Evolution of Genome-Organizing Long Non-coding RNAs in Metazoans

América Ramírez-Colmenero¹, Katarzyna Oktaba² and Selene L. Fernandez-Valverde*¹

¹Unidad de Genómica Avanzada (Langebio), Centro de Investigación y de Estudios Avanzados del IPN, Irapuato, México,
²Unidad Irapuato, Centro de Investigación y de Estudios Avanzados del IPN, Irapuato, México

Long non-coding RNAs (lncRNAs) have important regulatory functions across eukarya. It is now clear that many of these functions are related to gene expression regulation through their capacity to recruit epigenetic modifiers and establish chromatin interactions. Several lncRNAs have been recently shown to participate in modulating chromatin within the spatial organization of the genome in the three-dimensional space of the nucleus. The identification of lncRNA candidates is challenging, as it is their functional characterization. Conservation signatures of lncRNAs are different from those of protein-coding genes, making identifying lncRNAs under selection a difficult task, and the homology between lncRNAs may not be readily apparent. Here, we review the evidence for these higher-order genome organization functions of lncRNAs in animals and the evolutionary signatures they display.

Keywords: evolution, conservation, long-non-coding RNAs, chromatin conformation, three-dimensional chromatin conformation, genome topology, gene expression regulation

INTRODUCTION

The three-dimensional (3D) organization of DNA in the cell nucleus has become a significant subject of study, particularly its influence on gene regulation. Recent advances in chromatin conformation capture (3C) techniques, computational, and modeling approaches have made its study feasible on a genome-wide scale, giving insight into the structure and the dynamics of chromatin folding in space and time. Nuclear 3D organization has multiple levels and varies between cell types and biological conditions. For instance, chromosomes are subdivided into topologically associating domains (TADs) within which chromatin loops bring together regulatory elements and target loci separated in the linear genome (Dixon et al., 2012). These chromatin interactions are crucial for precise gene expression regulation (reviewed in Furlong and Levine, 2018; Schoenfelder and Fraser, 2019; Ibrahim and Mundlos, 2020). Importantly, changes in transcriptional programs result in variation in chromatin interactions within TADs, while TAD boundaries delimiting these domains are preserved (Dixon et al., 2015). TADs segregate in the nuclear space into transcriptionally active (A) and inactive (B) compartments. A/B compartments correlate well with histone modifications characteristic of euchromatin and heterochromatin, respectively, and are described as cell type-specific, being able to undergo switches during cell differentiation and lineage commitment (Lieberman-Aiden et al., 2009; Rao et al., 2014; Dixon et al., 2015; Fortin and Hansen, 2015).
In addition to DNA and histones, RNA is a major component of the cell nucleus (Rinn and Chang, 2012). High-throughput sequencing methods have revealed the pervasive transcription of thousands of non-coding RNA (ncRNA) molecules in the genome. Among the latter, long non-coding RNAs (lncRNAs) have emerged as important gene regulators in eukaryotes. lncRNAs are broadly defined as transcripts longer than 200 nucleotides, with little to no protein-coding potential (Mercer et al., 2009; Wang and Chang, 2011; Derrien et al., 2012). lncRNAs are more lowly expressed (Hezroni et al., 2015), display more tissue-restricted expression patterns (Necsulea et al., 2014), have fewer exons, and are shorter than protein-coding genes (Hezroni et al., 2015). In animals, several lncRNAs are essential to phenomena such as gene silencing, activation, and chromatin remodeling, with significant roles in development, immunity, and cancer (Guttman et al., 2011; Schmitt and Chang, 2016; Delás et al., 2017). lncRNA functions may predate the origin of metazoans, as several unicellular holozans possess lncRNAs that are distinct in terms of their histone marks as well as expression throughout their life cycle (Gaiti et al., 2017).

SIGNATURES OF CONSERVATION IN LNCRNAS

There has been a long debate on whether most lncRNAs are functional or not (van Bakel et al., 2010; Clark et al., 2011; Lindsay et al., 2013). This discussion was, in part, sparked by the fact that the sequence of lncRNAs is generally poorly conserved across species, suggesting that they are not under purifying selection (Babak et al., 2005; Ponjavic et al., 2007; Marques and Ponting, 2009). There are several examples of orthologous RNAs that preserve their function, but whose sequence is so divergent, they can no longer be identified as orthologs by sequence similarity alone (Ponjavic et al., 2007; Ulitsky et al., 2011; Ulitsky, 2016). Thus, the detection of conservation beyond sequence is paramount to annotate candidate lncRNAs for further functional characterization.

