

Review

Local Tandem Repeat Expansion in *Xist* RNA as a Model for the Functionalisation of ncRNA

Neil Brockdorff 

Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK; neil.brockdorff@bioch.ox.ac.uk

Received: 19 September 2018; Accepted: 16 October 2018; Published: 19 October 2018



Abstract: *Xist*, the master regulator of the X chromosome inactivation in mammals, is a 17 kb lncRNA that acts *in cis* to silence the majority of genes along the chromosome from which it is transcribed. The two key processes required for *Xist* RNA function, localisation *in cis* and recruitment of silencing factors, are genetically separable, at least in part. Recent studies have identified *Xist* RNA sequences and associated RNA-binding proteins (RBPs) that are important for these processes. Notably, several of the key *Xist* RNA elements correspond to local tandem repeats. In this review, I use examples to illustrate different modes whereby tandem repeat amplification has been exploited to allow orthodox RBPs to confer new functions for *Xist*-mediated chromosome inactivation. I further discuss the potential generality of tandem repeat expansion in the evolution of functional long non-coding RNAs (lncRNAs).

Keywords: X inactivation; *Xist*; lncRNA; tandem repeat

1. Introduction

The long non-coding RNA (lncRNA) *Xist* mediates X chromosome inactivation (X inactivation), the process that, in mammals, equalises levels of X-linked gene expression in XX female relative to XY male cells [1–5]. *Xist* RNA is expressed from the inactive chromosome-elect at the onset of X inactivation during early embryogenesis and then localises to a subnuclear domain corresponding to the X chromosome nuclear territory [4,6]. Localised *Xist* RNA induces chromosome inactivation by recruiting factors that modify underlying chromatin and repress gene activity [7]. The repressive chromatin state on the inactive X chromosome (Xi), once established, is maintained through subsequent cell divisions in development and adult life.

A deletion analysis using *Xist* transgenes suggested that the localisation of *Xist* RNA and *Xist*-mediated silencing are separable processes [8]. Further support for separation of gene-silencing and localisation of *Xist* RNA has come from recent studies identifying RNA binding proteins (RBPs) that interact with *Xist* RNA, and that function principally in *Xist* RNA localisation (hnRNPU and Ciz1) [9–11], or *Xist*-mediated gene silencing (Spen, Rbm15/m6A-methyltransferase complex, LBR, and hnRNPK) [12–17].

Further studies on *Xist* RBPs and the elements to which they bind have provided important advances in our understanding of *Xist* RNA localisation and *Xist*-mediated silencing. Several observations have led to the idea that *Xist* RNA is anchored to the Xi territory through the interaction with the nuclear matrix: Thus, *Xist* RNA domains are retained in cells following nuclear matrix extraction, a procedure which yields nuclei from which the bulk of chromatin and soluble proteins are removed [18]. Additionally, the knockdown/knockout of the *Xist* localization factors hnRNPU and Ciz1, both of which have been characterised as components of the insoluble nuclear matrix, leads to the delocalisation of *Xist* RNA to sites throughout the nucleoplasm [9–11]. Finally, analysis by 3D-SIM, a method for super-resolution light microscopy, has revealed that *Xist* RNA colocalises with hnRNPU and Ciz1 within channel networks that pervade interphase chromosome territories,

including on Xi [19,20]. Extrapolating from this idea, it can be inferred that anchored *Xist* RNA molecules function to nucleate chromosome silencing through the recruitment/enrichment of *Xist* silencing factors (Figure 1).

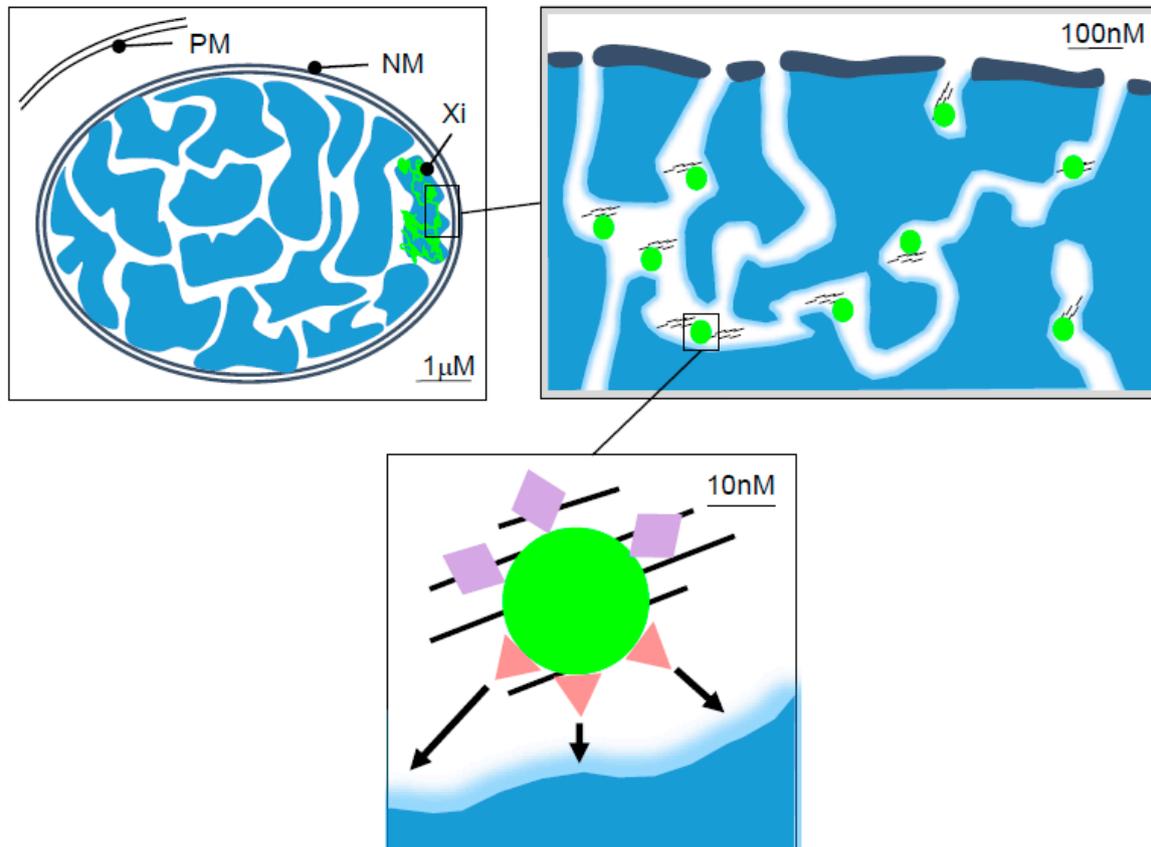


Figure 1. *Xist* RNA in the interphase nucleus. The schematic represents cross-sections of the nucleus illustrating the deduced relationship of *Xist* RNA (green) relative to the inactive X chromosome (Xi), interphase chromatin (blue), and interchromatin channels at different scales. The green circles represent single *Xist* RNA molecules. PM and NM denote the plasma- and nuclear-membrane respectively. The black lines represent the nuclear matrix proteins. Coloured shapes indicate the *Xist* RNA-associated proteins linked to tethering *Xist* RNA to the nuclear matrix (lilac diamond) or *Xist*-mediated chromatin silencing (red triangles). Arrows indicate the activity of the silencing factors towards proximal chromatin sites.

