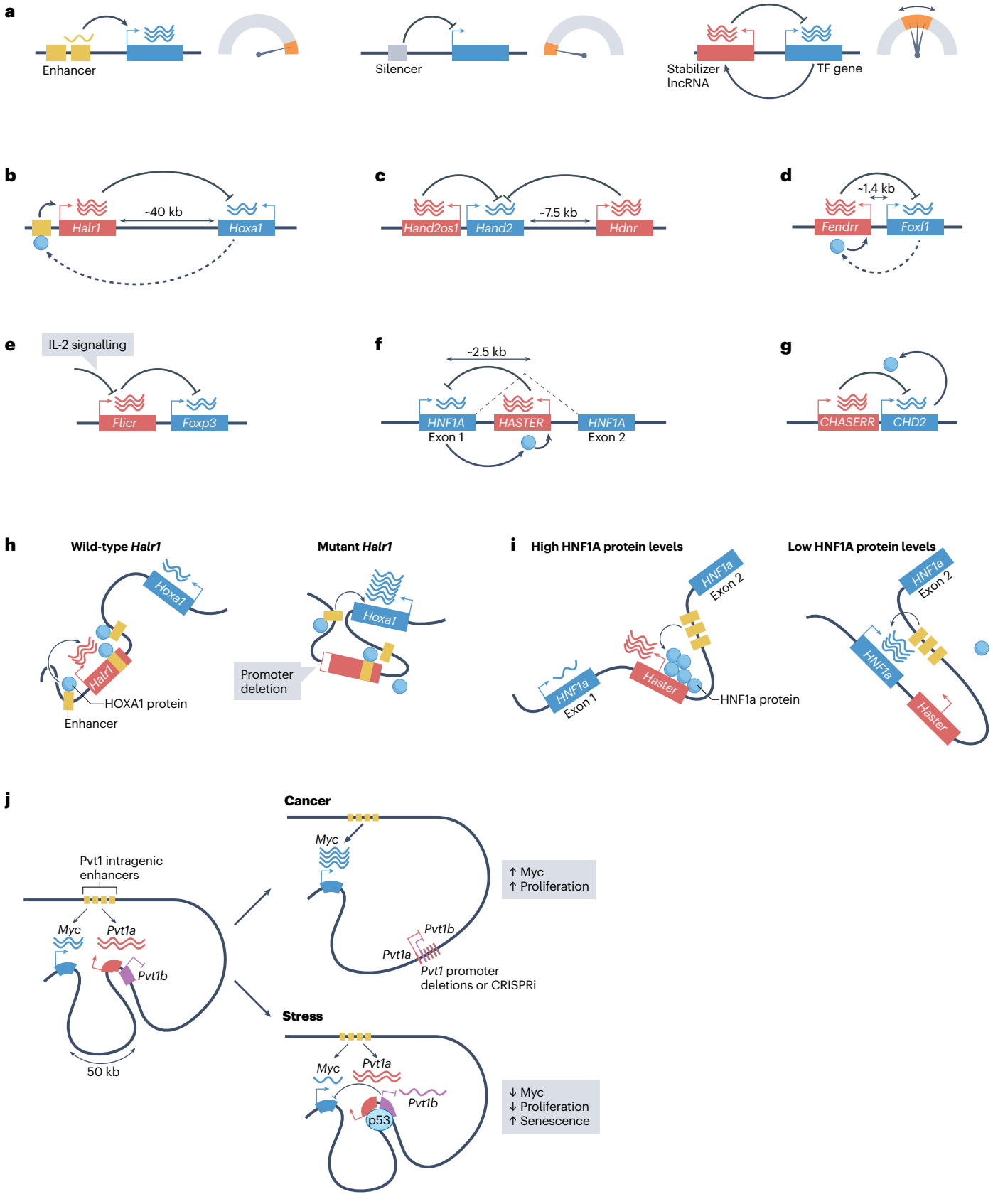

a
Enhancer
Silencer
Stabilizer lncRNA
TF gene
b
Halr1
~40 kb
Hoxa1
c
Hand2os1
Hand2
~7.5 kb
Hdnr
d
~1.4 kb
Fendrr
Foxf1
e
IL-2 signalling
Flicr
Foxp3
f
~2.5 kb
HNF1A
Exon 1
HASTER
HNF1A
Exon 2
g
CHASERR
CHD2
h
Wild-type Halr1
Hoxa1
Halr1
HOXA1 protein
Enhancer
Mutant Halr1
Hoxa1
Promoter deletion
i
High HNF1A protein levels
HNF1a
Exon 2
HNF1a
Exon 1
Haster
HNF1a protein
Low HNF1A protein levels
HNF1a
Exon 2
HNF1a
Haster
j
Pvt1 intragenic enhancers
Myc
Pvt1a
Pvt1b
50 kb
Cancer
Myc
Pvt1b
Pvt1a
Pvt1 promoter deletions or CRISPRi
↑ Myc
↑ Proliferation
Stress
Myc
Pvt1a
Pvt1b
p53
↓ Myc
↓ Proliferation
↑ Senescence

Fig. 2 | Long non-coding RNAs as *cis*-acting transcription stabilizers. **a**, Whereas enhancers promote cell type-specific gene activation and silencers prevent the expression of their target genes, transcription-stabilizing long non-coding RNAs (lncRNAs) act in *cis* to tune the transcription level of dosage-sensitive transcription-factor (TF) genes. **b**–**g**, Examples of lncRNAs (in red) that act as transcription stabilizers of adjacent genes, all of which encode transcription regulators (in blue). The lncRNAs are *Halr1* (**b**), *Hand2os1* (**c**), *FENDRR* (**d**), *Flicr* (**e**), *HASTER* (**f**) and *CHASERR* (**g**) – loss of function of all of these lncRNAs caused increased expression of the adjacent gene. HNF1 homeobox A (HNF1A) and chromodomain helicase DNA-binding protein 2 (CHD2) (blue circles) enhance the inhibitory effects of the adjacent lncRNAs, and therefore provide negative feedback. Homeobox A (HOXA1) and forkhead box F1 (FOXF1) proteins are positive regulators of the lncRNAs *Halr1* and *FENDRR*, respectively, suggesting they could also form a negative feedback loop. Some transcription-stabilizing lncRNAs modulate their target genes in a signal-responsive manner; for example, interleukin-2 (IL-2) acts on *Flicr* (**e**) to reduce high *Foxp3* expression levels in regulatory T cells. Two lncRNAs, *Hand2os1* and *Hdnr* (**c**), restrict *Hand2* expression. **h**,**i**, The promoters of transcription-stabilizing lncRNAs modulate interactions between their target TFs genes and local enhancers. **h**, Left: in pluripotent cells, *Halr1* binds and sequesters proximal enhancers of *Hoxa1*, which dampens retinoic acid-induced expression of *Hoxa1*. Right: deletion of the *Halr1* promoter increases enhancer–*Hoxa1* interactions. HOXA1 (blue circles) binds to local enhancers and activates *Halr1*, which restrains *Hoxa1* expression. Left and right in **h** depict retinoic acid-induced cells. **i**, The *Haster* promoter limits interactions between the *Hnf1a* promoter and intragenic enhancers. This effect is accentuated at high concentrations of HNF1A protein, thereby providing negative feedback on *Hnf1a* transcription. **j**, The active *Pvt1* lncRNA promoter acts as a boundary element that associates with enhancers located within the *Pvt1* gene body and limits access of the *Myc* promoter to these enhancers. Experimental inhibition of the transcription activity of the *Pvt1* promoter through targeted promoter deletions or CRISPR inactivation (CRISPRi) leads to increased *Myc* promoter–enhancer engagement, high *Myc* transcription and increased cellular proliferation. The *Pvt1* locus also harbours a p53-dependent isoform, *Pvt1b*, which downregulates *Myc* transcription during stress, decreases cell proliferation and increases cell senescence without apparent changes in *Myc*–enhancer contacts.

properties, such as the ability to bind nucleosomal DNA[82], and a capacity to activate silent genes in fibroblasts[83]. Accordingly, livers of *Haster* mutant mice, in which HNF1A concentrations are abnormally high, exhibit widespread genomic binding of HNF1A and chromatin opening at HNF1A recognition sequences that are normally inaccessible in hepatocytes, leading to aberrant, ectopic gene transcription[70]. Feedback from lncRNAs could therefore tune TF concentrations to ensure the specificity of cell type-specific gene programmes.

Studies of mouse and human *CHASERR* lncRNA have exemplified another remarkable feedback system that controls a chromatin remodeller instead of a DNA-binding TF[71]. The lncRNA *CHASERR* is transcribed upstream on the same strand of the *CHD2* gene (Fig. 2b). Heterozygous deletions of mouse *Chaserr* promoter or gene body, or lncRNA depletion, increased *Chd2* expression, indicating that *Chaserr* inhibits *Chd2* (ref. 71). The authors postulated that *Chaserr* transcription interferes with transcription of the downstream *Chd2* gene. They also found that CHD2 forms an autoregulatory feedback loop by binding to *Chaserr* transcripts – and to the *Chaserr* gene – which promotes *Chaserr* interference of *Chd2* (ref. 71) (Fig. 2g). Failure of this feedback in *Chaserr* mutants causes increased CHD2 expression, which in turn decreases the expression of many other genes that are located downstream of CHD2-bound transcription units.

Mechanisms of modulation of transcription-factor genes

Genetic experiments have begun to shed light on how stabilizer lncRNAs tune the expression of TF genes. Although lncRNA–TF gene pairs vary greatly in their relative orientations or genomic distance (Fig. 2b–g), perturbation studies have revealed commonalities in their mode of action.

Several studies have demonstrated that the mechanism by which transcription stabilizer lncRNAs modulate the expression of adjacent TF genes occurs in *cis*. The demonstration that a lncRNA exerts its effects in *cis* rules out the possibility that the local regulatory activity is carried out by RNA-encoded polypeptides, as well as other indirect mechanisms. *cis* effects have been demonstrated using heterozygous mutant models that can unequivocally ascertain whether only the chromosome that carries the mutant allele exhibits altered expression of the TF gene. In practice, this analysis can be carried out using either heterozygous lncRNA mutations bred on hybrid mouse strain backgrounds, thereby allowing to distinguish between the two chromosomes, or with compound heterozygotes in which the lncRNA and TF mutations are on separate chromosomes[70,71]. Other studies have pointed to a *cis* effect by showing that ectopic lncRNA expression does not rescue the lncRNA-null mice[69]. The case of *Fendrr* differs from that of other stabilizer lncRNAs in that ectopic expression was shown to partially rescue the mouse *Fendrr* mutant phenotype[68], which suggests that at least some effects of *Fendrr* occur in *trans*. This possibility was supported by another study, which deleted a sequence in *Fendrr* forming a putative DNA–DNA–RNA triplex with various potential target sequences, and found that it partially phenocopies other *Fendrr* mutants[84]. However, these experiments have not fully addressed whether the rheostat-like function of *Fendrr* on the *Foxf1* gene, with which it shares a promoter region and a closely related developmental phenotype, also occurs through this type of *trans* mechanism[68].

Cis-regulatory effects can be mediated through RNA-dependent or DNA-dependent mechanisms. A functional requirement of lncRNA transcription for expression stabilization has been demonstrated using transcription termination alleles or CRISPR-based transactivation for some, but not all, stabilizer lncRNAs[66,67,69,71]. Likewise, genetic perturbations or RNA degradation experiments have shown that the RNA itself is functionally important for the inhibitory effects of *Chaserr*[71], *Flicr*[69] and *Hdnr*[67] on adjacent genes. By contrast, blocking transcription of *HASTER* using nuclease-deficient Cas9 as a roadblock or by inserting a polyadenylation site, as well as activation of *HASTER* transcription through a modified CRISPR–Cas9 system did not have an effect on *HNF1A* mRNA expression[70]. Moreover, overexpression of HNF1A separation-of-function mutants that lacked the ability to transactivate genes, and therefore did not activate *HASTER* transcription, still resulted in feedback inhibition of the endogenous *HNF1A* gene, an effect that required an intact *HASTER* promoter[70]. Thus, HNF1A interactions with the *HASTER* promoter, but not *HASTER* transcript elongation or transcripts, appear to be important for the *HASTER*–*HNF1A* transcription feedback.

Chromatin conformation capture studies have shown that the promoters of several transcription stabilizer lncRNAs, including *HASTER*[70], *PVT1*[74], *Meteor*[16], *Halr1*[63,64] and *Chaserr*[71], limit interactions between enhancers and their target genes, which consequently dampens gene transcription (Fig. 2h–j). In the case of *HASTER*, increased HNF1A concentrations led to enhanced binding to the *HASTER* promoter, which

further limited interactions between *HNF1A* intronic enhancers and the *HNF1A* promoter (Fig. 2i). These independent studies suggest that enhancer competition may be a prevalent mode through which lncRNAs control the expression of TF genes (Fig. 2h–j). In summary, to maintain homeostatic expression levels of TFs, *cis*-acting lncRNAs deploy transcription-dependent and RNA-dependent mechanisms, but also compete with enhancers of their target genes, and several lncRNAs appear to simultaneously use more than one of these molecular mechanisms[14,71,74].