The conservation signals in lncRNAs can differ from those typically found in protein-coding genes (Diederichs, 2014; Ulitsky, 2016). For instance, conventional conservation analyses applied to coding sequences, such as calculating the rate between synonymous and non-synonymous mutations, are not suitable for these elements. Nevertheless, lncRNAs display some sequence conservation, generally in short sequence islands, potentially due to selection constraints on sequences necessary for interacting with other transcripts, proteins, or DNA (Kapusta and Feschotte, 2014; Quinn et al., 2016; Ulitsky, 2016). lncRNAs may also display constraints on the post-transcriptional processing of the transcript, leading to the conservation of splice sites across different species (Nitsche et al., 2015; Ulitsky, 2016). lncRNAs can also possess structural conservation – a constraint that may not be readily detectable at the sequence level (Smith et al., 2013; Tavares et al., 2019). Finally, lncRNAs can have positional conservation, and be expressed from syntenic loci despite having lost most or all sequence conservation. These modes of conservation are not mutually exclusive and may be present in a single lncRNA.

Beyond their apparent lack of conservation, many functionally characterized lncRNAs modulate the organization of higher-order chromatin structures in the nucleus (Saxena and Carninci, 2011; Marchese and Huarte, 2014). lncRNAs are involved in the formation of DNA loops and domains (Wang and Chang, 2011; Zhang et al., 2014), interchromosomal structures (Hacisuleyman et al., 2016), heterochromatic regions (Deng et al., 2009; Engreitz et al., 2013), subnuclear bodies (Mao et al., 2011), and the dynamic assembly of protein complexes (Tsai et al., 2010; Lin et al., 2014; Marin-Béjar et al., 2017). Several novel experimental methods allow the identification of lncRNAs binding to chromatin in vivo across the genome (Li et al., 2017; Sridhar et al., 2017; Bell et al., 2018; Bonetti et al., 2020; Gavrilov et al., 2020). Recruiting and binding to effector molecules is a prevalent mode of action of lncRNAs in both cis and trans activities.

Here, we summarize lncRNAs that affect, establish, or maintain three-dimensional chromatin organization in metazoans and the conservation signals that indicate they are under selection.

LNCRNAS THAT AFFECT TAD CONFORMATION AND THEIR CONSERVATION

Sequence Conservation

Sequence conservation in lncRNAs can range from very high to almost non-existent. Despite being generally presented as poorly conserved, a subset of lncRNAs can present significant sequence conservation across species (Necsulea et al., 2014; Hezroni et al., 2015). However, sequence conservation does not guarantee functional equivalence; a highly conserved lncRNA can be fundamental in one species while dispensable in others. For example, the lncRNA Metastasis Associated in Lung Adenocarcinoma Transcript 1 (MALAT1) is highly conserved from human to zebrafish (Figure 1A; Hutchinson et al., 2007; Lin et al., 2007). While the human MALAT1 functions in nuclear speckles, regulating alternative splicing (Hutchinson et al., 2007; Tripathi et al., 2010), cell-cycle associated genes (Yang et al., 2011), and cancer progression (Gutschner et al., 2013), the murine ortholog is neither essential for these functions nor mouse development (Eißmann et al., 2012; Nakagawa et al., 2012; Zhang et al., 2012). However, it is more common for lncRNAs to have short conserved motifs or domains that are important for their association with DNA or proteins that regulate chromatin conformation. For example, lncRNAs that affect 3D genome topology and arise from highly conserved syntenic loci, such as the Hox clusters, display contrasting patterns of sequence conservation compared to their protein counterparts in the same cluster. Hox genes, organized in mammals in four clusters (HoxA–HoxD), encode transcription factors crucial for patterning along the anterior-posterior axis. Numerous ncRNAs are
transcribed from the human HOX loci, and their expression relates to differential histone marks and transcriptional accessibility (Rinn et al., 2007).