Super-resolution microscopy experiments have further revealed that *Xist* RNA domains are comprised of around 100–200 foci, representing individual *Xist* ribonucleoprotein particles (RNPs) [20,21]. An important challenge is to understand how this relatively small number of *Xist* RNPs induces gene silencing across an entire chromosome, comprising some 150 Mb of DNA and around 1000 genes. One factor that may contribute is signal amplification through local tandemly repeated RBP binding sites within single *Xist* RNA molecules. Specifically, the sequence analysis of *Xist* RNA in mouse, human and other species has revealed that a significant proportion of the primary RNA sequence is comprised of blocks of local tandemly repeated elements, designated repeats A–F (Figure 2). Several of these blocks are conserved in *Xist* from different mammalian species [2,4,22], and moreover, in many cases, they have been shown to bind RBPs that play a role in *Xist* RNA localisation and *Xist*-mediated silencing [10,11,13,15,17]. In this review, I will discuss the emerging evidence regarding how *Xist* tandem repeats contribute to function, and in addition, the wider implications for these findings in understanding the evolution of functional lncRNAs.

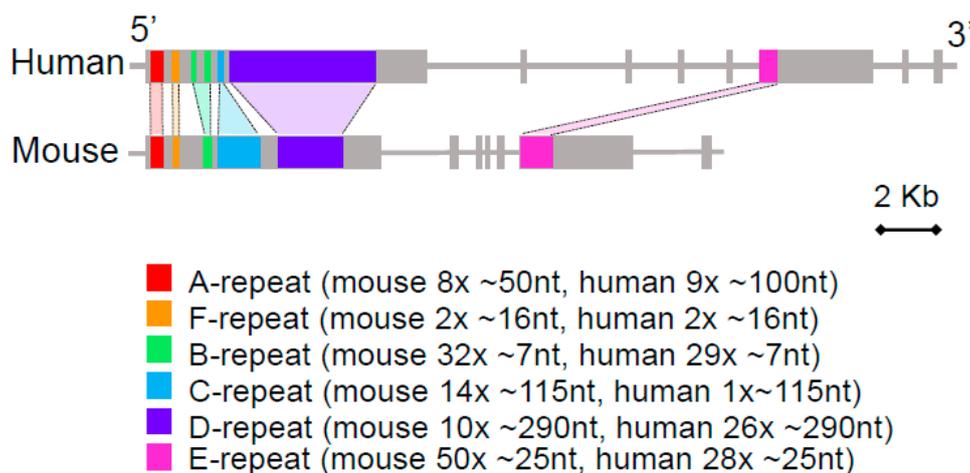


Figure 2. Local tandem repeats in *Xist* RNA. The schematic illustrates the intron/exon structure of human and mouse *Xist* genes with conserved tandem repeat blocks indicated in different colours. The key indicates the label, approximate copy number, and monomer length for each repeat block based on the mouse and human *Xist/XIST* RNA sequence.

2. Interaction of the RBPs *Spn* and *Rbm15* with the *Xist* A-Repeat Element

The *Xist* A-repeat, located at the 5' end of *Xist* RNA, comprises of 7–8 copies of a 24 nt GC-rich core sequence separated by uridine tracts of variable length [2,8]. The deletion of the A-repeat from full-length *Xist* RNA largely ablates *Xist*-mediated silencing, but does not grossly affect *Xist* RNA localisation [8]. However, in the context of a truncated *Xist* RNA transgene spanning the first 4 kb of sequence, deletion of the A-repeat does lead to loss of *Xist* RNA localisation [8,17], attesting to functional redundancy of localisation elements in full-length *Xist* RNA. Given the essential role of the A-repeats, there has been considerable interest in defining the RBPs that interact with this element. To this end recent proteomic and genetic screening strategies have identified that the RBP *Spn* is recruited to the A-repeat and functions as a critical factor in *Xist*-mediated silencing [12–15], recruiting the NCoR-HDAC3 complex to deacetylate chromatin at target sites [12]. A closely related RBP, *Rbm15*, that has been implicated in *Xist*-mediated silencing has also been shown to bind to the A-repeat region [14,23]. *Rbm15* interacts with the multiprotein enzyme complex that catalyses the N6-adenosine (m6A) methylation of RNA, and it is thought that *Rbm15* functions target m6A to sites across *Xist* RNA to facilitate the silencing function [14,23]. The precise mechanism of the *Rbm15*/m6A function in *Xist* silencing is not well defined but is thought to involve the m6A reader protein, YTHDC1 [23].

An interesting speculation is that the evolutionary expansion of A-repeat monomers has increased the number of *Spn* molecules that can concurrently bind to a single *Xist* RNA molecule, thereby enhancing the silencing activity. *Spn* is a large protein, some 3600 amino acids, which includes four closely spaced RNA binding (RRM) domains at the N-terminus, and a *Spn* Paralogue and Orthologue C-terminal (SPOC) domain located at the C-terminus [24–27]. The SPOC domain mediates the interaction with the NCoR-HDAC3 complex [26]. In vitro studies using a truncated N-terminal *Spn* protein spanning RRM 2–4 have determined that binding to the A-repeat monomer occurs with a 1:2 stoichiometry [28]. This conclusion is based on the observation that the overall A-repeat structure is comprised of pairs of monomers, and the finding that *Spn* RRM 2–4 binds to single-stranded regions, determined by in vitro CLiP (UV cross-link analysis). Thus, in relation to signal amplification, each *Xist* RNA molecule has the potential to bind 3–4 *Spn* molecules, a moderate enhancement. It is interesting to note that the complementation of the A-repeat deletion can be achieved using a synthetic X7.5 A-repeat consensus sequence, but not with an X5.5 A-repeat [8]. Whether or not the A-repeat is bound by multiple *Spn* molecules contemporaneously in vivo remains to be determined. Thus, the large size of the *Spn* protein may sterically hinder nearby binding sites. Additionally, it is likely

that Rbm15, which has a very similar configuration of N-terminal RRM domains to that of Spen [29], competes for A-repeat binding sites.