Dual positive and negative regulatory functions

Numerous lncRNA loci intertwine positive and negative *cis*-regulatory functions[14,16,64,70,74]. For example, a comprehensive genetic dissection of *Meteor* showed that its transcription is required to activate *Eomes* in pluripotent cells, whereas the *Meteor* promoter limits *Eomes* expression during neuronal differentiation[16]. Likewise, a deletion of the *Haster* promoter in mice led to increased HNF1A expression in all hepatocytes, whereas this same deletion caused variegated HNF1A expression in pancreatic β-cells, with some cells showing marked HNF1A overexpression and others complete *HNF1A* silencing[70]. The *Haster* promoter, therefore, has a cell type-specific *cis*-activating function in addition to its negative feedback.

Some of the observed dual phenotypes might occur because stabilizer lncRNAs are often embedded in enhancer clusters. Different genetic alterations in a locus with a complex interspersion of positive and negative regulatory elements can easily lead to opposite phenotypes. For example, an allele that produces a premature termination of *Hand2os1* transcription causes a severe loss of cardiac *Hand2* expression[65], whereas *Hand2os1* deletions cause *Hand2* upregulation[65,66]. Likewise, small deletions in *Halr1* and RNA perturbations have resulted in increased expression of HoxA genes, whereas other deletions in the *Halr1* locus have uncovered HoxA-activating sequences[62–64].

Signal-induced modulation of transcription-factor genes by lncRNAs

For some *cis*-regulatory lncRNAs, modulation of the neighbouring gene can be triggered by cellular and environmental perturbations. In this manner, lncRNAs can endow signal responsiveness to a single gene, rather than act on a wide gene expression programme. For example, in humans and mice, *PVT1* is a collection of alternatively spliced lncRNAs initiated ~50 kb downstream of the TF oncogene *MYC*, which accumulate locally and downregulate *MYC* transcription[14,74,85]. A study in human breast cancer cell lines showed that the *PVT1* promoter limits long-range interactions between the *MYC* promoter and *PVT1* intragenic enhancers, and therefore reduces *MYC* expression[74] (Fig. 2j). Human and mouse *PVT1* are part of an additional inhibitory mechanism that involves the expression of a stress-dependent, p53-inducible transcript isoform termed *Pvt1b*, which is initiated at a downstream transcription start site and whose production inhibits transcription of *Myc* without insulating *Myc* from its enhancers[14] (Fig. 2j). The stress-induced *Pvt1b* isoform reverses transcription activation by p53 into a local inhibitory signal, thereby limiting Myc levels and reducing cell proliferation. Interestingly, *Pvt1b* production leads to both proliferation arrest by *Myc* downregulation within hours of stress, and to long-term *Myc* repression, which is associated with activation of cell senescence, suggesting it has a role in epigenetic reprogramming of the *Myc* locus[14,86].

Flicr is another signal-responsive transcription stabilizer that dampens the expression of *Foxp3*, which encodes a forkhead TF that controls regulatory T cells[69] (Fig. 2e). The disruption of mouse *Flicr* promoters, a mutation of a *Flicr* splice donor site or targeted degradation of *Flicr* all led to increased expression of *Foxp3* and its target genes, and to a relative depletion of regulatory T cell subpopulations that express low *Foxp3* levels[69]. This process can be modulated by interleukin-2 (IL-2), which inhibits *Flicr* expression, thus promoting *Foxp3* expression and regulatory T cell expansion[69].

Whereas *Pvt1* and *Flicr* tune the transcription level of active genes, other signal-responsive lncRNAs elicit transcription switches. The mouse and human lncRNA *MORRBID*, for example, rapidly downregulates the neighbouring pro-apoptotic gene *BCL2L11* (also known as *BIM*), in response to cytokines and viral infections, thus promoting the survival of myeloid and $CD8^+$ T cells[87,88]. *Bcl2l11* downregulation is accompanied by deposition of the gene-repressive H3K27me3-modified chromatin[88]. A plant lncRNA named *COOLAIR*, which inactivates the expression of the vernalization TF FLOWERING LOCUS C (FLC) during the autumn to winter transition[89], provides fascinating insights into how a lncRNA carries out a signal-responsive binary switch (Box 1).

Cis-regulatory lncRNAs as allele-specific repressors

Another notable gene regulatory activity that cannot be explained solely through the general framework of TFs interacting with specific DNA sequences is the selective silencing of one of two homologous chromosomal loci. Some of the best characterized lncRNAs accomplish this type of function, including the lncRNA *XIST*, which regulates X-chromosome inactivation (XCI), and imprinted lncRNAs that control parent-of-origin allele-specific repression.

Box 1

A plant long non-coding RNA as a paradigm of environmental switch

The *Arabidopsis thaliana* long non-coding RNA (lncRNA) *COOLAIR* exemplifies how a lncRNA can function as a binary switch. *COOLAIR* is a gene comprising a group of antisense, alternatively spliced lncRNA isoforms that overlap the gene body and promoter of the vernalization transcription factor (TF) *FLC* gene[255–257]. *COOLAIR* responds to environmental cues such as the first seasonal frost, and switches off *FLC* expression during the autumn to winter transition[89]. In vivo analysis of structural conformations of individual *COOLAIR* RNA molecules revealed striking cold-dependent enrichment of specific structural isoforms[258]. Interestingly, these transcript structural variants preferentially occur in a key region of complementarity between *COOLAIR* and the transcription start site of *FLC*. The structural variability might influence the ability of the lncRNA to form an R-loop at the 5′ end of *FLC*, mediate DNA–DNA–RNA triplex formation between *COOLAIR* and the *FLC* transcription start site or promote the recruitment of a protein complex to the *FLC* transcription start site[259]. Although the exact mechanism by which *COOLAIR* suppresses *FLC* is unclear, this finding reveals a new dimension of RNA-based *cis*-regulation, namely the capacity to be dynamically altered by adopting alternative structural conformations in response to environmental cues[258].

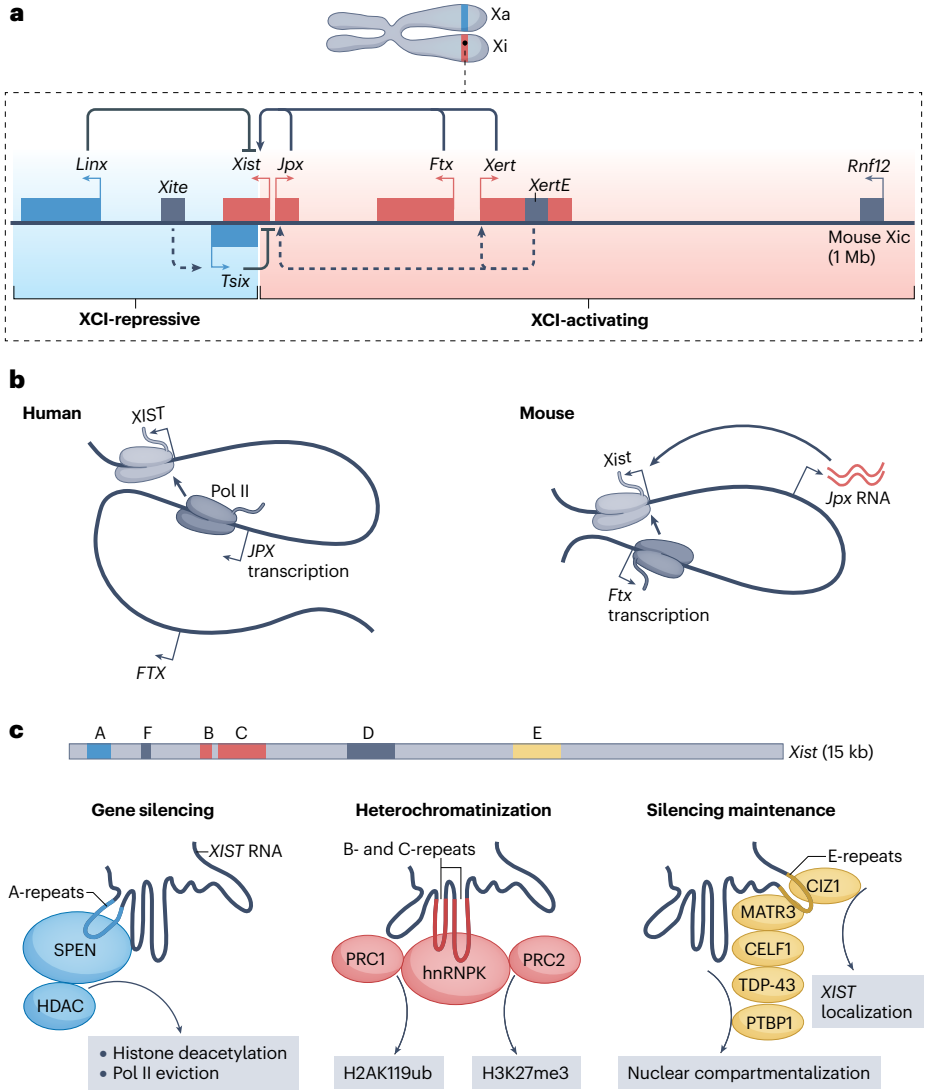


Fig. 3 | Control of X-chromosome inactivation by long non-coding RNAs. **a**, Transcription activation of mouse *Xist*. The X-inactivation centre (Xic) shows two topologically associating domains (TADs) in mouse cells. In one TAD (blue background), the long non-coding RNAs (lncRNAs) *Linx* and *Tsix* – antisense transcript of *Xist* – and the *Tsix* enhancers, termed *Xite*, are located. On the active X chromosome (Xa), *Tsix* transcription suppresses *Xist* transcription, whereas the *Linx* promoter acts across the TAD boundary to limit *Xist* expression in *cis*. In the other TAD (red background), the lncRNAs *Xist*, *Jpx*, *Ftx* and *Xert* are located. *Xert* enhancers, termed *XertE*, promote both *Xert* and *Xist* transcription on the inactive X chromosome (Xi). Following X-chromosome inactivation (XCI), *Jpx* and *Ftx* maintain *Xist* expression and accumulation at Xi. **b**, Similarities and differences between *XIST* regulation by *JPX* and *FTX* in human and mouse. In human, whereas *FTX* is not essential for *XIST* regulation, *JPX* transcription, but not the mature RNA, contributes to polymerase II (Pol II) loading and *XIST* transcription and accumulation. In mouse, *Ftx* transcription promotes *Xist* transcription, whereas mature *Jpx* transcript is responsible for *Xist* transcriptional activation and accumulation. **c**, *XIST* mediates transcriptional gene silencing at the X chromosome. *XIST* RNA highlighting its repeat regions A–F and showing the role of A-repeats in promoting the initial steps of gene silencing through SPEN-mediated and histone deacetylase (HDAC)-mediated histone deacetylation and RNA Pol II eviction; the role of B-repeats and C-repeats in heterochromatinization through recruitment of Polycomb repressive complex 1 (PRC1) and PRC2 downstream of heterogeneous nuclear ribonucleoprotein K (hnRNPK); and the role E-repeats in the CIP1-interacting zinc finger protein 1 (CIZ1)-dependent maintenance of *XIST* localization at Xi and in recruiting RNA-binding proteins to mediate the nuclear compartmentalization of Xi. H2AK119ub, histone H2A ubiquitylated at lysine 119; H3K27me3, histone H3 trimethylated at lysine 27; MATR3, matrin 3; PTBP1, polypyrimidine tract-binding protein 1; TDP-43, TAR DNA-binding protein 43.

X-chromosome inactivation

XCI in mammals ensures X-linked dosage compensation between cells of females and males by inactivating one of the two X chromosomes in female cells[90–92]. This process is controlled by the X-inactivation centre (Xic), a genomic region that integrates X-chromosome counting information with random selection of one of the two female X chromosomes for inactivation[93–96].

The lncRNA gene *XIST*, which has a central role in XCI, is located in the Xic and is selectively transcribed from what will become the silenced X chromosome (Xi)[90,97] (Fig. 3a). *XIST* transcription is regulated by

neighbouring activating and repressive *cis*-regulatory lncRNAs, which are partitioned into two adjacent TADs[98]. In mice, the lncRNAs *Jpx, Ftx* and *Xert* are located in the same TAD as *Xist* and promote its transcription in *cis*[99,100] (Fig. 3a). This regulation has been demonstrated also in human cell models, and by a 453-kb deletion in a human female that overlaps *JPX* and *FTX* and caused markedly skewed *XIST* expression from the intact chromosome[101]. Mechanistically, *XIST* transcription and accumulation depend on *JPX* transcription in human cells, or on the accumulation of mature *Jpx* RNA in mice[100] (Fig. 3b). Furthermore, *Jpx* can also act in *cis* and in *trans* to activate *Xist* by binding and displacing CTCF from the *Xist* promoter[102].

By contrast, the lncRNA *Tsix*, which inhibits *Xist* transcription in *cis*, is located in an adjacent TAD, along with the *Xite* enhancer elements, which promote *Tsix* activation[103,104] (Fig. 3a). The lncRNA *Linx* also maps to this TAD and acts as a distant *cis*-inhibitor of *Xist*[105]. This effect is exerted by the active *Linx* promoter, and is independent of *Linx* transcript elongation or effects on *Tsix*[105]. Heterozygous inactivation of *Tsix, Xite* or *Linx* prior to the onset of XCI shows that they are essential *cis*-regulators of *Xist* expression in mice[104–106].