The HOX antisense intergenic RNA (HOTAIR) lncRNA is transcribed from the boundary between domains with differential chromatin marks at the HOXc locus but acts in trans repressing...
transcription of coding and non-coding genes on the HOXD locus (Rinn et al., 2007). A chromatin loop established between HOTAIR locus and the HOXC distal enhancer (HDE) located downstream of HOTAIR promotes transcription of the lncRNA. This loop is disrupted by the recruitment of hepatocyte nuclear factor 4-α (HNF4α), a master regulator of epithelial differentiation, to the HDE (Battistelli et al., 2019). HOTAIR exists across mammals, albeit poorly conserved in sequence; it is only highly conserved in primates (He et al., 2011). Noteworthy, a highly conserved domain in exon 6, possibly the backbone of HOTAIR, appeared first in kangaroos suggesting the ab initio generation of HOTAIR in marsupials (He et al., 2011). Despite its low sequence conservation across mammals, key secondary structural elements of HOTAIR contain protein-binding motifs and have significant conservation or covariation (He et al., 2011; Somarowthu et al., 2015). However, studies evaluating the functional conservation of murine HOTAIR (mHotaír) present contradictory results. On the one hand, the deletion of the HoxC cluster, including mHotaír, did not affect HoxD silencing in vivo (Schorderet and Duboule, 2011). In contrast, mice homozygous for mHotaír KO presented homeostatic spine transformation and malformation of metacarpal bones, and derived fibroblasts showed altered expression and levels of epigenetic marks at hundreds of genes, including HoxD genes (Li et al., 2013). Interestingly, human and mouse HOTAIR differ in number, arrangement, and degree of sequence conservation among their exons. The absence of exons with protein-binding motifs in mHotaír may partially explain differences in their function.

Another lncRNA expressed from HOX clusters is HOXA transcript at the distal tip (HOTTIP), transcribed from the 5’ end of the HOXA locus in mammals and conserved in avians (Wang et al., 2011). Chromosomal looping brings HOTTIP into spatial proximity to its target genes in cis, allowing HOTTIP to activate transcription by binding the WD repeat domain 5/mixed lineage leukemia (WDR5/MLL) complex, driving H3K4me3 (Wang et al., 2011). HOTTIP and its association with CCCTC-binding factor (CTCF), which delineates active and inactive TADs within the HOXA cluster, also influence the expression of HoxA genes (Narendra et al., 2015; Wang et al., 2018).

Long non-coding RNAs also enable the establishment of inter-chromosomal structures. The Functional intergenic repeating RNA element (Firre) is an lncRNA involved in pluripotency, hematopoiesis, and adipogenesis (Hacisuleyman et al., 2014; Lewadowski et al., 2019). Firre accumulates across a ~5 Mb domain around its transcription site on the X chromosome (Hacisuleyman et al., 2014), located between two TADs, and highly enriched in CTCF binding sites, required for Firre transcription (Barutcu et al., 2018). This domain colocalizes with five regions on different chromosomes that contain genes with roles in adipogenesis. The formation of this structure depends on the interaction of Firre with Heterogeneous Nuclear Ribonucleoprotein U (HNRNPU), through a 156-bp repeating RNA domain (RRD; Hacisuleyman et al., 2014). This RRD is unique to Firre, and functions as a lineage-specific nuclear retention signal in mice and humans. The RRD and other local repeats (LRs) are conserved to different extents across Firre orthologs in mammals. Firre is also required for the super-loop formation of the inactive X chromosome (Xi), H3K27me3 deposition, and the localization of the Xi to the perinuclear region (Yang et al., 2015; Barutcu et al., 2018).

The 3D architecture of TADs enables a group of multi-exonic lncRNAs, termed immune gene-priming lncRNAs (IPLs), to direct the active priming of the promoters of immune genes, necessary for a rapid and robust pro-inflammatory response as part of trained immunity (Fanucchi et al., 2019). Upon induction of transcription of immune genes by the tumor necrosis factor (TNF), chromatin contacts increase TNF-induced genes and the lncRNAs loci. IPLs are somewhat conserved between mouse and human; the majority possess an Alu element in their first intron and share putative transcription-factor binding motifs at their promoters.