The interaction strength of Spen and Rbm15 with *Xist* RNA and with NCoR-HDAC3 or the m6A MTase complex, respectively, is also potentially important in relation to signal amplification. Thus, relatively weak and, hence, dynamic interactions could contribute to signal amplification by creating a high local concentration of effector proteins in interchromatin spaces (Figure 3). In the case of Spen, the in vitro analysis indicates that Spen RRMs bind A-repeat sequences with a high affinity, ~10 nM [28], suggesting a stable interaction. However, the interaction of Spen with NCoR, NCoR with HDAC3 could be relatively weak and dynamic, potentially underpinning the local concentration of the effectors to amplify function. A similar argument can be put forward for Rbm15 interactions with the A-repeat and m6A MTase complex respectively.

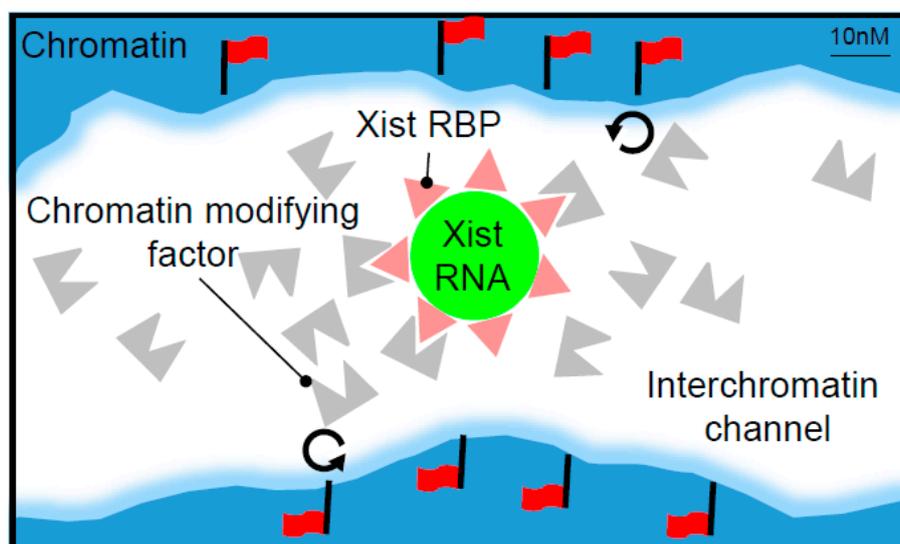


Figure 3. *Xist* ribonucleoprotein particles (RNPs) generate local concentrations of silencing factors within interchromatin channels. The schematic illustrating a model for how *Xist* RNA bound RNA binding proteins (RBPs) (red triangles) can function to generate a local concentration of chromatin-modifying factors (grey shape). Multiple RBP molecules that are strongly bound to a tandemly repeated element on the anchored *Xist* RNA molecule (green) interact weakly/transiently with the chromatin-modifying factor such that the local concentration of unbound molecules increases within the interchromatin channel. The unbound chromatin-modifying factor can then act at the available chromatin sites (circular arrows) indiscriminately within a zone surrounding the *Xist* RNP, resulting in a widespread deposition or removal of specific chromatin modifications (red flags).

3. Recruitment of the Polycomb System by hnRNPK Bound to the *Xist* B/C-Repeat

A recent study has shown that the *Xist* RNA B-repeat, together with a short part of the C-repeat (Figure 2), directs the recruitment of Polycomb repressive complexes (PRC) 1 and 2, which also contribute to *Xist*-mediated gene silencing [17]. Briefly, the B-repeat is comprised of around 32 copies of a cytidine-rich heptameric sequence element which is bound by the RBP hnRNPK. hnRNPK is a multifunctional RBP characterised by 3 KH domains that mediate RNA binding, and an unstructured domain that links the protein to diverse interaction partners [30]. The recently identified hnRNPK interacting proteins include Pcgf3 and Pcgf5, homologous Polycomb proteins that function as core subunits of the non-canonical Pcgf3/5-PRC1 complex, responsible for initiating Polycomb recruitment by *Xist* RNA [17,31]. Pcgf3/5-PRC1 catalyses mono-ubiquitylation of lysine 119 in histone H2A (H2AK119ub1), a repressive chromatin modification which, in addition to contributing directly to gene repression in X inactivation, initiates a positive feedback cascade, recruiting other non-canonical PRC1 complexes, and a second complex, PRC2, which catalyses histone H3 lysine 27 tri-methylation

(H3K27me3) [31]. The role of the C-repeat region in this pathway has not been clearly defined but it is also bound by hnRNPK, albeit at a low level [32].

As argued above for Spen, the recruitment of multiple hnRNPK molecules by B/C-repeat monomers may serve to enhance *Xist*-mediated chromosome silencing. Each KH domain in hnRNPK binds to separate 6–7 nt C-rich tracts in a cooperative manner [33], and it follows that up to 10 hnRNPK molecules could bind to the B-repeat element of a single *Xist* molecule. This has not yet been tested either in vitro or in vivo, and as argued above, factors such as steric hindrance or competition with other RBPs may influence the occupancy of the B-repeat. Similar to Spen, hnRNPK binds RNA with a high affinity, in the low nanomolar range [34], whereas the interaction of hnRNPK with Pcgf3/5 appears to be relatively weak, as judged from the low stoichiometry of hnRNPK in mass spectrometry analysis using Pcgf3/5 as bait (unpublished work). Thus, in vivo, hnRNPK stably bound to the *Xist* B-repeat may function to amplify the local concentration of the Pcgf3/5-PRC1 effector complex within interchromatin spaces of Xi, with consequential widespread modification of the underlying chromatin, as depicted in Figure 3. Consistent with this idea, Polycomb mediated H3K27me3 is evenly distributed over Xi chromatin [19,35–37], contrasting with conventional Polycomb sites where the deposition is limited to relatively discrete elements at the promoters of target genes [38].

A further contribution to signal amplification in this example comes from the positive feedback between different Polycomb complexes and their respective histone modifications, as detailed above. Overall, the successive amplification steps in the B-repeat-mediated recruitment of the Polycomb system provide a compelling rationale for how relatively small numbers of *Xist* RNA molecules establish chromatin modification and gene silencing across an entire chromosome.