Elegant studies have shown that once the Xi is selected, Xi-specific expression and accumulation of *Xist* is both necessary and sufficient for chromosome-wide gene repression in *cis*. Early experiments established that an inducible *Xist* transgene can silence autosomes in *cis* in embryonic stem cells[107,108]. Molecular and genetic deletion studies have since revealed that repetitive sequences and structural elements within the *Xist* RNA are central to its ability to recruit regulatory proteins to Xi[109–113] (Fig. 3c). The A-repeats of *Xist* adopt structural features that are recognized by SPEN (also known as MSX2-interacting protein), a transcription co-repressor that mediates the recruitment of chromatin modifying complexes that promote histone deacetylation, evict Pol II and contribute to the early steps of X-linked gene silencing[111,114–117]. B-repeats and C-repeats have been implicated in the scaffolding of hnRNPK, which mediates the recruitment of Polycomb repressive complex 1 (PRC1) and the subsequent activity of PRC2 at Xi[118–122], although a recent study also describes independent binding of PRC2 to A-repeats during initiation of XCI[123]. The E-repeats mediate the assembly of RNA-binding proteins (RBPs) such as CIP1-interacting zinc finger protein (CIZ1), polypyrimidine tract-binding protein 1 (PTBP1), matrin 3 (MATR3), TAR DNA-binding protein 43 (TDP-43) and CELF1, which promotes chromatin compaction and maintenance of late-stage Xi in a phase-separated compartment[124,125].

Ultimately, the accumulation of *Xist*-scaffolded protein complexes at Xi begins a succession of events that initiate, spread and maintain transcriptional gene silencing[97,126] through the formation of a repressive chromatin state[127–129] and reconfiguration of the chromosomal architecture[112,130,131].

Parent-of-origin allelic repression

Imprinted loci provide another mechanism of lncRNA-dependent *cis*-regulation. There are at least 29 imprinted domains in the mouse and human genomes, harbouring more than 150 imprinted genes that are organized in clusters[132]. Imprinting of such loci is critically dependent on differentially methylated regions that span ~1–3 kb and acquire parent-of-origin specified epigenetic states during gametogenesis[133]. A seminal discovery in genomic imprinting was the identification of lncRNAs that are transcribed from the unmethylated allele in a differentially methylated region, and contribute to the repression of imprinted genes from the same locus[134–140].

Although imprinted lncRNAs exhibit considerable sequence and gene-structure diversity, they also share key similarities. Imprinted lncRNAs accumulate in the chromatin at the loci from which they are expressed[141], frequently exerting bidirectional, long-range silencing of multiple genes in *cis*[142]. Furthermore, transcription-based and RNA-based mechanisms have been proposed to cooperatively contribute to allele-specific transcription repression by these lncRNAs[142].

Imprinted lncRNAs cause allele-specific silencing of coding genes with which they overlap[143,144]. Genetic experiments in mice have shown that promoter deletions or premature transcription termination of imprinted lncRNAs disrupts their silencing functions[145–148]. For example, prevention of *Airn* or *Kcnq1ot1* transcription read-through into the coding genes *Igf2r* and *Kcnq1*, respectively, causes reactivation of the paternal alleles of these genes[146–150]. At the well-studied *Airn–Igf2r* locus, transcription interference has been ascribed either to promoter occlusion, where *Airn* transcription of *Igf2r* antisense prevents the recruitment of the transcription initiation machinery at the *Igf2r* promoter, or to a mechanism in which the *Airn* transcript actively removes the transcription machinery from the nascent *Igf2r* transcript[147] (Fig. 4). A related proposed mechanism is the collision of converging elongating Pol II complexes, exemplified by transcription of the lncRNA *Ube3a-ATS*, which leads to premature *Ube3a* termination at the paternal allele[151]. Support for this model comes from the observation that although both maternal and paternal *Ube3a* promoters are actively transcribed[152], either premature termination or antisense oligonucleotide-mediated degradation of paternal *Ube3a-ATS* prior to the overlap with the paternal *Ube3a* transcripts de-repress paternal *Ube3a* expression[152,153]. Despite the unequivocal evidence that antisense transcription is important for silencing by imprinted lncRNAs, it remains to be established why only some antisense transcripts evoke this effect, given that mammalian genomes harbour a myriad of convergent sense–antisense transcripts that are co-expressed in the same cells[154].

Imprinted loci, however, harbour multiple genes that do not overlap with imprinted lncRNAs yet show allele-specific silencing. In these cases, lncRNAs serve as scaffolds for other regulatory complexes. Repressed alleles at several imprinted loci are heavily enriched in H3K27me3, and several studies describe cell type-specific direct interactions of the lncRNAs *Airn, Kcnq1ot1* and *Meg3* with PRC1, PRC2 and the H3K9 methyltransferase G9a (also known as EHMT2)[139,155–162] (Fig. 4). Two recent studies have highlighted the role of *Airn* in PRC spreading over a 15 Mb domain in mouse trophoblast stem cells, showing a strong correlation between *Airn* expression, PRC occupancy at CpG islands and local changes in the chromatin architecture[137,163]. This finding is broadly consistent with some studies showing that RNA interactions are essential for genomic occupancy of PRC2 (ref. 164). Such mechanisms therefore explain how local RNA-dependent functions contribute to silencing of non-overlapping genes. One open question is whether the interaction of imprinted lncRNAs with protein-binding factors is mediated by specific lncRNA sequences and/or structures, as proposed for *XIST*.

Imprinted loci also undergo profound monoallelic changes in CTCF binding and local 3D genome organization, which can insulate genes from enhancers[165]. There is evidence that RNA–protein interactions are important for CTCF binding[166,167], and imprinted lncRNAs have been proposed to contribute to local allele-specific 3D genome changes, although more evidence is needed to define the precise role of imprinted lncRNAs in 3D genome organization[168,169].

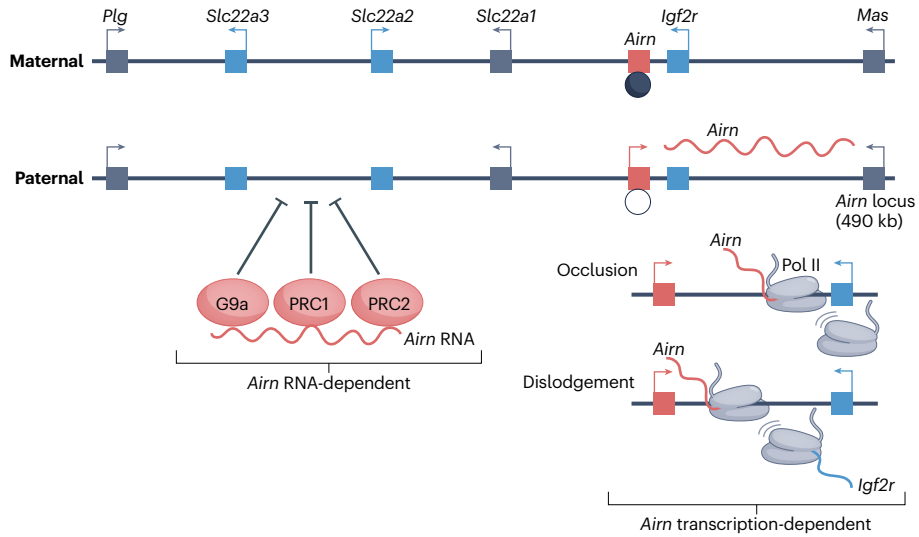


Fig. 4 | Long non-coding RNA-mediated allele-specific repression at imprinted loci. The mouse imprinted *Igf2r* locus, which includes the long non-coding RNA (lncRNA) *Airn*. Maternally expressed (blue), paternally expressed (red), and biallelic-expressed (grey) genes are shown. The circle underneath *Airn* denotes a differentially methylated region, with a black filling indicating methylation of the maternal allele. Bottom right: occlusion and dislodgement models for transcription-mediated repression of *Igf2r* by *Airn*. Bottom left: *Airn* RNA-dependent recruitment of the gene-repressing chromatin modifier complexes Polycomb repressive complex 1 (PRC1), PRC2 and G9a, which repress the *Airn*-distal genes *Slc22a2* and *Slc22a3*. *Airn* RNA is drawn as an unstructured molecule owing to lack of structural information. Pol II, polymerase II.

The analysis of imprinted lncRNAs, therefore, has offered unique insights into how transcription-based and RNA-based mechanisms cooperate to enact gene repression in *cis*.

Transcription regulation by lncRNAs in *trans*

Single-molecule imaging studies have shown that some lncRNA molecules are exclusively localized at their transcribed locus, whereas other lncRNAs disperse throughout the nucleus and could thus function in *trans*. In early studies, global changes in gene expression observed following lncRNA inhibition implied that the lncRNAs have such a *trans*-regulatory function, but in several cases the changes were later attributed to secondary events or to off-target effects of the perturbation tools. The validation of *trans*-regulatory functions of lncRNAs requires considerations such as the physiological stoichiometry of the lncRNA and its targets[29], evidence for direct engagement of the lncRNA with putative target regions and the ability to rescue loss-of-function phenotypes with exogenously expressed lncRNAs[170].

The archetypal *trans*-acting lncRNA is *HOTAIR*, which is expressed from the mouse and human HOXC locus and was proposed to regulate the expression of genes in the distant HOXD locus through PRC2 targeting[171]. A large *Hotair* deletion in mice confirmed a role in transcription repression of HoxD genes and several imprinted loci, leading to developmental defects[172]. However, the contribution of *Hotair* to homeotic transformation and the specificity of its interaction with PRC2 were challenged by subsequent studies, one of which used more selective mutations to show that *Hotair* RNA primarily acts as a negative regulator in *cis* of adjacent HoxC TF genes[73,173–175]. *Hotair* highlights the need for using complementary experimental tools to understand the function of *trans*-regulatory lncRNAs.

Global transcription control through nuclear assemblies

An emerging concept in transcription control is the role of nuclear compartmentalization of regulatory factors, mediated by interactions between lncRNAs and RBPs with intrinsically disordered regions[176]. Two well-characterized examples of lncRNA-containing nuclear assemblies are nuclear speckles and paraspeckles, which compartmentalize the highly abundant lncRNAs *MALAT1* and *NEAT1*, respectively, and have a role in the global regulation of transcription and RNA processing[177,178].

Recent studies have expanded the list of lncRNA-scaffolded nuclear assemblies (Fig. 5a). The intron-retaining, nuclear isoform of the lncRNA *Charme* specifically recruits MATR3–PTBP1 into nuclear aggregates that regulate chromatin at myogenic loci[179–181]. Another study directly visualized the definitive endoderm-specific lncRNA *DIGIT* (also known as *GSC-DT*) in condensates that contained the acetylated H3K18 reader, bromodomain-containing protein 3 (BRD3)[182]. Deletion of the retained intron of *Charme* or of *DIGIT* disrupted condensate formation and perturbed their respective downstream developmental programmes[182]. Analogously, the breast cancer-associated lncRNA mammary tumour-associated RNA 25 (*MaTAR25*) was found to interact with the complex purine-rich element-binding protein A (PURA)–PURB and was proposed to guide their association with the promoter of tensin-1, a key mediator of cell–matrix adhesion and metastatic migration[183]. How lncRNAs are targeted to one or many distant genomic sites remains an open question. In the context of nuclear assemblies, it is possible that locus specificity may be determined by either the lncRNA or the RBP.

Recent studies have also demonstrated a more general role for abundant nascent transcripts in maintaining regional chromatin compaction[184]. Analysis of chromatin-associated pre-mRNAs, lncRNAs and non-coding RNAs produced from repetitive regions has identified an RNA–protein scaffold that serves to counteract chromatin compaction and maintain active chromosome territories[184]. lncRNAs can also promote chromosomal reorganization by bringing genomic locations from different chromosomes in spatial proximity within the nucleus[185]. The X-linked lncRNA functional intergenic repeating RNA element (*Firre*) has been proposed to promote the formation of such an inter-chromosomal nuclear compartment, which contains co-regulated genes with a shared function in energy metabolism[186,187] (Fig. 5b).