The region comprising an IPL, Upstream master lncRNA of the inflammatory chemokine locus (UMLILO), engages in chromosomal contacts with CXCL chemokine genes belonging to the same TAD, but UMLILO does not have enhancer-RNA-like characteristics. In contrast to other IPLs, UMLILO is not conserved in mice and only partially conserved in pigs, suggesting that IPLs are not essential across species, but have a complementary role in ensuring robust gene expression. UMLILO has short conserved sequence motifs and interacts with WDR5 through its conserved exon 3, directing WDR5/MLL1 to chemokine gene promoters, mediating H3K4me3. Transcription of chemokines in UMLILO knocked-down cells was restored by insertion of another WDR5-binding lncRNA, HOTTIP, under the control of the UMLILO promoter (Fanucchi et al., 2019). The ability of HOTTIP to rescue the loss of UMLILO is an example of convergent functional evolution, as they share minimal sequence similarity.

Another group of chromatin-modifying lncRNAs arises from the syntenic estrogen receptor 1 (ESR1) locus. ESR1 is strongly upregulated in cancerous cells undergoing estrogen deprivation. A cluster of ncRNAs, ESR1 locus enhancing and activating non-coding RNAs (Eleanors), are transcribed from introns in a large chromatin cluster within a TAD that contains the ESR1 locus (Tomita et al., 2015). These Eleanors form a chromatin-associated RNA cloud that delineates the TAD and cis-activate transcription. This TAD interacts with another active TAD that contains the apoptotic transcription factor forkhead Box O3 (FOXO3; Abdalla et al., 2019). Knockdown of a promoter-associated Eleanor, pa-Eleanor(S), induced repression of the rest of the Eleanors and the genes within the TAD, including ESR1 (Abdalla et al., 2019). The abundant and highly conserved Eleanor2 increases chromatin accessibility in the ESR1 upstream region by destabilizing nucleosomes, activating ESR1, and is required for the formation of the RNA cloud (Fujita et al., 2020).

Positional Conservation
Long non-coding RNAs may be expressed from syntenic loci, suggesting a common origin, but may have lost the majority of sequence conservation (Figure 1B). The functions of these lncRNAs are thought to rely primarily on their transcription...
(Diederichs, 2014; Ulitsky, 2016). Thus, the evolutionary signature would be expected to reside outside the transcribed region (Ulitsky, 2016). Indeed, many IncRNAs have a very conserved promoter but little to no conservation in their transcribed region (Guttman et al., 2009). A substantial difficulty in this classification is defining when sequence conservation is entirely lost. As outlined above, several IncRNAs only retain small patches of conservation considered negligible by some authors and meaningful by others.

Examples of this conundrum are dosage compensation IncRNAs in Drosophila melanogaster (Figure 1B). Detailed syntenic analysis of Drosophilid genomes revealed 47 new orthologs, where only 19 had been identified by sequence similarity (Quinn et al., 2016). Importantly, it was shown that the roX RNA itself, only its transcription, is necessary for dosage compensation (Quinn et al., 2016). Furthermore, a distant roX RNA ortholog rescues the loss of roX between two distant species (D. melanogaster and Drosophila busckii) despite almost no sequence conservation outside an eight nucleotide-long conserved patch of microhomology (Quinn et al., 2016).

A more traditional example of positional conservation is the IncRNA antisense to Igf2r RNA non-coding (Airn), required for paternal-specific silencing of imprinted genes in the insulin-like growth factor 2 (Igf2r) cluster (Sleutels et al., 2002). The function of Airn is conserved between human and mouse despite them sharing little conserved sequence (Yotova et al., 2008). The Igf2r silencing function of Airn was shown to be dependent on transcriptional overlap and not on the transcribed RNAs themselves (Latos et al., 2012). However, recent evidence shows that this is only the case for nearby imprinted genes, as the murine Airn IncRNA itself is necessary for the recruitment of chromatin-modifying complexes to distant non-overlapping genes in the cluster (Andergassen et al., 2019).

**Structural Conservation**

Structural conservation is potentially the most telling signal of conservation in IncRNAs, yet the most difficult to identify. The basic premise is that structural domains may be preserved despite changes in the sequence, as long as complementary base pairs are maintained.

The non-coding isoform of the steroid receptor RNA activator (SRA), ncSRA, has a four-domain secondary structure with varying levels of sequence conservation (Figure 1C). ncSRA functions as a coactivator of several human hormone receptors by modifying chromatin structure (Novikova et al., 2012). ncSRA associates with CTCF and the DEAD-BOX helicase 5 (DDX5), and this association is necessary for the insulator activity of CTCF in vivo (Yao et al., 2010). The functional RNA structure is conserved in all mammals, while its sequence is not. Furthermore, several of the varying positions in other species show changes predicted to help stabilize its structural elements (Novikova et al., 2012).