4. Ciz1, an Anchoring Factor, Is Recruited by the *Xist* E-Repeat

As noted above, the *Xist* RBP Ciz1 facilitates the anchoring of *Xist* RNA to the Xi territory [10,11]. This function is only evident in specific somatic cell types, fibroblasts, and B- and T-lymphocytes, where Ciz1 loss of function causes the dispersal of *Xist* RNA particles throughout the nucleoplasm [10]. As Ciz1 colocalises with *Xist* RNA in all cell types, the cell type specificity of the phenotype likely reflects redundancy in the mechanism for *Xist* RNA localisation. A known alternative pathway involves hnRNPU, the depletion of which results in *Xist* RNA dispersal in many cell types, including embryonic stem cells [9,12,13,39,40]. There is evidence that hnRNPU, which has an RRM domain, interacts directly with *Xist* RNA, although an analysis by CLiP indicates that the binding sites are broadly distributed throughout the transcripts, not centred on any specific unique or local tandem repeat element [32].

Ciz1 was originally identified as a protein that interacts with the cell cycle inhibitor CDKN1A and was subsequently shown to have a role in initiating DNA replication, and in cell cycle progression at the G1/S checkpoint [41]. Biochemical fractionation experiments found that Ciz1 is enriched in the nuclear matrix compartment [42]. A role as an RBP was suggested due to the presence of putative RNA binding zinc-finger domains, including a matrin-like zinc-finger. Evidence for an interaction with a known RNA, specifically *Xist*, came only recently [10,11]. Ciz1 binding was mapped to the *Xist* E-repeat, comprising of an approximately 20–25 nt sequence element tandemly repeated to approximately 50 copies (Figure 2). Whilst the sequence requirements and stoichiometry of Ciz1 binding to the E-repeat remain to be determined, the strong enrichment of Ciz1 in association with *Xist* RNPs compared to elsewhere in the nucleus, as determined by super-resolution microscopy [10], indicates that there are indeed multiple Ciz1 molecules bound to each *Xist* RNA molecule. The precise relationship of Ciz1 and hnRNPU in *Xist* RNA localisation is not clear, but the hnRNPU knockout in fibroblasts does result in the dispersal of *Xist* particles that continue to be associated with Ciz1 [11]. This latter observation suggests that hnRNPU is required for Ciz1 to confer *Xist* RNA anchoring, either due to a direct interaction or indirectly, for example playing a role in the formation of the nuclear matrix, the substrate for anchoring. The latter idea is supported by recent evidence that hnRNPU regulates interphase chromosome structure via oligomerization with chromatin-associated RNAs [43].

It is interesting to consider that the association of multiple Ciz1 molecules with the E-repeat underpins its role in anchoring *Xist* RNA particles. Specifically, if we assume that Ciz1 is relatively immobile as a consequence of its interaction with the nuclear matrix, but interacts with typical target RNAs that have a single or a few Ciz1 binding sites dynamically/transiently, then an RNA that has evolved multiple recognition motifs would be predicted to have a significantly increased dwell/retention time, as illustrated in Figure 4. This idea is conceptually distinct from *Xist* RNA repeats serving to increase the local concentration of the effector complexes proposed above in relation to the recruitment of Spn/Rbm15 and hnRNPk (Figure 3).

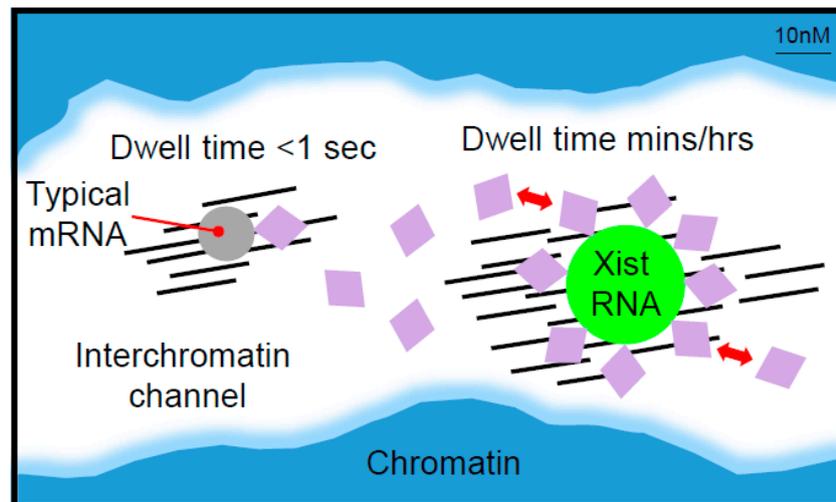


Figure 4. The amplification of RBP binding sites as a driver for *Xist* RNP anchoring. A schematic illustrating a model for how the amplification of RBP-binding sites on *Xist* RNA facilitates RNA anchoring. For a typical messenger RNA (mRNA) (grey circle), an RBP (lilac diamond) that interacts transiently with nuclear matrix proteins (black lines) immobilises the mRNA in interchromatin channels for a short time (hypothetical dwell time $< 1 \text{ s}$). For *Xist* RNA (green circle), amplification of the number of binding sites for the RBP increases the dwell time hypothetically up to minutes or even hours in a manner proportional to the local concentration of the RBP, the interaction strength with *Xist* RNA, and the number of binding sites.

The *Xist* A-repeat and C-repeat have also been suggested to have a role in *Xist* RNA localisation [8,44,45] and, although the mechanistic basis for this is uncertain, it could also involve amplified binding sites for RBPs that enhance *Xist* RNA retention in the nuclear matrix fraction. This speculation is consistent with a redundancy in *Xist* RNA localisation pathways [8].

5. Local Tandem Repeat Amplification and the Evolution of *Xist* RNA Function

The accumulated evidence that local tandem repeats in *Xist* RNA are central to its role in X inactivation suggests a simple model for how *Xist* RNA evolved. Thus, tandem duplication of sequences that encompass the binding site for a common RBP in the archetypal *Xist* RNA, likely as a consequence of DNA replication errors, could have generated enhanced or modified RBP functions on which natural selection could act on. The theory of sex chromosome evolution implies that the dosage compensation evolved incrementally, with allelic repression of X-linked genes being selected initially within a relatively small region of the prototypic X chromosome (corresponding to the region in which the prototypic Y homologues become functionally compromised due to recombination suppression) [46]. With the progressive erosion of the prototypic Y chromosome, there would be a selective pressure for the X inactivation signal to spread further, eventually encompassing the entire X chromosome. Thus, the amplification of binding sites for RBPs that interact transiently with the nuclear matrix would be predicted to incrementally increase the dwell time of prototypic *Xist* RNPs,

with the increased number of bound RBP molecules contributing to the strength of the interaction with the nuclear matrix. Evolution of strong interactions with the nuclear matrix would first serve to limit *Xist* RNP localisation *in cis*, and second, would be important in terms of modulating the range or distance of spread of *Xist* RNPs.