Engagement of targets through triplex formation

Some lncRNAs were proposed to specifically target genomic locations through the formation of hybrid DNA–DNA–RNA triplex structures. Initially, this model was put forth to explain *cis*-regulation by overlapping antisense lncRNAs, such as *KHPS1* (refs. 188,189) and *PARTICLE*[190]. This model has been expanded to address genome-wide triplex formation based on computational identification of regions of lncRNA–DNA complementarity[191–193] or experimental identification

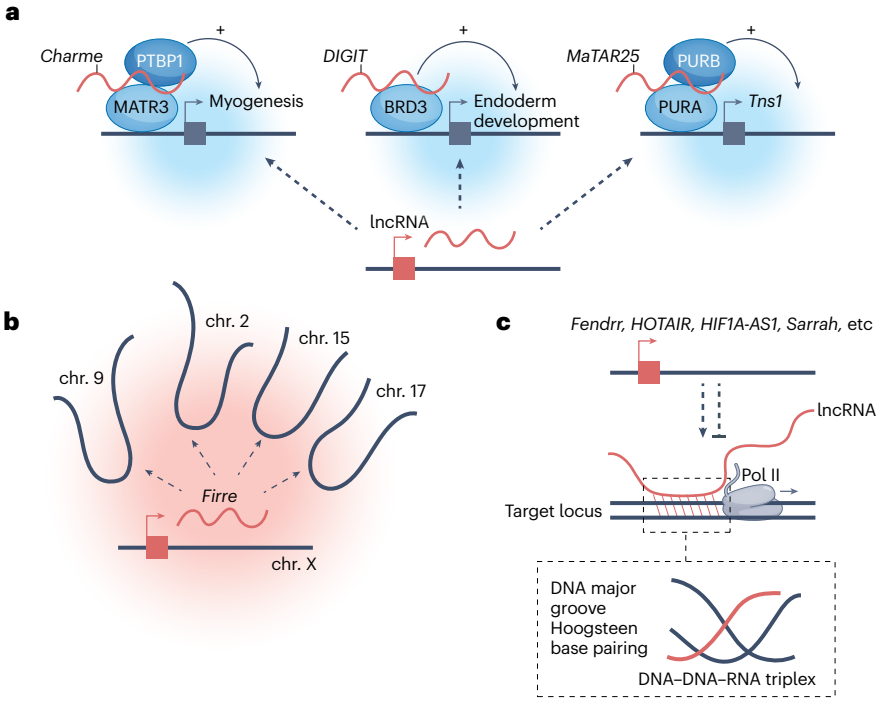


Fig. 5 | Mechanisms of transcription regulation by *trans*-acting long non-coding RNAs. a, Association of the long non-coding RNAs (lncRNAs) *Charme*, *DIGIT* and mammary tumour-associated RNA 25 (*MaTAR25*) with polypyrimidine tract-binding protein 1 (PTBP1)–matrin 3 (MATR3), bromodomain-containing protein 3 (BRD3) and purine-rich element-binding protein A (PURA)–PURB, respectively, drives condensate formation (blue background) and localization at target genes. This localization promotes the activation of broad developmental or disease-associated transcription programmes. **b**, The lncRNA *Firre* promotes inter-chromosomal contacts, which facilitates the co-regulation of a genes with shared functions in energy metabolism. **c**, Sequence complementarity between lncRNAs and one or many genomic regions enables targeting of the lncRNA to specific loci in *trans* through the formation of a DNA–DNA–(lnc)RNA triplex that involves DNA major groove Hoogstein base pairing. Various lncRNAs, including *Fendrr*, *HOTAIR*, *HIF1A-AS1*, *Sarrah* and others, have been shown to engage in triplex formation and exert positive or negative effects on target-gene expression. Pol II, polymerase II.

of RNAse H-resistant lncRNA–DNA heteroduplexes pulled down using an RNA–DNA-specific antibody[193,194]. Examples of lncRNAs proposed to engage this mechanism to repress or activate networks of genes in *trans* include *Fendrr* in mid-gestational embyos[191], *HOTAIR* in mesenchymal stem cells[192], *HIF1A-AS1* in endothelial cells[193] and *Sarrah* (also known as *Oxct1as*) in cardiomyocytes[194] (Fig. 5c). A recent study proposed an additional role for hybrid triplexes, showing that *KCNQ1OT1* forms triplexes to target gene-repressing complexes to transposable elements[195]. More work, however, is needed to define the extent to which DNA–DNA–RNA triplex structures are formed by lncRNAs for site-specific regulation at distant genomic sites.

Dual *cis* and *trans* regulation

Some lncRNAs have been shown to mediate regulatory activities both in *cis* and in *trans*. Notable examples include lncRNAs such as *MEG3*, which controls imprinting in *cis* but also mediates the p53 stress response[196] or engages in triplex formation[197]; the auxin-inducible lncRNA *Apolo* in forming R-loops in *cis* and in *trans* as a regulator of auxin-responsive genes in plants[198,199]; and the *cis*-activating *lincRNA-Cox2*, which controls the expression of the neighbouring gene *Ptgs2*, but also modulates the expression of a wide range of immune genes through an unknown mechanism[18]. In particular, the *trans* activity of *lincRNA-Cox2* was demonstrated in a mouse model by rescuing a *lincRNA-Cox2* deletion with a lncRNA-expressing transgene[18].

Roles of transcription-regulating lncRNAs in disease

As we begin to grasp the biological purpose of different types of regulatory lncRNAs, it becomes possible to explore their involvement in human Mendelian and polygenic diseases and oncogenesis, and their potential role as therapeutic targets.

Mendelian diseases

Identifying lncRNA gene mutations that cause Mendelian diseases poses major challenges because, unlike protein-coding sequences, there are no rules to predict the functionality of lncRNA sequence variants. Even in cases of lncRNA deletions, it is challenging to ascertain that phenotypes are not due to disruption of other functional elements, such as enhancers, located in the deleted region. Making this distinction usually requires complex genetic engineering approaches, such as combining deletions, transcription termination signals and insertion of RNA ribozymes.

Several lncRNA genes are located within genetically mapped loci that harbour Mendelian or monogenic mutations. For example, deletions encompassing the *FENDRR* locus lead to alveolar capillary dysplasia and misalignment of pulmonary veins (AVCD-MPV), although those deletions also disrupt elements that regulate the nearby gene *FOXF1*, which also harbours causal AVCD-MPV mutations[76,200] (Table 1). Nonetheless, mice with disrupted *Fendrr* transcription recapitulate

features of AVCD-MPV[68,201]. Likewise, variants in the lncRNA *HELLPAR* co-segregate with haemolysis, elevated liver enzymes, low platelets (HELLP) syndrome, although more conclusive evidence is needed to prove the causality of distinct variants[202]. These examples, together with knowledge that lncRNA genes often contain or are adjacent to enhancers, highlight some of the serious challenges facing efforts to demonstrate the pathogenicity of lncRNA defects in Mendelian diseases.

The analysis of *Maenli*, discussed above, has illustrated how a Mendelian phenotype can be followed up with careful mouse genetic studies to specifically assess the role of a lncRNA in the disease. Two children with a limb malformation were found to harbour the same homozygous 27 kb non-coding deletion in the *EN1* locus, whereas another individual with the same phenotype had a larger deletion in the same region in one allele, and an insertion in the other allele[17] (Fig. 6a). They examined the syntenic mouse sequence which contains *Maenli* lncRNA, and either deleted the *Maenli* promoter or blocked its transcription by inserting a polyadenylation sequence, both of which phenocopied the human disease[17]. These cases suggested that the severe human developmental phenotype was caused by germ-line deletions of a *cis*-activating lncRNA, raising the question of how many rare or common genetic variants in lncRNAs might influence human health.

Phenotypic relevance of transcription-stabilizing lncRNAs

So far, *FENDRR* and *PVT1* are the only lncRNAs that restrain the transcription of adjacent TF genes for which there is genetic evidence of a role in human disease[76,200]. It is reasonable to expect that large-scale functional screens will uncover many more *cis*-regulatory lncRNAs, and whole genome sequencing has the potential to discover genetic defects in lncRNAs. However, not all functional genetic elements are vulnerable to disruptive mutations. For example, enhancers are thought to be relatively robust to loss of function owing to functional redundancy[203–205].

Several considerations nevertheless indicate that genetic defects in stabilizer lncRNAs can result in strong phenotypes. Many of the TFs that are controlled by known stabilizer lncRNAs are very sensitive to gene dosage (Table 1). Thus, haploinsufficient germ-line mutations in *HAND2, CHD2*, *FOXF1, HNF1A* and *PITX2* cause Mendelian diseases, and in some of these cases an increased dosage in mice or humans also has phenotypic consequences[206–213] (Fig. 6b). Interestingly, somatic gain-of-function mutations or increased expression of *MYC*, the target of the lncRNA *PVT1* (Table 1), constitute an established oncogenic mechanism[214].

It is thus not surprising that mutations in lncRNAs that control the expression levels of these TFs frequently have phenotypes in mice (Fig. 6b). Different *Hand2os1* deletions cause either abnormal heart function or heart malformations and embryonic lethality in mice[66]. Likewise, *Hdnr* disruption leads to increased *Hand2* and severe cardiac malformations[67,215], and deletion of *Flicr* decreases susceptibility for autoimmune diabetes in non-obese diabetic mice[69]. Pancreatic or germ-line *Haster* mutations cause diabetes[70], and *Chaserr* null mutations cause embryonic lethality or severe growth retardation[71] (Table 1).

Importantly, these in vivo lncRNA phenotypes are not simply associated with silencing of their target TF genes, as would be expected if there was inadvertent disruption of an enhancer, but are instead linked to increased expression of the adjacent TF genes. Interestingly, most stabilizer lncRNA mutants with organismal phenotypes exhibit only moderate changes in the expression of their target TF genes, which underscores the importance of maintaining TF concentrations within a narrow range. The strong mutant phenotypes also suggest that, unlike enhancers, *cis*-inhibitory lncRNAs have limited built-in redundancy. Thus, despite the challenges in annotating lncRNA mutations that are deleterious in humans, mouse genetics indicate that lncRNA defects can lead to disease.

Polygenic diseases

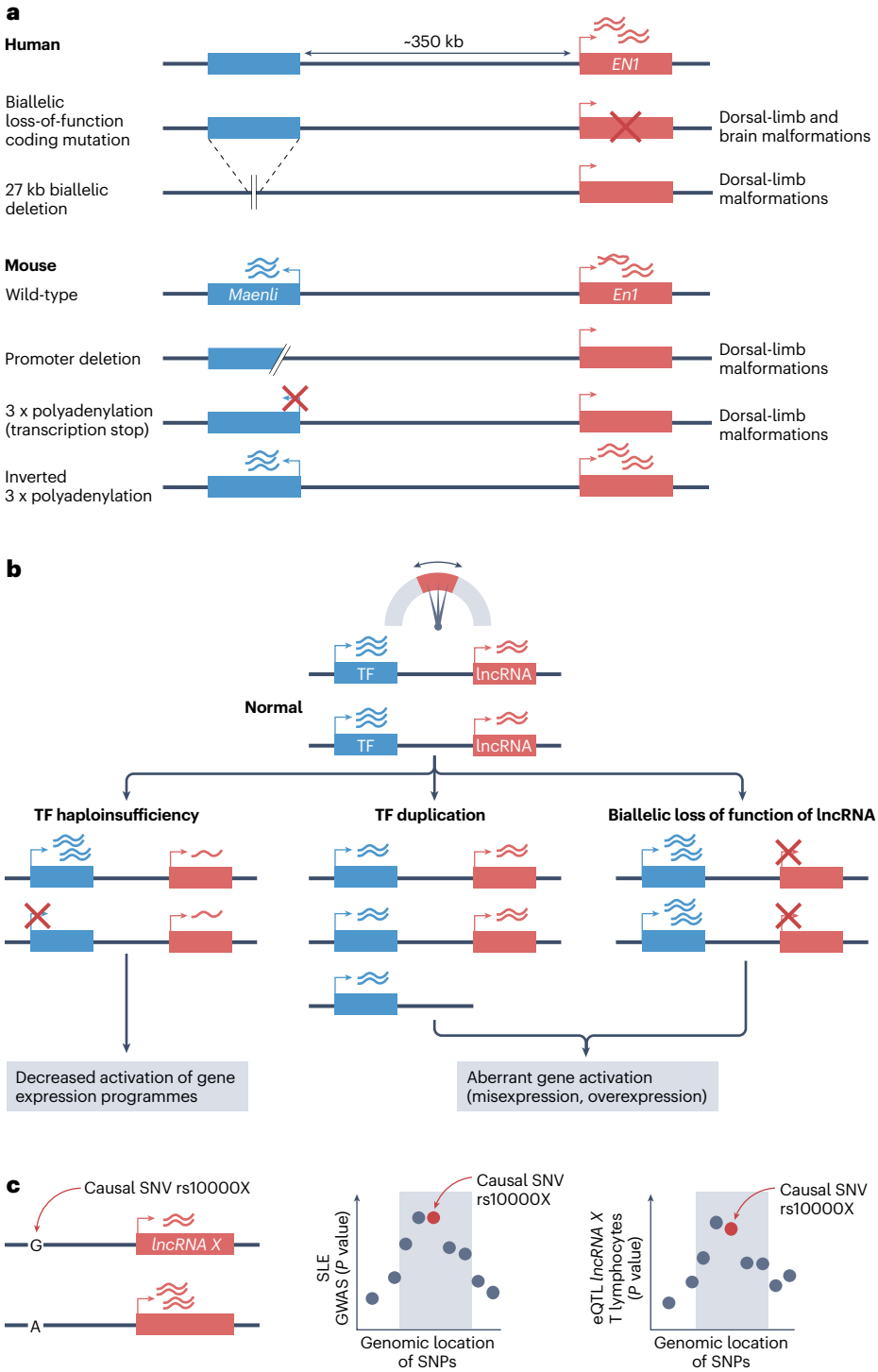
The most prevalent chronic human diseases, such as Alzheimer disease, coronary artery disease or type 2 diabetes, reflect the interplay of environmental factors with a large number of genetic variants. Although individual variants typically have very small effects on disease risk, the fact that they demonstrably influence human disease processes has the potential to shed light on causal mechanisms. A major fraction of susceptibility variants for common diseases identified in genome-wide association studies are non-coding, but the extent to which they act through lncRNAs is still unknown.