Dosage compensation IncRNAs (see next section) show patches of structural conservation of biological importance. The Repeat A (RepA) region of X-inactive specific transcript (Xist), essential to the establishment of X chromosome inactivation, interacts with proteins such as the polycomb repressive complex 2 (PRC2; Zhao et al., 2008), ATRX chromatin remodeler (Sarma et al., 2014), and SHARP repressor protein (McHugh et al., 2015). RepA was experimentally shown to have a complex structure that is preserved despite rapid changes across mammalian evolution, strongly suggesting that this structure is indispensable for Xist function (Liu et al., 2017). IncRNAs involved in dosage compensation in drosophilids, roX1 and roX2, have conserved boxes that correspond precisely with stems that are necessary for binding to the male-specific lethal (MSL) proteins. Domains outside these interaction zones are not conserved and lack structure (Ilik et al., 2013; Quinn et al., 2016).

**HOTAIR** has also been shown to have a complex secondary structure, with some evidence of conservation in mammals acquired from computational methods (Somarowthu et al., 2015). However, there is some debate as to whether there is enough evidence to suggest that HOTAIR's structure is conserved in mammals (Rivas et al., 2017). Similarly, secondary-structure predictions on Firre indicated that the RRD is a highly structured domain (Nakagawa and Hirano, 2014), consistent with LRs representing potential binding platforms for the specific targeting of proteins to specific genomic regions by IncRNAs.

**Functional Convergence: The Case of Dosage Compensation IncRNAs**

The IncRNAs involved in the process of dosage compensation are extraordinary examples of de novo emergence of novel IncRNAs of unrelated evolutionary origins (Figure 1D). A prominent example is the Xist IncRNA, required for dosage compensation in the sex-chromosomes of eutherians (Penny et al., 1996). Random X-chromosome inactivation in females is necessary to balance the transcriptional output to that of males. Xist localizes at the X inactivation center (XIC) and is expressed exclusively from the inactivated X (Xi; Brown et al., 1991). During the onset of X inactivation, Xist accumulates at the XIC (Clemson et al., 1996), and then targets gene-rich regions that are spatially close to its transcription site (Engreitz et al., 2013; Simon et al., 2013), incorporating them into the Xist silencing domain and spreading further to cover the complete future Xi (Engreitz et al., 2013). Xist-mediated inactivation involves the transcriptional silencing of most genes on the Xi, and its compaction and recruitment to the nuclear lamina (Zhao et al., 2008; Hasegawa et al., 2010; Chu et al., 2015; McHugh et al., 2015; Minajigi et al., 2015).

While exonic sequences of Xist are well-conserved among eutherians, there are differences in the exon-intron structure, length, and sequence between species (Nesterova et al., 2001; Elisaphenko et al., 2008). This indicates that either Xist genes present a high adaptation level or that their sequence and structure are not essential (Elisaphenko et al., 2008). Xist is not present in non-eutherian vertebrates, including marsupials, despite common epigenetic features on the Xi, such as loss of active histone marks and exclusion of RNA polymerase II (Chu et al., 2011). Homology of Xist with promoters and exonic sequences of the protein-coding gene ligand of
numb-protein x 3 (Lnx3) found in marsupials, chicken, and fish suggests that Xist emerged through pseudogenization of Lnx3, possibly by the insertion of tandem repeats from transposable elements (Duret et al., 2006; Elisaphenko et al., 2008).

Interestingly, in marsupials, X-chromosome inactivation is imprinted, tissue-specific, and somewhat incomplete compared to eutherians, and thought to be achieved by female-specific expression of the lncRNA RNA on the silent X (RxS), which is transcribed from and coats the paternal chromosome (Grant et al., 2012). The independent evolution of Xist and Rsx adds to the notion of dosage systems rapidly evolving from ancient silencing mechanisms common to all eukaryotes through the use of lncRNAs (Gendrel and Heard, 2014; Graves, 2016). The discoveries on the regulation of Xist by non-coding elements located at its own and the neighboring TAD and the impact of this 3D conformation on the regulatory landscape adds another layer of complexity to the mechanisms for dosage compensation (van Bemmelen et al., 2019; Galupa et al., 2020).