Using a similar argument, the amplification of binding sites for RBPs that generate a local concentration of silencing factors could be selected for on the basis of improved efficiency/completeness of gene dosage compensation. One possibility is that an RBP that binds to a single site in the prototypic *Xist* lncRNA may have a previously evolved interaction with a known transcriptional repressor. An example of this would be the RBP Spen, for which interaction with NCoR-HDAC3 is predicted to be evolutionarily conserved [26]. Alternatively, an RBP that binds to a unique site in the prototypic *Xist* lncRNA may acquire an interaction with a known repressor complex as a neofunctionality. This could be the case for hnRNPK which has roles in diverse aspects of RNA biogenesis [30], of which the recruitment of Pcgf3/5-PRC1 complexes to chromatin has only been documented to occur in the context of *Xist* RNA [17]. Thus, the amplification of hnRNPK-binding sites in *Xist* RNA may have been exploited to enhance a relatively weak interaction with the Pcgf3/5 Polycomb protein to a level sufficient to initiate the Polycomb cascade.

In the context of the above models, the long-term retention of *Xist* RNPs on the nuclear matrix is likely to be important in enabling the local concentration of silencing factors to reach a critical threshold. Accordingly, Spen/hnRNPK binding sites in other messenger RNAs (mRNAs) or lncRNAs may not lead to significant levels of chromatin modification *in cis*, i.e., in cases where the RNA is not anchored to the nuclear matrix for a significant time. Here it is interesting to note that hnRNPK binding to C-rich motifs has been implicated in retention of nuclear RNAs, including those arising from the Alu family of dispersed repeats [47].

6. A Role for Local Tandem Repeat Expansion in Other LncRNAs

Our growing appreciation of the importance of local tandem repeat expansion in functionalisation of *Xist* RNA leads to the question of how general this model may be in the evolution of functional lncRNA. There are indeed examples that illustrate that this mechanism may be more widely utilised. Thus, the lncRNA RNA on the Silent X (*Rsx*), which evolved as the master regulator of X inactivation in marsupial mammals entirely independently of *Xist*, is also characterised by the presence of large blocks of local tandem repeated sequences [48]. The sequence similarity of these elements and those present in *Xist* RNA is limited, and at present nothing is known about their functional importance or bound RBPs. However, close parallels in the X inactivation process in eutherian and metatherian mammals, notably chromosome-wide hypoacetylation of histones and recruitment of the Polycomb system, imply that commonalities could extend to key *Xist* silencing factors such as Spen and hnRNPK. There may also be overlaps in terms of the factors and mechanisms regulating *Rsx* RNA localisation, as, like *Xist* RNA, *Rsx* RNA localises strictly *in cis* on the X chromosome from which it is transcribed [48].

Whilst the focus of this review up to this point has been on local repeat sequences that are tandemly arranged, it should be noted that many of the arguments put forward apply equally well to local repeats that are dispersed within a given locus, referred to simply as local repeats. An example of this has emerged from recent studies on the lncRNA Firre/FIRRE [49,50]. Thus, a 156 bp local repeat termed RRD, present in 8 copies, is required for the nuclear localisation of FIRRE RNA, potentially through providing binding sites for the nuclear matrix protein hnRNPU. RRD is also present in mouse Firre RNA (16 copies). Although primary sequence conservation relative to primate RRD is only moderate (~60%), interaction with hnRNPU is conserved and knockdown experiments indicate hnRNPU-dependent nuclear retention both in human and mouse cells [50]. Another well-documented example of an lncRNA in which local repeat sequences play a functional role is NORAD, which regulates the genome stability by sequestering PUMILIO proteins that control mRNA stability [51,52].

Recent studies on the lncRNA Neat1, which is required to nucleate the formation of paraspeckles, a phase-separated membrane-less organelle [53], also point to the importance of the multivalent

interaction of RBPs in lncRNA function. Specifically, the recruitment of proteins required to establish paraspeckles is mediated by several redundant modules present in the middle part of Neat1 lncRNA [54]. In this example, there is no available evidence for a locally repeated sequence common to independent modules. Possible explanations for this are that dissimilar sequences recruit the key RBPs required for paraspeckle assembly, or alternatively, that each independent module has a conserved secondary RNA structure that forms from apparently disparate primary sequences.

Whilst the amplification of RBP binding sites likely represents a key mechanism for the functionalisation of lncRNAs, it should be noted that the amplification of binding sites for other RNAs may also be an important evolutionary mechanism. Indeed, certain circular RNAs have evolved as molecular sponges, presenting multiple binding sites for specific miRNAs [55].

7. Summary

The expansion of local tandem repeats during the evolution of *Xist* RNA provides an instructive example of how lncRNAs may become functionalised through natural selection. Conversely, the analysis of local tandem repeat expansions, or for that matter dispersed local repeats, in lncRNA, together with the identification of RBPs that bind to them, could provide a useful approach towards understanding the function of specific lncRNAs. An interesting starting point would be to analyse lncRNAs associated with imprinted gene clusters as, at least in some cases, these likely function analogously to *Xist* to induce chromatin repression over contiguous genomic regions [56,57]. Similarly, lncRNAs transcribed from chromosomal regions comprising tandem DNA repeats, for example at telomeres [58], pericentric heterochromatin [59], and other repetitive regions [60], are interesting candidates. Determining which lncRNAs to focus on in terms of the potential for functionalisation through repeat expansion will be greatly facilitated by SEEKR, a recently described bioinformatics tool that analyses the k-mer (short motif) content within defined lncRNA sequences [61].

Funding: Wellcome Trust, 103768.

Acknowledgments: I would like to thank Tatyana Nesterova, Guifeng Wei and Heather Coker for suggestions and critical reading of this review. Neil Brockdorff is funded by a grant from Wellcome (103768).

Conflicts of Interest: The author declares no conflict of interest.