Numerous disease-risk variants have tight genetic co-localization with expression quantitative trait loci (eQTLs) for lncRNAs in disease-relevant cells (Fig. 6c). For example, common DNA variants that influence the expression of the lncRNA named *ANRIL* (also known as *CDKN2B-AS1*) co-localize with genetic association signals for coronary heart disease and type 2 diabetes[216,217], and pancreatic islet eQTLs for *LINC01512* (also known as *HI-LNC77*) as well as a splicing quantitative trait locus for *LINC00261* co-localize with type 2 diabetes genetic association signals[216]. A recent systematic analysis of expression and splicing quantitative trait loci from a broad panel of tissues revealed that DNA variants influencing the expression of more than 100 lncRNAs co-localize with variants associated with 66 polygenic phenotypes. For more than 50% of these loci, the effect on the lncRNA eQTL appears to be exclusive, or stronger than effects on any protein-coding gene eQTL[218]. These data warrant in-depth studies to examine how specific lncRNAs can act as molecular effectors of genetic susceptibility for common diseases.

lncRNA defects in cancer

Recurrent somatic copy number variants have been identified in several lncRNA loci. Examples include *PVT1* structural mutations in multiple cancer types[74,219,220], amplification of *FAL1* in approximately 10% of liver cancer[221], amplification of *SAMMSON* in 10% of melanoma[222] and loss of *ANRIL* in >50% of glioblastomas[223]. These regions, however, also harbour proto-oncogenes (*MYC*, *MCL1*, *MITF*) or the tumour suppressor *CDKN2A*, and are linked with enhancers, which hamper the ability to assess the pathogenic role of lncRNA defects[224]. Nonetheless, cancer-associated somatic structural variants such as focal deletions have been reported at the *PVT1* promoter region in breast cancer and in large B cell lymphomas, as well as chromosomal rearrangements that separate *PVT1* from *MYC*[74,225]. Furthermore, experimental deletions or transcription inhibition of the *PVT1* promoter cause high *MYC* expression and increased cellular proliferation[14,74].

In addition to these genetic changes, abnormal expression of many lncRNAs has been linked to cancer progression[224]. Well-studied examples are *MALAT1* overexpression, which is a strong predictor of metastasis in lung adenocarcinoma[226], and increased *HOTAIR* expression, which correlates with progression to metastasis and poor outcomes in breast cancer[227]. Abnormal expression of such lncRNAs could contribute to oncogenesis regardless of their function in normal physiology. For example, a recent preprint reports that in a mouse model of lung adenocarcinoma, *Malat1* overexpression is a driver of metastasis through



dysregulation of gene expression and reprogramming of the tumour microenvironment[228]. Other studies have shown that increased *HOTAIR* abundance can alter the stoichiometry of PRCs, resulting in deregulation of gene expression[227]. Furthermore, overexpression of *cis*-acting lncRNAs, such as the imprinted lncRNA *H19* (refs. 229–231), the gene-inhibiting lncRNAs *PVT1* (refs. 85,232) and *MORRBID*[233,234], and the gene-activating *LINCRNA-P21* (ref. 235) also have oncogenic *trans*-regulatory activities.

X-chromosome inactivation defects

XCI is essential for development[236–238] and conditional mouse deletions of *Xist* in the haematopoietic system cause aberrant epigenetic states and oncogenic transformation[239,240]. However, recent studies have revealed that XCI is not permanent in all cell lineages, as reversals can be observed in specific adult immune cell subtypes[241]. This reversal has been proposed to underlie the female-specific predisposition for

Fig. 6 | Involvement of long non-coding RNAs in genetic diseases. a, The human *EN1* locus, which encodes a homeobox transcription factor (TF). *EN1* harbours recessive coding mutations in an individual with limb and brain malformations, whereas far-upstream biallelic non-coding deletions (or a compound heterozygous deletion and insertion not shown here) cause dorsal-limb malformations. *Maenli* is a mouse long non-coding RNA (lncRNA) mapped to the orthologous minimal deleted region in humans. Deletion of the *Maenli* promoter, or insertion of polyadenylation signals of transcription termination, recapitulate limb malformations and lead to reduced *En1* expression, whereas an inverted termination signal has no effect. **b**, Transcription stabilizers control dosage-sensitive TF genes. Small deviations in the expression levels of certain TFs can be caused by heterozygous loss-of-function mutations or duplications of TF genes, or by biallelic loss of function of the stabilizer lncRNA, causing abnormal cellular transcription with organismal phenotypes. In several examples, the defects in stabilizer lncRNAs and the dosage alterations of their target genes have the same phenotype (Table 1). **c**, An expression quantitative trait locus (eQTL), in which a single nucleotide variant (SNV) influences expression of a lncRNA (left). rs10000X is depicted as the identifier of a fictitious regulatory SNV that is causal for this eQTL. The two graphs on the right depict statistically significant *P* values of a group of adjacent SNVs for association with the presence of the autoimmune disease systemic lupus erythematosus (SLE) (top) or with a lncRNA eQTL detected in T cell lymphocytes, a cell type that is relevant to SLE (bottom). The co-localization of both sets of association *P* values means that the lncRNA is a plausible mediator of the disease association. Several hundred instances such as this have been identified, indicating that variation in lncRNA expression contributes to the susceptibility of common diseases. GWAS, genome-wide association studies.

autoimmune diseases through gene dosage increase from the X chromosome, which is known to have a high density of immunity regulating genes[241–245]. Another recent study, reported in a preprint, has directly implicated the immunogenicity of *XIST* ribonucleoprotein complexes in autoimmune disease mechanisms[246].

Therapeutic modulation of lncRNAs

Regardless of whether a disease is caused by a lncRNA defect, it is sometimes possible to envision therapeutic targeting of a lncRNA to modulate a disease-relevant process, using sequence-specific RNA degradation, RNA mimetics or genome editing tools that control the transcription of lncRNAs[247,248]. A good example is Angelman syndrome, which is an imprinting neurodevelopmental disorder caused by deletions or mutations of the active maternal *UBE3A* allele. Several clinical trials are underway to activate the silenced paternal *UBE3A* allele with antisense oligonucleotides that block or cause degradation of the lncRNA *UBE3A-ATS*, following proof-of-concept studies in model systems[153,248,249].

Another important therapeutic application has been the use of a lncRNA to correct abnormal gene dosage. Insertion of an inducible *XIST* transgene into one copy of chromosome 21 in induced pluripotent stem cells from an individual with trisomy 21 was shown to cause silencing in *cis* of the chromosome, and reversal of major transcriptional and other cellular defects[250].

Small molecules that target and modulate the activity and stability of lncRNAs represent an additional therapeutic avenue. Compounds have been developed to target the stabilizing 3′-end triple helix structure of *MALAT1*, given its strong association with metastasis in solid tumours[251–253]. Another example is X1, a small molecule that binds *XIST* A-repeats and displaces PRC2 and SPEN, thereby blocking XCI[254]. This approach has been proposed to de-repress wild-type alleles in X-linked disorders such as Rett syndrome.

Conclusion and future perspective

Ever since the discovery of lncRNAs, efforts to understand their biological significance have met daunting challenges. Investigations started from a blank slate, with no sense of what type of molecular entities might exist under the lncRNA umbrella, or which experimental tools could be used to elucidate their function. Despite these obstacles, the past years have witnessed major breakthroughs, many of which have come from exhaustive systematic efforts to dissect the function of single lncRNAs. These studies have led to the discovery of a spectrum of lncRNA functions, including essential gene-activating functions, and specialized *cis*-regulatory feedback mechanisms.

These findings have also raised a long list of pressing new questions. For example, what fraction of the catalogued lncRNAs is functional, and how many lncRNAs belong to the categories identified so far? How many functional lncRNAs act through RNA-dependent mechanisms, as opposed to those that primarily involve the activation of lncRNA promoters or transcription? For lncRNAs with an ascribed function, our knowledge of the underlying mechanisms is fragmented, which surely explains several apparent paradoxes, such as the observation that the transcription of different lncRNAs can lead to either silencing or activation of their antisense genes. Likewise, transcription activation of lncRNAs has been shown to be essential for both *cis*-activating and inhibiting functions, but the rules that underlie these outcomes are not well understood. Another crucial gap in the field is our lack of knowledge of the major sequence determinants of lncRNA functions, and their vulnerability to disease-causing variation. Understanding the molecular underpinnings of different types of functional lncRNAs, combined with knowledge of which lncRNAs act in disease-relevant processes, holds promise for the development of new therapeutic strategies.

Published online: 19 January 2024

References

1. Lipshitz, H. D., Peattie, D. A. & Hogness, D. S. Novel transcripts from the ultrabithorax domain of the bithorax complex. *Genes. Dev.* **1**, 307–322 (1987).
2. Cumberledge, S., Zaratzian, A. & Sakonju, S. Characterization of two RNAs transcribed from the *cis*-regulatory region of the abd-A domain within the *Drosophila* bithorax complex. *Proc. Natl Acad. Sci. USA* **87**, 3259–3263 (1990).
3. Mattick, J. S. et al. Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* **24**, 430–447 (2023).
4. Anderson, D. M. et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell* **160**, 595–606 (2015).
5. Morgado-Palacin, L. et al. The TINCR ubiquitin-like microprotein is a tumor suppressor in squamous cell carcinoma. *Nat. Commun.* **14**, 1328 (2023).
6. Frankish, A. et al. GENCODE: reference annotation for the human and mouse genomes in 2023. *Nucleic Acids Res.* **51**, D942–D949 (2023).
7. Gil, N. & Ulitsky, I. Regulation of gene expression by *cis*-acting long non-coding RNAs. *Nat. Rev. Genet.* **21**, 102–117 (2020).
8. Rinn, J. L. & Chang, H. Y. Long noncoding RNAs: molecular modalities to organismal functions. *Annu. Rev. Biochem.* **89**, 283–308 (2020).
9. Moran, I. et al. Human β-cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. *Cell Metab.* **16**, 435–448 (2012).
10. Cabili, M. N. et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes. Dev.* **25**, 1915–1927 (2011).
11. Hon, C. C. et al. An atlas of human long non-coding RNAs with accurate 5′ ends. *Nature* **543**, 199–204 (2017).
12. Akerman, I. et al. Human pancreatic β cell lncRNAs control cell-specific regulatory networks. *Cell Metab.* **25**, 400–411 (2017).
13. Dimitrova, N. et al. *LincRNA-p21* activates *p21* in *cis* to promote Polycomb target gene expression and to enforce the G1/S checkpoint. *Mol. Cell* **54**, 777–790 (2014).