References

1. Lyon, M.F. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* **1961**, *190*, 372–373. [[CrossRef](#)] [[PubMed](#)]
2. Brockdorff, N.; Ashworth, A.; Kay, G.F.; McCabe, V.M.; Norris, D.P.; Cooper, P.J.; Swift, S.; Rastan, S. The product of the mouse *Xist* gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **1992**, *71*, 515–526. [[CrossRef](#)]
3. Penny, G.D.; Kay, G.F.; Sheardown, S.A.; Rastan, S.; Brockdorff, N. Requirement for *Xist* in X chromosome inactivation. *Nature* **1996**, *379*, 131–137. [[CrossRef](#)] [[PubMed](#)]
4. Brown, C.J.; Hendrich, B.D.; Rupert, J.L.; Lafreniere, R.G.; Xing, Y.; Lawrence, J.; Willard, H.F. The human *XIST* gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **1992**, *71*, 527–542. [[CrossRef](#)]
5. Lee, J.T.; Jaenisch, R. Long-range cis effects of ectopic X-inactivation centres on a mouse autosome. *Nature* **1997**, *386*, 275–279. [[CrossRef](#)] [[PubMed](#)]
6. Duthie, S.M.; Nesterova, T.B.; Formstone, E.J.; Keohane, A.M.; Turner, B.M.; Zakian, S.M.; Brockdorff, N. *Xist* RNA exhibits a banded localization on the inactive X chromosome and is excluded from autosomal material *in cis*. *Hum. Mol. Genet.* **1999**, *8*, 195–204. [[CrossRef](#)] [[PubMed](#)]
7. Cerase, A.; Pintacuda, G.; Tattermusch, A.; Avner, P. *Xist* localization and function: New insights from multiple levels. *Genome Biol.* **2015**, *16*, 166. [[CrossRef](#)] [[PubMed](#)]
8. Wutz, A.; Rasmussen, T.P.; Jaenisch, R. Chromosomal silencing and localization are mediated by different domains of *Xist* RNA. *Nat. Genet.* **2002**, *30*, 167–174. [[CrossRef](#)] [[PubMed](#)]

9. Hasegawa, Y.; Brockdorff, N.; Kawano, S.; Tsutui, K.; Tsutui, K.; Nakagawa, S. The matrix protein hnRNP U is required for chromosomal localization of *Xist* RNA. *Dev. Cell* **2010**, *19*, 469–476. [[CrossRef](#)] [[PubMed](#)]
10. Ridings-Figueroa, R.; Stewart, E.R.; Nesterova, T.B.; Coker, H.; Pintacuda, G.; Godwin, J.; Wilson, R.; Haslam, A.; Lilley, F.; Ruigrok, R.; et al. The nuclear matrix protein CIZ1 facilitates localization of *Xist* RNA to the inactive X-chromosome territory. *Genes Dev.* **2017**, *31*, 876–888. [[CrossRef](#)] [[PubMed](#)]
11. 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* **2017**, *114*, 10654–10659. [[CrossRef](#)] [[PubMed](#)]
12. McHugh, C.A.; Chen, C.K.; Chow, A.; Surka, C.F.; Tran, C.; McDonel, P.; Pandya-Jones, A.; Blanco, M.; Burghard, C.; Moradian, A.; et al. The *Xist* lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **2015**, *521*, 232–236. [[CrossRef](#)] [[PubMed](#)]
13. Chu, C.; Zhang, Q.C.; da Rocha, S.T.; Flynn, R.A.; Bharadwaj, M.; Calabrese, J.M.; Magnuson, T.; Heard, E.; Chang, H.Y. Systematic discovery of *Xist* RNA binding proteins. *Cell* **2015**, *161*, 404–416. [[CrossRef](#)] [[PubMed](#)]
14. Moindrot, B.; Cerase, A.; Coker, H.; Masui, O.; Grijzenhout, A.; Pintacuda, G.; Schermelleh, L.; Nesterova, T.B.; Brockdorff, N. A pooled shRNA screen identifies Rbm15, Spen, and Wtap as factors required for *Xist* RNA-mediated silencing. *Cell Rep.* **2015**, *12*, 562–572. [[CrossRef](#)] [[PubMed](#)]
15. Monfort, A.; Di Minin, G.; Postlmayr, A.; Freimann, R.; Arieti, F.; Thore, S.; Wutz, A. Identification of Spen as a crucial factor for *Xist* function through forward genetic screening in haploid embryonic stem cells. *Cell Rep.* **2015**, *12*, 554–561. [[CrossRef](#)] [[PubMed](#)]
16. Minajigi, A.; Froberg, J.E.; Wei, C.; Sunwoo, H.; Kesner, B.; Colognori, D.; Lessing, D.; Payer, B.; Boukhali, M.; Haas, W.; et al. Chromosomes. A comprehensive *Xist* interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* **2015**, *349*. [[CrossRef](#)] [[PubMed](#)]
17. Pintacuda, G.; Wei, G.; Roustan, C.; Kirmizitas, B.A.; Solcan, N.; Cerase, A.; Castello, A.; Mohammed, S.; Moindrot, B.; Nesterova, T.B.; et al. hnRNPK recruits PCGF3/5-PRC1 to the *Xist* RNA B-repeat to establish polycomb-mediated chromosomal silencing. *Mol. Cell* **2017**, *68*, 955–969. [[CrossRef](#)] [[PubMed](#)]
18. Clemson, C.M.; McNeil, J.A.; Willard, H.F.; Lawrence, J.B. *XIST* RNA paints the inactive X chromosome at interphase: Evidence for a novel RNA involved in nuclear/chromosome structure. *J. Cell Biol.* **1996**, *132*, 259–275. [[CrossRef](#)] [[PubMed](#)]
19. Cerase, A.; Smeets, D.; Tang, Y.A.; Gdula, M.; Kraus, F.; Spivakov, M.; Moindrot, B.; Leleu, M.; Tattermusch, A.; Demmerle, J.; et al. Spatial separation of *Xist* RNA and polycomb proteins revealed by superresolution microscopy. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 2235–2240. [[CrossRef](#)] [[PubMed](#)]
20. Smeets, D.; Markaki, Y.; Schmid, V.J.; Kraus, F.; Tattermusch, A.; Cerase, A.; Sterr, M.; Fiedler, S.; Demmerle, J.; Popken, J.; et al. Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct *Xist* RNA foci. *Epigenetics Chromatin* **2014**, *7*, 8. [[CrossRef](#)] [[PubMed](#)]
21. Sunwoo, H.; Wu, J.Y.; Lee, J.T. The *Xist* RNA-PRC2 complex at 20-nm resolution reveals a low *Xist* stoichiometry and suggests a hit-and-run mechanism in mouse cells. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E4216–E4225. [[CrossRef](#)] [[PubMed](#)]
22. Nesterova, T.B.; Slobodyanyuk, S.Y.; Elisaphenko, E.A.; Shevchenko, A.I.; Johnston, C.; Pavlova, M.E.; Rogozin, I.B.; Kolesnikov, N.N.; Brockdorff, N.; Zakian, S.M. Characterization of the genomic *Xist* locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence. *Genome Res.* **2001**, *11*, 833–849. [[CrossRef](#)] [[PubMed](#)]
23. Patil, D.P.; Chen, C.K.; Pickering, B.F.; Chow, A.; Jackson, C.; Guttman, M.; Jaffrey, S.R. m⁶A RNA methylation promotes *XIST*-mediated transcriptional repression. *Nature* **2016**, *537*, 369–373. [[CrossRef](#)] [[PubMed](#)]
24. Oswald, F.; Kostezka, U.; Astrahantseff, K.; Bourteele, S.; Dillinger, K.; Zechner, U.; Ludwig, L.; Wilda, M.; Hameister, H.; Knochel, W.; et al. SHARP is a novel component of the Notch/RBP-J κ signalling pathway. *EMBO J.* **2002**, *21*, 5417–5426. [[CrossRef](#)] [[PubMed](#)]
25. Kuang, B.; Wu, S.C.; Shin, Y.; Luo, L.; Kolodziej, P. Split ends encodes large nuclear proteins that regulate neuronal cell fate and axon extension in the *Drosophila* embryo. *Development* **2000**, *127*, 1517–1529. [[PubMed](#)]
26. Ariyoshi, M.; Schwabe, J.W. A conserved structural motif reveals the essential transcriptional repression function of Spen proteins and their role in developmental signaling. *Genes Dev.* **2003**, *17*, 1909–1920. [[CrossRef](#)] [[PubMed](#)]

27. Shi, Y.; Downes, M.; Xie, W.; Kao, H.Y.; Ordentlich, P.; Tsai, C.C.; Hon, M.; Evans, R.M. Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. *Genes Dev.* **2001**, *15*, 1140–1151. [[CrossRef](#)] [[PubMed](#)]
28. Lu, Z.; Zhang, Q.C.; Lee, B.; Flynn, R.A.; Smith, M.A.; Robinson, J.T.; Davidovich, C.; Gooding, A.R.; Goodrich, K.J.; Mattick, J.S.; et al. RNA duplex map in living cells reveals higher order transcriptome structure. *Cell* **2016**, *165*, 1267–1279. [[CrossRef](#)] [[PubMed](#)]
29. Mercher, T.; Coniat, M.B.; Monni, R.; Mauchauffe, M.; Nguyen Khac, F.; Gressin, L.; Mugneret, F.; Leblanc, T.; Dastugue, N.; Berger, R.; et al. Involvement of a human gene related to the *Drosophila spen* gene in the recurrent t(1;22) translocation of acute megakaryocytic leukemia. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 5776–5779. [[CrossRef](#)] [[PubMed](#)]
30. Bomsztyk, K.; Denisenko, O.; Ostrowski, J. hnRNP K: One protein multiple processes. *Bioessays* **2004**, *26*, 629–638. [[CrossRef](#)] [[PubMed](#)]
31. Almeida, M.; Pintacuda, G.; Masui, O.; Koseki, Y.; Gdula, M.; Cerase, A.; Brown, D.; Mould, A.; Innocent, C.; Nakayama, M.; et al. PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. *Science* **2017**, *356*, 1081–1084. [[CrossRef](#)] [[PubMed](#)]
32. Cirillo, D.; Blanco, M.; Armaos, A.; Bunes, A.; Avner, P.; Guttman, M.; Cerase, A.; Tartaglia, G.G. Quantitative predictions of protein interactions with long noncoding RNAs. *Nat. Methods* **2016**, *14*, 5–6. [[CrossRef](#)] [[PubMed](#)]
33. Paziewska, A.; Wyrwicz, L.S.; Bujnicki, J.M.; Bomsztyk, K.; Ostrowski, J. Cooperative binding of the hnRNP K three KH domains to mRNA targets. *FEBS Lett.* **2004**, *577*, 134–140. [[CrossRef](#)] [[PubMed](#)]
34. Moritz, B.; Lilie, H.; Naarmann-de Vries, I.S.; Urlaub, H.; Wahle, E.; Ostareck-Lederer, A.; Ostareck, D.H. Biophysical and biochemical analysis of hnRNP K: Arginine methylation, reversible aggregation and combinatorial binding to nucleic acids. *Biol. Chem.* **2014**, *395*, 837–853. [[CrossRef](#)] [[PubMed](#)]
35. Marks, H.; Chow, J.C.; Denissov, S.; Francoijs, K.J.; Brockdorff, N.; Heard, E.; Stunnenberg, H.G. High-resolution analysis of epigenetic changes associated with X inactivation. *Genome Res.* **2009**, *19*, 1361–1373. [[CrossRef](#)] [[PubMed](#)]
36. Calabrese, J.M.; Sun, W.; Song, L.; Mugford, J.W.; Williams, L.; Yee, D.; Starmer, J.; Mieczkowski, P.; Crawford, G.E.; Magnuson, T. Site-specific silencing of regulatory elements as a mechanism of X inactivation. *Cell* **2012**, *151*, 951–963. [[CrossRef](#)] [[PubMed](#)]
37. Pinter, S.F.; Sadreyev, R.I.; Yildirim, E.; Jeon, Y.; Ohsumi, T.K.; Borowsky, M.; Lee, J.T. Spreading of X chromosome inactivation via a hierarchy of defined Polycomb stations. *Genome Res.* **2012**, *22*, 1864–1876. [[CrossRef](#)] [[PubMed](#)]
38. Mendenhall, E.M.; Koche, R.P.; Truong, T.; Zhou, V.W.; Issac, B.; Chi, A.S.; Ku, M.; Bernstein, B.E. GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS Genet.* **2010**, *6*, e1001244. [[CrossRef](#)] [[PubMed](#)]
39. Sakaguchi, T.; Hasegawa, Y.; Brockdorff, N.; Tsutsui, K.; Tsutsui, K.M.; Sado, T.; Nakagawa, S. Control of chromosomal localization of *Xist* by hnRNP U family molecules. *Dev. Cell* **2016**, *39*, 11–12. [[CrossRef](#)] [[PubMed](#)]
40. Kolpa, H.J.; Fackelmayer, F.O.; Lawrence, J.B. SAF-A Requirement in anchoring *XIST* RNA to chromatin varies in transformed and primary cells. *Dev. Cell* **2016**, *39*, 9–10. [[CrossRef](#)] [[PubMed](#)]
41. Pauzaite, T.; Thacker, U.; Tollitt, J.; Copeland, N.A. Emerging roles for Ciz1 in cell cycle regulation and as a driver of tumorigenesis. *Biomolecules* **2016**, *7*, 1. [[CrossRef](#)] [[PubMed](#)]
42. Ainscough, J.F.; Rahman, F.A.; Sercombe, H.; Sedo, A.; Gerlach, B.; Coverley, D. C-terminal domains deliver the DNA replication factor Ciz1 to the nuclear matrix. *J. Cell Sci.* **2007**, *120*, 115–124. [[CrossRef](#)] [[PubMed](#)]
43. Nozawa, R.S.; Boteva, L.; Soares, D.C.; Naughton, C.; Dun, A.R.; Buckle, A.; Ramsahoye, B.; Bruton, P.C.; Saleeb, R.S.; Arnedo, M.; et al. SAF-A regulates interphase chromosome structure through oligomerization with chromatin-associated RNAs. *Cell* **2017**, *169*, 1214–1227.e18. [[CrossRef](#)] [[PubMed](#)]
44. Beletskii, A.; Hong, Y.K.; Pehrson, J.; Egholm, M.; Strauss, W.M. PNA interference mapping demonstrates functional domains in the noncoding RNA *Xist*. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9215–9220. [[CrossRef](#)] [[PubMed](#)]
45. Sarma, K.; Levasseur, P.; Aristarkhov, A.; Lee, J.T. Locked nucleic acids (LNAs) reveal sequence requirements and kinetics of *Xist* RNA localization to the X chromosome. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 22196–22201. [[CrossRef](#)] [[PubMed](#)]