14. Olivero, C. E. et al. p53 activates the long noncoding RNA Pvt1b to inhibit myc and suppress tumorigenesis. *Mol. Cell* **77**, 761–774.e8 (2020).
This study identifies a p53-induced isoform of the lncRNA *Pvt1*, which acts in *cis* to suppress *Myc* transcription in response to stress and thus limits cellular proliferation.
15. Winkler, L. et al. Functional elements of the *cis*-regulatory *lincRNA-p21*. *Cell Rep.* **39**, 110687 (2022).
Systematic genetic dissection of the *lincRNA-p21* locus reveals that transcription initiation of *lincRNA-p21* is sufficient for stimulation of *p21* expression in *cis*.
16. Gil, N. et al. Complex regulation of Eomes levels mediated through distinct functional features of the Meteor long non-coding RNA locus. *Cell Rep.* **42**, 112569 (2023).
17. Allou, L. et al. Non-coding deletions identify Maenli lncRNA as a limb-specific *En1* regulator. *Nature* **592**, 93–98 (2021).
Genetically engineered mouse models reveal that transcript elongation of the lncRNA *Maenli* promotes *En1* expression and supports limb development. This study demonstrates that a human developmental limb disorder is likely caused by a monogenic lncRNA defect.
18. Elling, R. et al. Genetic models reveal *cis* and *trans* immune-regulatory activities for lincRNA-Cox2. *Cell Rep.* **25**, 1511–1524.e6 (2018).
19. Isoda, T. et al. Non-coding transcription instructs chromatin folding and compartmentalization to dictate enhancer–promoter communication and T cell fate. *Cell* **171**, 103–119.e18 (2017).
20. Perry, R. B., Hezroni, H., Goldrich, M. J. & Ulitsky, I. Regulation of neuroregeneration by long noncoding RNAs. *Mol. Cell* **72**, 553–567.e5 (2018).
21. De Santa, F. et al. A large fraction of extragenic RNA Pol II transcription sites overlap enhancers. *PLoS Biol.* **8**, e1000384 (2010).
22. Andersson, R. et al. An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014).
23. Heintzman, N. D. et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**, 311–318 (2007).
24. Paralkar, V. R. et al. Lineage and species-specific long noncoding RNAs during erythro-megakaryocytic development. *Blood* **123**, 1927–1937 (2014).
25. Paralkar, V. R. et al. Unlinking an lncRNA from its associated *cis* element. *Mol. Cell* **62**, 104–110 (2016).
26. Espinosa, J. M. Revisiting lncRNAs: how do you know yours is not an eRNA? *Mol. Cell* **62**, 1–2 (2016).
27. Wang, K. C. et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* **472**, 120–124 (2011).
28. Pradeepa, M. M. et al. Psip1/p52 regulates posterior Hoxa genes through activation of lncRNA Hottip. *PLoS Genet.* **13**, e1006677 (2017).
29. Unfried, J. P. & Ulitsky, I. Substoichiometric action of long noncoding RNAs. *Nat. Cell Biol.* **24**, 608–615 (2022).
30. Henninger, J. E. et al. RNA-mediated feedback control of transcriptional condensates. *Cell* **184**, 207–225.e24 (2021).
31. Sharp, P. A., Chakraborty, A. K., Henninger, J. E. & Young, R. A. RNA in formation and regulation of transcriptional condensates. *RNA* **28**, 52–57 (2022).
32. Oksuz, O. et al. Transcription factors interact with RNA to regulate genes. *Mol. Cell* **83**, 2449–2463.e13 (2023).
33. Ntini, E. et al. Long ncRNA A-ROD activates its target gene DKK1 at its release from chromatin. *Nat. Commun.* **9**, 1636 (2018).
34. Huarte, M. et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **142**, 409–419 (2010).
35. Groff, A. F. et al. In vivo characterization of Linc-p21 reveals functional *cis*-regulatory DNA elements. *Cell Rep.* **16**, 2178–2186 (2016).
36. Furuhata, R. et al. LincRNA-p21 exon 1 expression correlates with Cdkn1a expression in vivo. *Genes. Cell* **27**, 14–24 (2022).
37. Engreitz, J. M. et al. Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature* **539**, 452–455 (2016).
38. Gil, N. & Ulitsky, I. Production of spliced long noncoding RNAs specifies regions with increased enhancer activity. *Cell Syst.* **7**, 537–547.e3 (2018).
39. Haerty, W. & Ponting, C. P. Unexpected selection to retain high GC content and splicing enhancers within exons of multiexonic lncRNA loci. *RNA* **21**, 333–346 (2015).
40. Schuler, A., Ghanbarian, A. T. & Hurst, L. D. Purifying selection on splice-related motifs, not expression level nor RNA folding, explains nearly all constraint on human lincRNAs. *Mol. Biol. Evol.* **31**, 3164–3183 (2014).
41. Tan, J. Y. & Marques, A. C. The activity of human enhancers is modulated by the splicing of their associated lncRNAs. *PLoS Comput. Biol.* **18**, e1009722 (2022).
42. Canzio, D. et al. Antisense lncRNA transcription mediates DNA demethylation to drive stochastic protocadherin α promoter choice. *Cell* **177**, 639–653.e15 (2019).
43. Heinz, S. et al. Transcription elongation can affect genome 3D structure. *Cell* **174**, 1522–1536.e22 (2018).
44. Zhang, S., Ubelmesser, N., Barbieri, M. & Papantonis, A. Enhancer–promoter contact formation requires RNAPII and antagonizes loop extrusion. *Nat. Genet.* **55**, 832–840 (2023).
45. Banigan, E. J. et al. Transcription shapes 3D chromatin organization by interacting with loop extrusion. *Proc. Natl Acad. Sci. USA* **120**, e2210480120 (2023).
46. Lai, F. et al. Activating RNAs associate with mediator to enhance chromatin architecture and transcription. *Nature* **494**, 497–501 (2013).
47. Li, W. et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* **498**, 516–520 (2013).
48. Melo, C. A. et al. eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol. Cell* **49**, 524–535 (2013).
49. Orom, U. A. et al. Long noncoding RNAs with enhancer-like function in human cells. *Cell* **143**, 46–58 (2010).
50. Kalantari, R., Chiang, C. M. & Corey, D. R. Regulation of mammalian transcription and splicing by nuclear RNAi. *Nucleic Acids Res.* **44**, 524–537 (2016).
51. Khvorova, A. Modulation of DNA transcription: the future of ASO therapeutics? *Cell* **185**, 2011–2013 (2022).
52. Marasco, L. E. et al. Counteracting chromatin effects of a splicing-correcting antisense oligonucleotide improves its therapeutic efficacy in spinal muscular atrophy. *Cell* **185**, 2057–2070.e15 (2022).
53. Lee, J. H. et al. Enhancer RNA m^6A methylation facilitates transcriptional condensate formation and gene activation. *Mol. Cell* **81**, 3368–3385.e9 (2021).
54. Rahnamoun, H. et al. RNAs interact with BRD4 to promote enhanced chromatin engagement and transcription activation. *Nat. Struct. Mol. Biol.* **25**, 687–697 (2018).
55. Liang, L. et al. Complementary Alu sequences mediate enhancer–promoter selectivity. *Nature* **619**, 868–875 (2023).
56. Barshad, G. et al. RNA polymerase II dynamics shape enhancer–promoter interactions. *Nat. Genet.* **55**, 1370–1380 (2023).
57. Dao, L. T. M. et al. Genome-wide characterization of mammalian promoters with distal enhancer functions. *Nat. Genet.* **49**, 1073–1081 (2017).
58. Guttman, M. et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227 (2009).
59. Ulitsky, I., Shkumatava, A., Jan, C. H., Sive, H. & Bartel, D. P. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**, 1537–1550 (2011).
60. Liu, N. et al. Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. *Cell* **173**, 430–442.e17 (2018).
61. Pang, B., van Weerd, J. H., Hamoen, F. L. & Snyder, M. P. Identification of non-coding silencer elements and their regulation of gene expression. *Nat. Rev. Mol. Cell Biol.* **24**, 383–395 (2023).
62. Maamar, H., Cabili, M. N., Rinn, J. & Raj, A. linc-HOXA1 is a noncoding RNA that represses Hoxa1 transcription in *cis*. *Genes. Dev.* **27**, 1260–1271 (2013).
63. Su, G. et al. Enhancer architecture-dependent multilayered transcriptional regulation orchestrates RA signaling-induced early lineage differentiation of ESCs. *Nucleic Acids Res.* **49**, 11575–11595 (2021).
64. Yin, Y. et al. Opposing roles for the lncRNA haunt and its genomic locus in regulating HOXA gene activation during embryonic stem cell differentiation. *Cell Stem Cell* **16**, 504–516 (2015).
65. Anderson, K. M. et al. Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. *Nature* **539**, 433–436 (2016).
66. Han, X. et al. The lncRNA *Hand2os1/Uph* locus orchestrates heart development through regulation of precise expression of *Hand2*. *Development* **146**, dev176198 (2019).
67. Ritter, N. et al. The lncRNA locus handsdown regulates cardiac gene programs and is essential for early mouse development. *Dev. Cell* **50**, 644–657.e8 (2019).
68. Grote, P. et al. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev. Cell* **24**, 206–214 (2013).
69. Zemmour, D., Pratama, A., Loughhead, S. M., Mathis, D. & Benoist, C. Flicr, a long noncoding RNA, modulates Foxp3 expression and autoimmunity. *Proc. Natl Acad. Sci. USA* **114**, E3472–E3480 (2017).
70. Beucher, A. et al. The HASTER lncRNA promoter is a *cis*-acting transcriptional stabilizer of HNF1A. *Nat. Cell Biol.* **24**, 1528–1540 (2022).
This study shows that the promoter of the lncRNA *HASTER* ensures that levels of the transcripion factor HNF1A are maintained within a narrow homeostatic range. *Haster* deficiency causes abnormal HNF1A genomic occupancy and diabetes in mice.
71. Rom, A. et al. Regulation of CHD2 expression by the Chaserr long noncoding RNA gene is essential for viability. *Nat. Commun.* **10**, 5092 (2019).
72. Bond, A. M. et al. Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. *Nat. Neurosci.* **12**, 1020–1027 (2009).
73. Amandio, A. R., Necsulea, A., Joye, E., Mascrez, B. & Duboule, D. Hotair is dispensible for mouse development. *PLoS Genet.* **12**, e1006232 (2016).
74. Cho, S. W. et al. Promoter of lncRNA gene PVT1 is a tumor-suppressor DNA boundary element. *Cell* **173**, 1398–1412.e22 (2018).
75. Chen, F. L. et al. The long noncoding RNA *Playrr* regulates *Pitx2* dosage and protects against cardiac arrhythmias. Preprint at *bioRxiv* https://doi.org/10.1101/2022.09.20.508562 (2022).
76. Szafranski, P., Gambin, T., Karolak, J. A., Popek, E. & Stankiewicz, P. Lung-specific distant enhancer *cis* regulates expression of FOXF1 and lncRNA FENDRR. *Hum. Mutat.* **42**, 694–698 (2021).
77. Ali, T. & Grote, P. Beyond the RNA-dependent function of lncRNA genes. *eLife* **9**, e60583 (2020).
78. Ghildiyal, R. et al. Loss of long noncoding RNA NXTAR in prostate cancer augments androgen receptor expression and enzalutamide resistance. *Cancer Res.* **82**, 155–168 (2022).
79. Kribelbauer, J. F., Rastogi, C., Bussemaker, H. J. & Mann, R. S. Low-affinity binding sites and the transcription factor specificity paradox in eukaryotes. *Annu. Rev. Cell Dev. Biol.* **35**, 357–379 (2019).
80. Golson, M. L. & Kaestner, K. H. Fox transcription factors: from development to disease. *Development* **143**, 4558–4570 (2016).

81. Servitja, J. M. et al. Hnf1α (MODY3) controls tissue-specific transcriptional programs and exerts opposed effects on cell growth in pancreatic islets and liver. *Mol. Cell Biol.* **29**, 2945–2959 (2009).
82. Fernandez Garcia, M. et al. Structural features of transcription factors associating with nucleosome binding. *Mol. Cell* **75**, 921–932.e6 (2019).
83. Huang, P. et al. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature* **475**, 386–389 (2011).
84. Ali, T. et al. Fendrr synergizes with Wnt signalling to regulate fibrosis related genes during lung development via its RNA:dsDNA triplex element. *Nucleic Acids Res.* **51**, 6227–6237 (2023).
85. Colombo, T., Farina, L., Macino, G. & Paci, P. PVT1: a rising star among oncogenic long noncoding RNAs. *Biomed. Res. Int.* **2015**, 304208 (2015).
86. Tesfaye, E. et al. The p53 transcriptional response across tumor types reveals core and senescence-specific signatures modulated by long noncoding RNAs. *Proc. Natl Acad. Sci. USA* **118**, e2025539118 (2021).
87. Kotzin, J. J. et al. The long noncoding RNA morrbid regulates CD8 T cells in response to viral infection. *Proc. Natl Acad. Sci. USA* **116**, 11916–11925 (2019).
88. Kotzin, J. J. et al. The long non-coding RNA morrbid regulates Bim and short-lived myeloid cell lifespan. *Nature* **537**, 239–243 (2016).
89. Zhao, Y. et al. Natural temperature fluctuations promote COOLAIR regulation of FLC. *Genes. Dev.* **35**, 888–898 (2021).
90. Jegu, T., Aeby, E. & Lee, J. T. The X chromosome in space. *Nat. Rev. Genet.* **18**, 377–389 (2017).
91. Deng, X., Berletch, J. B., Nguyen, D. K. & Disteche, C. M. X chromosome regulation: diverse patterns in development, tissues and disease. *Nat. Rev. Genet.* **15**, 367–378 (2014).
92. Brockdorff, N., Bowness, J. S. & Wei, G. Progress toward understanding chromosome silencing by Xist RNA. *Genes. Dev.* **34**, 733–744 (2020).
93. Augui, S., Nora, E. P. & Heard, E. Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat. Rev. Genet.* **12**, 429–442 (2011).
94. Galupa, R. & Heard, E. X-chromosome inactivation: a crossroads between chromosome architecture and gene regulation. *Annu. Rev. Genet.* **52**, 535–566 (2018).
95. Furlan, G. & Galupa, R. Mechanisms of choice in X-chromosome inactivation. *Cells* **11**, 535 (2022).
96. Mutzel, V. & Schulz, E. G. Dosage sensing, threshold responses, and epigenetic memory: a systems biology perspective on random X-chromosome inactivation. *Bioessays* **42**, e1900163 (2020).
97. Jacobson, E. C., Pandya-Jones, A. & Plath, K. A lifelong duty: how Xist maintains the inactive X chromosome. *Curr. Opin. Genet. Dev.* **75**, 101927 (2022).
98. van Bemmel, J. G. et al. The bipartite TAD organization of the X-inactivation center ensures opposing developmental regulation of Tsix and Xist. *Nat. Genet.* **51**, 1024–1034 (2019).
99. Gjaltema, R. A. F. et al. Distal and proximal *cis*-regulatory elements sense X chromosome dosage and developmental state at the Xist locus. *Mol. Cell* **82**, 190–208.e17 (2022).
100. Rosspopoff, O. et al. Species-specific regulation of XIST by the JPX/FTX orthologs. *Nucleic Acids Res.* **51**, 2177–2194 (2023).
Functional similarities and differences between the human and mouse lncRNAs orthologues *JPX* and *Jpx* and *FTX* and *Ftx* highlight the complementary roles of lncRNA transcription and the mature lncRNAs in XCI.
101. Quesada-Espinosa, J. F. et al. First female with Allan–Herndon–Dudley syndrome and partial deletion of X-inactivation center. *Neurogenetics* **22**, 343–346 (2021).
102. Sun, S. et al. Jpx RNA activates Xist by evicting CTCF. *Cell* **153**, 1537–1551 (2013).
103. Loda, A., Collombet, S. & Heard, E. Gene regulation in time and space during X-chromosome inactivation. *Nat. Rev. Mol. Cell Biol.* **23**, 231–249 (2022).
104. Ogawa, Y. & Lee, J. T. Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol. Cell* **11**, 731–743 (2003).
105. Galupa, R. et al. A conserved noncoding locus regulates random monoallelic xist expression across a topological boundary. *Mol. Cell* **77**, 352–367.e8 (2020).
106. Lee, J. T. & Lu, N. Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell* **99**, 47–57 (1999).
107. Wutz, A. & Jaenisch, R. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell* **5**, 695–705 (2000).
108. Lee, J. T., Strauss, W. M., Dausman, J. A. & Jaenisch, R. A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell* **86**, 83–94 (1996).
109. Chu, C. et al. Systematic discovery of Xist RNA binding proteins. *Cell* **161**, 404–416 (2015).
110. Markaki, Y. et al. Xist nucleates local protein gradients to propagate silencing across the X chromosome. *Cell* **184**, 6212 (2021).
111. McHugh, C. A. et al. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521**, 232–236 (2015).
112. Minajigi, A. et al. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **349**, aab2276 (2015).
113. Raposo, A. C., Casanova, M., Gendrel, A. V. & da Rocha, S. T. The tandem repeat modules of Xist lncRNA: a Swiss army knife for the control of X-chromosome inactivation. *Biochem. Soc. Trans.* **49**, 2549–2560 (2021).
114. Carter, A. C. et al. Spen links RNA-mediated endogenous retrovirus silencing and X chromosome inactivation. *eLife* **9**, e54508 (2020).
115. Dossin, F. et al. SPEN integrates transcriptional and epigenetic control of X-inactivation. *Nature* **578**, 455–460 (2020).
116. Jachowicz, J. W. et al. Xist spatially amplifies SHARP/SPEN recruitment to balance chromosome-wide silencing and specificity to the X chromosome. *Nat. Struct. Mol. Biol.* **29**, 239–249 (2022).
117. Zylicz, J. J. et al. The implication of early chromatin changes in X chromosome inactivation. *Cell* **176**, 182–197.e23 (2019).
118. Bousard, A. et al. The role of Xist-mediated Polycomb recruitment in the initiation of X-chromosome inactivation. *EMBO Rep.* **20**, e48019 (2019).
119. Jansz, N. et al. Smchd1 targeting to the inactive X is dependent on the Xist–HnrnpK–PRC1 pathway. *Cell Rep.* **25**, 1912–1923.e9 (2018).
120. Pintacuda, G. et al. hnRNPK recruits PCGF3/5–PRC1 to the Xist RNA B-repeat to establish polycomb-mediated chromosomal silencing. *Mol. Cell* **68**, 955–969.e10 (2017).
121. Wang, C. Y., Colognori, D., Sunwoo, H., Wang, D. & Lee, J. T. PRC1 collaborates with SMCHD1 to fold the X-chromosome and spread Xist RNA between chromosome compartments. *Nat. Commun.* **10**, 2950 (2019).
122. Wang, C. Y., Jegu, T., Chu, H. P., Oh, H. J. & Lee, J. T. SMCHD1 merges chromosome compartments and assists formation of super-structures on the inactive X. *Cell* **174**, 406–421.e25 (2018).
123. Colognori, D., Sunwoo, H., Wang, D., Wang, C. Y. & Lee, J. T. Xist repeats A and B account for two distinct phases of X inactivation establishment. *Dev. Cell* **54**, 21–32.e5 (2020).
124. Pandya-Jones, A. et al. A protein assembly mediates Xist localization and gene silencing. *Nature* **587**, 145–151 (2020).
Assembly of multiple RNA-binding proteins on *Xist* E-repeats promotes homotypic and heterotypic interactions that result in the formation of a condensate, which is essential for gene silencing. Once formed, this condensate can sustain XCI in absence of *Xist*.
125. Sunwoo, H., Colognori, D., Froberg, J. E., Jeon, Y. & Lee, J. T. Repeat E anchors Xist RNA to the inactive X chromosomal compartment through CDKN1A-interacting protein (CIZ1). *Proc. Natl Acad. Sci. USA* **114**, 10654–10659 (2017).
126. Strehle, M. & Guttman, M. Xist drives spatial compartmentalization of DNA and protein to orchestrate initiation and maintenance of X inactivation. *Curr. Opin. Cell Biol.* **64**, 139–147 (2020).
127. de Napoles, M. et al. Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7**, 663–676 (2004).
128. Sun, B. K., Deaton, A. M. & Lee, J. T. A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. *Mol. Cell* **21**, 617–628 (2006).
129. Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J. & Lee, J. T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322**, 750–756 (2008).
130. Engreitz, J. M. et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**, 1237973 (2013).
131. Simon, M. D. et al. High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* **504**, 465–469 (2013).
132. Tucci, V., Isles, A. R., Kelsey, G., Ferguson-Smith, A. C. & Erice Imprinting, G. Genomic imprinting and physiological processes in mammals. *Cell* **176**, 952–965 (2019).
133. Barlow, D. P. & Bartolomei, M. S. Genomic imprinting in mammals. *Cold Spring Harb. Perspect. Biol.* **6**, a018382 (2014).
134. Guenzl, P. M. & Barlow, D. P. Macro lncRNAs: a new layer of *cis*-regulatory information in the mammalian genome. *RNA Biol.* **9**, 731–741 (2012).
135. Latos, P. A. & Barlow, D. P. Regulation of imprinted expression by macro non-coding RNAs. *RNA Biol.* **6**, 100–106 (2009).
136. Kota, S. K. et al. ICR noncoding RNA expression controls imprinting and DNA replication at the Dlk1–Dio3 domain. *Dev. Cell* **31**, 19–33 (2014).
137. Schertzer, M. D. et al. lncRNA-induced spread of Polycomb controlled by genome architecture, RNA abundance, and CpG island DNA. *Mol. Cell* **75**, 523–537.e10 (2019).
138. Seidl, C. I., Stricker, S. H. & Barlow, D. P. The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *EMBO J.* **25**, 3565–3575 (2006).
139. Terranova, R. et al. Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. *Dev. Cell* **15**, 668–679 (2008).
140. Tibbit, C. J. et al. Antisense activity across the *Nesp* promoter is required for *Nespas*-mediated silencing in the imprinted *Gnas* cluster. *Noncoding RNA* **1**, 246–265 (2015).
141. Quinodoz, S. A. et al. RNA promotes the formation of spatial compartments in the nucleus. *Cell* **184**, 5775–5790.e30 (2021).
142. MacDonald, W. A. & Mann, M. R. W. Long noncoding RNA functionality in imprinted domain regulation. *PLoS Genet.* **16**, e1008930 (2020).
143. Pauler, F. M., Koerner, M. V. & Barlow, D. P. Silencing by imprinted noncoding RNAs: is transcription the answer? *Trends Genet.* **23**, 284–292 (2007).
144. Hao, N., Palmer, A. C., Dodd, I. B. & Shearwin, K. E. Directing traffic on DNA — how transcription factors relieve or induce transcriptional interference. *Transcription* **8**, 120–125 (2017).
145. Andergassen, D. et al. The Airn lncRNA does not require any DNA elements within its locus to silence distant imprinted genes. *PLoS Genet.* **15**, e1008268 (2019).
146. Latos, P. A. et al. Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science* **338**, 1469–1472 (2012).
147. Santoro, F. et al. Imprinted Igf2r silencing depends on continuous Airn lncRNA expression and is not restricted to a developmental window. *Development* **140**, 1184–1195 (2013).

148. Sleutels, F., Zwart, R. & Barlow, D. P. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415**, 810–813 (2002).
149. Golding, M. C. et al. Depletion of Kcnq1ot1 non-coding RNA does not affect imprinting maintenance in stem cells. *Development* **138**, 3667–3678 (2011).
150. Mancini-Dinardo, D., Steele, S. J., Levorse, J. M., Ingram, R. S. & Tilghman, S. M. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes. Dev.* **20**, 1268–1282 (2006).
151. Meng, L., Person, R. E. & Beaudet, A. L. Ube3a-ATS is an atypical RNA polymerase II transcript that represses the paternal expression of Ube3a. *Hum. Mol. Genet.* **21**, 3001–3012 (2012).
152. Meng, L. et al. Truncation of Ube3a-ATS unsilences paternal Ube3a and ameliorates behavioral defects in the Angelman syndrome mouse model. *PLoS Genet.* **9**, e1004039 (2013).
153. Meng, L. et al. Towards a therapy for Angelman syndrome by targeting a long non-coding RNA. *Nature* **518**, 409–412 (2015).
154. Katayama, S. et al. Antisense transcription in the mammalian transcriptome. *Science* **309**, 1564–1566 (2005).
155. Lewis, A. et al. Epigenetic dynamics of the Kcnq1 imprinted domain in the early embryo. *Development* **133**, 4203–4210 (2006).
156. Lewis, A. et al. Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat. Genet.* **36**, 1291–1295 (2004).
157. Mohammad, F., Mondal, T., Guseva, N., Pandey, G. K. & Kanduri, C. Kcnq1ot1 noncoding RNA mediates transcriptional gene silencing by interacting with Dnmt1. *Development* **137**, 2493–2499 (2010).
158. Nagano, T. et al. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **322**, 1717–1720 (2008).
159. Pandey, R. R. et al. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell* **32**, 232–246 (2008).
160. Redrup, L. et al. The long noncoding RNA Kcnq1ot1 organises a lineage-specific nuclear domain for epigenetic gene silencing. *Development* **136**, 525–530 (2009).
161. Umlauf, D. et al. Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat. Genet.* **36**, 1296–1300 (2004).
162. Wagschal, A. et al. G9a histone methyltransferase contributes to imprinting in the mouse placenta. *Mol. Cell Biol.* **28**, 1104–1113 (2008).
163. Braceros, A. K. et al. Proximity-dependent recruitment of Polycomb repressive complexes by the lncRNA Airn. *Cell Rep.* **42**, 112803 (2023).
164. Long, Y. et al. RNA is essential for PRC2 chromatin occupancy and function in human pluripotent stem cells. *Nat. Genet.* **52**, 931–938 (2020).
165. Lleres, D. et al. CTCF modulates allele-specific sub-TAD organization and imprinted gene activity at the mouse Dlk1–Dio3 and Igf2–H19 domains. *Genome Biol.* **20**, 272 (2019).
166. Hansen, A. S. et al. Distinct classes of chromatin loops revealed by deletion of an RNA-binding region in CTCF. *Mol. Cell* **76**, 395–411.e13 (2019).
167. Saldana-Meyer, R. et al. RNA interactions are essential for CTCF-mediated genome organization. *Mol. Cell* **76**, 412–422.e5 (2019).
168. Kurukuti, S. et al. CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc. Natl Acad. Sci. USA* **103**, 10684–10689 (2006).
169. Murrell, A., Heeson, S. & Reik, W. Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nat. Genet.* **36**, 889–893 (2004).
170. Kopp, F. & Mendell, J. T. Functional classification and experimental dissection of long noncoding RNAs. *Cell* **172**, 393–407 (2018).
171. Rinn, J. L. et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311–1323 (2007).
172. Li, L. et al. Targeted disruption of Hotair leads to homeotic transformation and gene derepression. *Cell Rep.* **5**, 3–12 (2013).
173. Davidovich, C. et al. Toward a consensus on the binding specificity and promiscuity of PRC2 for RNA. *Mol. Cell* **57**, 552–558 (2015).
174. Schorderet, P. & Duboule, D. Structural and functional differences in the long non-coding RNA hotair in mouse and human. *PLoS Genet.* **7**, e1002071 (2011).
175. Selleri, L. et al. A hox-embedded long noncoding RNA: is it all hot air? *PLoS Genet.* **12**, e1006485 (2016).
176. Smith, K. P., Hall, L. L. & Lawrence, J. B. Nuclear hubs built on RNAs and clustered organization of the genome. *Curr. Opin. Cell Biol.* **64**, 67–76 (2020).
177. Hutchinson, J. N. et al. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8**, 39 (2007).
178. West, J. A. et al. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol. Cell* **55**, 791–802 (2014).
179. Ballarino, M. et al. Deficiency in the nuclear long noncoding RNA charme causes myogenic defects and heart remodeling in mice. *EMBO J.* **37**, e99697 (2018).
180. Desideri, F. et al. Intronic determinants coordinate charme lncRNA nuclear activity through the interaction with MATR3 and PTBP1. *Cell Rep.* **33**, 108548 (2020).
181. Taliani, V. et al. The long noncoding RNA Charme supervises cardiomyocyte maturation by controlling cell differentiation programs in the developing heart. *eLife* **12**, e81360 (2023).
182. Daneshvar, K. et al. lncRNA DIGIT and BRD3 protein form phase-separated condensates to regulate endoderm differentiation. *Nat. Cell Biol.* **22**, 1211–1222 (2020).
183. Chang, K. C. et al. MaTAR25 lncRNA regulates the Tensin1 gene to impact breast cancer progression. *Nat. Commun.* **11**, 6438 (2020).
184. Creamer, K. M., Kolpa, H. J. & Lawrence, J. B. Nascent RNA scaffolds contribute to chromosome territory architecture and counter chromatin compaction. *Mol. Cell* **81**, 3509–3525.e5 (2021).
185. Mele, M. & Rinn, J. L. "Cat's cradling" the 3D genome by the act of lncRNA transcription. *Mol. Cell* **62**, 657–664 (2016).
186. Andergassen, D. et al. In vivo firre and Dxz4 deletion elucidates roles for autosomal gene regulation. *eLife* **8**, e47214 (2019).
187. Hacisuleyman, E. et al. Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* **21**, 198–206 (2014).
188. Blank-Giwojna, A., Postepska-Igielska, A. & Grummt, I. lncRNA KHPS1 activates a poised enhancer by triplex-dependent recruitment of epigenomic regulators. *Cell Rep.* **26**, 2904–2915.e4 (2019).
189. Postepska-Igielska, A. et al. lncRNA Khps1 regulates expression of the proto-oncogene SPHK1 via triplex-mediated changes in chromatin structure. *Mol. Cell* **60**, 626–636 (2015).
190. O'Leary, V. B. et al. PARTICLE, a triplex-forming long ncRNA, regulates locus-specific methylation in response to low-dose irradiation. *Cell Rep.* **11**, 474–485 (2015).
191. Grote, P. & Herrmann, B. G. The long non-coding RNA Fendrr links epigenetic control mechanisms to gene regulatory networks in mammalian embryogenesis. *RNA Biol.* **10**, 1579–1585 (2013).
192. Kalwa, M. et al. The lncRNA HOTAIR impacts on mesenchymal stem cells via triple helix formation. *Nucleic Acids Res.* **44**, 10631–10643 (2016).
193. Leisegang, M. S. et al. HIF1α-AS1 is a DNA:DNA:RNA triplex-forming lncRNA interacting with the HUSH complex. *Nat. Commun.* **13**, 6563 (2022).
194. Trembinski, D. J. et al. Aging-regulated anti-apoptotic long non-coding RNA Sarrah augments recovery from acute myocardial infarction. *Nat. Commun.* **11**, 2039 (2020).
195. Zhang, X. et al. KCNQ1OT1 promotes genome-wide transposon repression by guiding RNA-DNA triplexes and HP1 binding. *Nat. Cell Biol.* **24**, 1617–1629 (2022).
196. Uroda, T. et al. Conserved pseudoknots in lncRNA MEG3 are essential for stimulation of the p53 pathway. *Mol. Cell* **75**, 982–995.e9 (2019).
197. Mondal, T. et al. MEG3 long noncoding RNA regulates the TGF-β pathway genes through formation of RNA–DNA triplex structures. *Nat. Commun.* **6**, 7743 (2015).
198. Ariel, F. et al. Noncoding transcription by alternative RNA polymerases dynamically regulates an auxin-driven chromatin loop. *Mol. Cell* **55**, 383–396 (2014).
199. Ariel, F. et al. R-loop mediated trans action of the APOLO long noncoding RNA. *Mol. Cell* **77**, 1055–1065.e4 (2020).
200. Szafranski, P. et al. Small noncoding differentially methylated copy-number variants, including lncRNA genes, cause a lethal lung developmental disorder. *Genome Res.* **23**, 23–33 (2013).
201. Sauvageau, M. et al. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *eLife* **2**, e01749 (2013).
202. van Dijk, M. et al. HELLP babies link a novel lincRNA to the trophoblast cell cycle. *J. Clin. Invest.* **122**, 4003–4011 (2012).
203. Kvon, E. Z., Waymack, R., Gad, M. & Wunderlich, Z. Enhancer redundancy in development and disease. *Nat. Rev. Genet.* **22**, 324–336 (2021).
204. Miguel-Escalada, I. et al. Pancreas agenesis mutations disrupt a lead enhancer controlling a developmental enhancer cluster. *Dev. Cell* **57**, 1922–1936.e9 (2022).
205. Osterwalder, M. et al. Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**, 239–243 (2018).
206. Chenier, S. et al. CHD2 haploinsufficiency is associated with developmental delay, intellectual disability, epilepsy and neurobehavioural problems. *J. Neurodev. Disord.* **6**, 9 (2014).
207. Cohen, A. S. A. et al. Haploinsufficiency of the basic helix–loop–helix transcription factor HAND2 causes congenital heart defects. *Am. J. Med. Genet. A* **182**, 1263–1267 (2020).
208. Dirkx, E. et al. Nfat and miR-25 cooperate to reactivate the transcription factor Hand2 in heart failure. *Nat. Cell Biol.* **15**, 1282–1293 (2013).
209. Tamura, M. et al. Overdosage of Hand2 causes limb and heart defects in the human chromosomal disorder partial trisomy distal 4q. *Hum. Mol. Genet.* **22**, 2471–2481 (2013).
210. Yamagata, K. et al. Mutations in the hepatocyte nuclear factor-1α gene in maturity-onset diabetes of the young (MODY3). *Nature* **384**, 455–458 (1996).
211. Luco, R. F. et al. A conditional model reveals that induction of hepatocyte nuclear factor-1α in Hnf1α-null mutant β-cells can activate silenced genes postnatally, whereas overexpression is deleterious. *Diabetes* **55**, 2202–2211 (2006).
212. Gage, P. J., Suh, H. & Camper, S. A. Dosage requirement of Pitx2 for development of multiple organs. *Development* **126**, 4643–4651 (1999).
213. Tumer, Z. & Bach-Holm, D. Axenfeld–Rieger syndrome and spectrum of PITX2 and FOXC1 mutations. *Eur. J. Hum. Genet.* **17**, 1527–1539 (2009).
214. Cole, M. D. The myc oncogene: its role in transformation and differentiation. *Annu. Rev. Genet.* **20**, 361–384 (1986).
215. George, M. R. et al. Minimal in vivo requirements for developmentally regulated cardiac long intergenic non-coding RNAs. *Development* **146**, dev185314 (2019).
216. Atla, G. et al. Genetic regulation of RNA splicing in human pancreatic islets. *Genome Biol.* **23**, 196 (2022).
217. Holdt, L. M. & Teupser, D. Long noncoding RNA ANRIL: lnc-ing genetic variation at the chromosome 9p21 locus to molecular mechanisms of atherosclerosis. *Front. Cardiovasc. Med.* **5**, 145 (2018).