46. Graves, J.A. Evolution of vertebrate sex chromosomes and dosage compensation. *Nat. Rev. Genet.* **2016**, *17*, 33–46. [[CrossRef](#)] [[PubMed](#)]
47. Lubelsky, Y.; Ulitsky, I. Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature* **2018**, *555*, 107–111. [[CrossRef](#)] [[PubMed](#)]
48. Grant, J.; Mahadevaiah, S.K.; Khil, P.; Sangrithi, M.N.; Royo, H.; Duckworth, J.; McCarrey, J.R.; VandeBerg, J.L.; Renfree, M.B.; Taylor, W.; et al. *Rsx* is a metatherian RNA with *Xist*-like properties in X-chromosome inactivation. *Nature* **2012**, *487*, 254–258. [[CrossRef](#)] [[PubMed](#)]
49. Hacısuleyman, E.; Shukla, C.J.; Weiner, C.L.; Rinn, J.L. Function and evolution of local repeats in the *Firre* locus. *Nat. Commun.* **2016**, *7*, 11021. [[CrossRef](#)] [[PubMed](#)]
50. Hacısuleyman, E.; Goff, L.A.; Trapnell, C.; Williams, A.; Henao-Mejia, J.; Sun, L.; McClanahan, P.; Hendrickson, D.G.; Sauvageau, M.; Kelley, D.R.; et al. Topological organization of multichromosomal regions by the long intergenic noncoding RNA *Firre*. *Nat. Struct. Mol. Biol.* **2014**, *21*, 198–206. [[CrossRef](#)] [[PubMed](#)]
51. Lee, S.; Kopp, F.; Chang, T.C.; Sataluri, A.; Chen, B.; Sivakumar, S.; Yu, H.; Xie, Y.; Mendell, J.T. Noncoding RNA NORAD regulates genomic stability by sequestering PUMILIO proteins. *Cell* **2016**, *164*, 69–80. [[CrossRef](#)] [[PubMed](#)]
52. Tichon, A.; Gil, N.; Lubelsky, Y.; Havkin Solomon, T.; Lemze, D.; Itzkovitz, S.; Stern-Ginossar, N.; Ulitsky, I. A conserved abundant cytoplasmic long noncoding RNA modulates repression by Pumilio proteins in human cells. *Nat. Commun.* **2016**, *7*, 12209. [[CrossRef](#)] [[PubMed](#)]
53. Clemson, C.M.; Hutchinson, J.N.; Sara, S.A.; Ensminger, A.W.; Fox, A.H.; Chess, A.; Lawrence, J.B. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* **2009**, *33*, 717–726. [[CrossRef](#)] [[PubMed](#)]
54. Yamazaki, T.; Souquere, S.; Chujo, T.; Kobelke, S.; Chong, Y.S.; Fox, A.H.; Bond, C.S.; Nakagawa, S.; Pierron, G.; Hirose, T. Functional domains of NEAT1 architectural lncRNA induce paraspeckle assembly through phase separation. *Mol. Cell* **2018**, *70*, 1038–1053. [[CrossRef](#)] [[PubMed](#)]
55. Memczak, S.; Jens, M.; Elefsinioti, A.; Torti, F.; Krueger, J.; Rybak, A.; Maier, L.; Mackowiak, S.D.; Gregersen, L.H.; Munschauer, M.; et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **2013**, *495*, 333–338. [[CrossRef](#)] [[PubMed](#)]
56. Sleutels, F.; Zwart, R.; Barlow, D.P. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **2002**, *415*, 810–813. [[CrossRef](#)] [[PubMed](#)]
57. Redrup, L.; Branco, M.R.; Perdeaux, E.R.; Krueger, C.; Lewis, A.; Santos, F.; Nagano, T.; Cobb, B.S.; Fraser, P.; Reik, W. The long noncoding RNA *Kcnq1ot1* organises a lineage-specific nuclear domain for epigenetic gene silencing. *Development* **2009**, *136*, 525–530. [[CrossRef](#)] [[PubMed](#)]
58. Azzalin, C.M.; Reichenbach, P.; Khorიაuli, L.; Giulotto, E.; Lingner, J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **2007**, *318*, 798–801. [[CrossRef](#)] [[PubMed](#)]
59. Martens, J.H.; O’Sullivan, R.J.; Braunschweig, U.; Opravil, S.; Radolf, M.; Steinlein, P.; Jenuwein, T. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.* **2005**, *24*, 800–812. [[CrossRef](#)] [[PubMed](#)]
60. Delilhas, N. Formation of a family of long intergenic noncoding RNA genes with an embedded translocation breakpoint motif in human chromosomal low copy repeats of 22q11.2—Some surprises and questions. *Non-Coding RNA* **2018**, *4*, 16. [[CrossRef](#)] [[PubMed](#)]
61. Kirk, J.M.; Kim, S.O.; Inoue, K.; Smola, M.J.; Lee, D.M.; Schertzer, M.D.; Wooten, J.S.; Baker, A.R.; Sprague, D.; Collins, D.W.; et al. Functional classification of long non-coding RNAs by k-mer content. *Nat. Genet.* **2018**, *50*, 1474–1482. [[CrossRef](#)] [[PubMed](#)]