218. de Goede, O. M. et al. Population-scale tissue transcriptomics maps long non-coding RNAs to complex disease. *Cell* **184**, 2633–2648.e19 (2021).
Identification of numerous lncRNAs as candidate mediators of genetic association signals that underly susceptibility for prevalent human diseases.
219. Cory, S., Graham, M., Webb, E., Corcoran, L. & Adams, J. M. Variant (6;15) translocations in murine plasmacytomas involve a chromosome 15 locus at least 72 kb from the c-myc oncogene. *EMBO J.* **4**, 675–681 (1985).
220. Graham, M., Adams, J. M. & Cory, S. Murine T lymphomas with retroviral inserts in the chromosomal 15 locus for plasmacytoma variant translocations. *Nature* **314**, 740–743 (1985).
221. Hu, X. et al. A functional genomic approach identifies FAL1 as an oncogenic long noncoding RNA that associates with BMI1 and represses p21 expression in cancer. *Cancer Cell* **26**, 344–357 (2014).
222. Leucci, E. et al. Melanoma addiction to the long non-coding RNA SAMMSON. *Nature* **531**, 518–522 (2016).
223. Hoadley, K. A. et al. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell* **173**, 291–304.e6 (2018).
224. Olivero, C. E. & Dimitrova, N. Identification and characterization of functional long noncoding RNAs in cancer. *FASEB J.* **34**, 15360–15646 (2020).
225. Hilton, L. K. et al. The double-hit signature identifies double-hit diffuse large B-cell lymphoma with genetic events cryptic to FISH. *Blood* **134**, 1528–1532 (2019).
226. Gutschner, T., Hammerle, M. & Diederichs, S. MALAT1 — a paradigm for long noncoding RNA function in cancer. *J. Mol. Med.* **91**, 791–801 (2013).
227. Gupta, R. A. et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071–1076 (2010).
228. Martinez-Terroba, E. et al. Overexpressed malat1 drives metastasis through inflammatory reprogramming of lung adenocarcinoma microenvironment. Preprint at *bioRxiv* https://doi.org/10.1101/2023.03.20.533534 (2023).
229. Hibi, K. et al. Loss of H19 imprinting in esophageal cancer. *Cancer Res.* **56**, 480–482 (1996).
230. Kondo, M. et al. Frequent loss of imprinting of the H19 gene is often associated with its overexpression in human lung cancers. *Oncogene* **10**, 1193–1198 (1995).
231. Rainier, S. et al. Relaxation of imprinted genes in human cancer. *Nature* **362**, 747–749 (1993).
232. Tseng, Y. Y. & Bagchi, A. The PVT1–MYC duet in cancer. *Mol. Cell Oncol.* **2**, e974467 (2015).
233. Cai, Z. et al. Targeting bim via a lncRNA morrbid regulates the survival of preleukemic and leukemic cells. *Cell Rep.* **31**, 107816 (2020).
234. Cai, Z. et al. Role of lncRNA Morrbid in PTPN11(Shp2)^{E76K}-driven juvenile myelomonocytic leukemia. *Blood Adv.* **4**, 3246–3251 (2020).
235. Huang, Y. et al. The role of lincRNA-p21 in regulating the biology of cancer cells. *Hum. Cell* **35**, 1640–1649 (2022).
236. Borensztein, M. et al. Xist-dependent imprinted X inactivation and the early developmental consequences of its failure. *Nat. Struct. Mol. Biol.* **24**, 226–233 (2017).
237. Marahrens, Y., Panning, B., Dausman, J., Strauss, W. & Jaenisch, R. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes. Dev.* **11**, 156–166 (1997).
238. Takagi, N. & Abe, K. Detrimental effects of two active X chromosomes on early mouse development. *Development* **109**, 189–201 (1990).
239. Yang, L., Yildirim, E., Kirby, J. E., Press, W. & Lee, J. T. Widespread organ tolerance to Xist loss and X reactivation except under chronic stress in the gut. *Proc. Natl Acad. Sci. USA* **117**, 4262–4272 (2020).
240. Yildirim, E. et al. Xist RNA is a potent suppressor of hematologic cancer in mice. *Cell* **152**, 727–742 (2013).
241. Syrett, C. M. et al. Loss of Xist RNA from the inactive X during B cell development is restored in a dynamic YY1-dependent two-step process in activated B cells. *PLoS Genet.* **13**, e1007050 (2017).
242. Spaziano, A. & Cantone, I. X-chromosome reactivation: a concise review. *Biochem. Soc. Trans.* **49**, 2797–2805 (2021).
243. Syrett, C. M. et al. Altered X-chromosome inactivation in T cells may promote sex-biased autoimmune diseases. *JCI Insight* **4**, e12671 (2019).
Findings in mouse and human suggesting that *XIST* dysregulation and abnormal X- chromosome inactivation in T cells underlie the increased prevalence of systemic lupus erythematosus in women.
244. Wang, J. et al. Unusual maintenance of X chromosome inactivation predisposes female lymphocytes for increased expression from the inactive X. *Proc. Natl Acad. Sci. USA* **113**, E2029–E2038 (2016).
245. Yu, B. et al. B cell-specific XIST complex enforces X-inactivation and restrains atypical B cells. *Cell* **184**, 1790–1803.e17 (2021).
246. Dou, D. R. et al. XIST ribonucleoproteins promote female sex-biased autoimmunity. Preprint at *bioRxiv* https://doi.org/10.1101/2022.11.05.515306 (2022).
247. Li, Y. et al. A noncoding RNA modulator potentiates phenylalanine metabolism in mice. *Science* **373**, 662–673 (2021).
248. Dindot, S. V. et al. An ASO therapy for Angelman syndrome that targets an evolutionarily conserved region at the start of the UBE3A-AS transcript. *Sci. Transl. Med.* **15**, eabf4077 (2023).
249. Wolter, J. M. et al. Cas9 gene therapy for Angelman syndrome traps Ube3a-ATS long non-coding RNA. *Nature* **587**, 281–284 (2020).
250. Jiang, J. et al. Translating dosage compensation to trisomy 21. *Nature* **500**, 296–300 (2013).
251. Abulwerdi, F. A. et al. Selective small-molecule targeting of a triple helix encoded by the long noncoding RNA, MALAT1. *ACS Chem. Biol.* **14**, 223–235 (2019).
252. Donlic, A., Zafferani, M., Padroni, G., Puri, M. & Hargrove, A. E. Regulation of MALAT1 triple helix stability and in vitro degradation by diphenylfurans. *Nucleic Acids Res.* **48**, 7653–7664 (2020).
253. Zafferani, M. et al. Multiassay profiling of a focused small molecule library reveals predictive bidirectional modulation of the lncRNA MALAT1 triplex stability in vitro. *ACS Chem. Biol.* **17**, 2437–2447 (2022).
254. Aguilar, R. et al. Targeting Xist with compounds that disrupt RNA structure and X inactivation. *Nature* **604**, 160–166 (2022).
255. Rosa, S., Duncan, S. & Dean, C. Mutually exclusive sense-antisense transcription at FLC facilitates environmentally induced gene repression. *Nat. Commun.* **7**, 13031 (2016).
256. Li, P., Tao, Z. & Dean, C. Phenotypic evolution through variation in splicing of the noncoding RNA COOLAIR. *Genes. Dev.* **29**, 696–701 (2015).
257. Wu, Z., Fang, X., Zhu, D. & Dean, C. Autonomous pathway: flowering locus C repression through an antisense-mediated chromatin-silencing mechanism. *Plant. Physiol.* **182**, 27–37 (2020).
258. Yang, M. et al. In vivo single-molecule analysis reveals COOLAIR RNA structural diversity. *Nature* **609**, 394–399 (2022).
Temperature-dependent structural alterations in the lncRNA *COOLAIR* underly its role as a transcription repressive switch during seasonal transition in plants.
259. Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N. J. & Dean, C. R-loop stabilization represses antisense transcription at the *Arabidopsis* FLC locus. *Science* **340**, 619–621 (2013).

Acknowledgements
The authors thank T. Graff, M. Cuenca-Ardura and B. Payer for critical reading of this manuscript. This work was supported by European Research Council (789055) and Spanish Ministry of Science and Innovation (PID2021-122522OB-I00) grants to J.F., and by National Institute of Health (R01CA262286 and R37CA230580) grants to N.D.

Author contributions
The authors contributed equally to all aspects of the article.

Competing interests
The authors declare no competing interests.

Additional information
Peer review information *Nature Reviews Molecular Cell Biology* thanks Takayuki Nojima, Igor Ulitsky and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.