lncRNAs are also the effectors of dosage compensation in drosophilids, but they differ in both origin and mechanism to those in mammals. Here, the roX1 and roX2 lncRNAs mediate the upregulation of genes on the single male X chromosome to equalize expression of the two X chromosomes in females. roX1 and roX2 associate to the MSL proteins, forming the MSL complex that localizes to numerous specific sites along the male X (Franke and Baker, 1999), mediating histone acetylation and increasing transcription. The MSL complex does not alter the global architecture of the X chromosome, but it does spread via spatial proximity from high-affinity sites – enriched at TAD boundaries – to other regions (Ramírez et al., 2015). Contrary to Xist, whose activity is limited to the chromosome from which it is expressed (Wutz and Jaenisch, 2000), roX transgenes target the X chromosome in trans and rescue roX1 and roX2 mutant males (Meller and Rattner, 2002).

The independent origin of Xist in mammals, Rsx in marsupials, and roX1 and roX2 in flies suggests that lncRNAs may be one of the fastest mechanisms to evolve novel epigenetic controls. As these lncRNAs participate in dosage compensation but have emerged independently in several lineages, they are extraordinarily difficult to identify as functionally convergent. Additional examples of functionally equivalent lncRNAs with no evolutionary relationship may likely have gone undetected.

**DISCUSSION**

Distinctly, lncRNAs have emerged as an additional layer of complexity involved in shaping the three-dimensional organization of the genome by interacting and modifying the structure of chromatin. Several lncRNAs affect chromatin conformation and display a combination of conservation signals that may be difficult to identify solely by looking at traditional genomic conservation metrics (summarized in Table 1). These signatures could prove useful to identify and prioritize lncRNA

| TABLE 1 | Characterized long non-coding RNAs (lncRNAs) that are involved in nuclear genome topology. |
|---------|------------------------------------------------------------------------------------------|
| IncRNA  | Function | Mode of action | Interacting proteins | Association with chromatin topology | Conservation | References |
| Xist    | X chromosome inactivation in mammals | In cis | PRC1, PRC2, HNRNPU, RBM15, SHARP, WTAP, HNRNPK, LBR, and many others | The organization of the XIC into two topologically associating domains (TADs) ensures the proper interaction of Xist and its antisense lncRNA 70-kb with regulatory elements. During X inactivation, Xist spreads along the chromosome exploiting the three-dimensional (3D) organization, resulting in compaction and recruitment to the nuclear lamina. | Present only in eutherian mammals. Presence of common core exonic sequences, despite species-specific unique sequences, and variation in length and gene structure. | Nesterova et al., 2001; Plath et al., 2003; Elisaphenko et al., 2008, Zhao et al., 2008; Hasegawa et al., 2010; Engreitz et al., 2013; Chu et al., 2015; McHugh et al., 2015; Minaigri et al., 2016; Moinard et al., 2015; Chen et al., 2016; Pintacuda et al., 2017; van Bemmelen et al., 2019; Galupa et al., 2020 |
| HOTTIP  | Gene control of HOXA locus for distal identity | In cis | WDR5/MLL and CTCF | A chromatin loop gets HOTTIP into spatial proximity to HOXA genes. Associates with CTCF to define functional TADs at HOXA cluster | Portions conserved in mammals and avians | Wang et al., 2011, 2018 |
| Aim     | Its transcription prevents overexpression of Igfr2 locus in a paternal-specific manner | In cis | EHMT2 | Forms an RNA cloud, creating a repressive domain | Tandem direct repeats at the CpG island at 5' end are conserved in human and mouse at an organizational level but not by sequence | Lyle et al., 2000; Seidl et al., 2006; Nagano et al., 2008; Lutos et al., 2009, 2012; Koerner et al., 2012; Santoro et al., 2013 |

(Continued)
TABLE 1 | Continued

| IncRNA   | Function                                      | Mode of action | Interacting proteins          | Association with chromatin topology                                                                 | Conservation                                      | References                                      |
|----------|-----------------------------------------------|----------------|-------------------------------|------------------------------------------------------------------------------------------------------|--------------------------------------------------|------------------------------------------------|
| Kcnq1ot1 | Silencing at imprinted Kcnq1 locus in a paternal-specific manner | In cis         | EHMT2, PRC2, PRC1, and DNMT1  | Formation of repressive chromatin loop on the imprinting control region of the locus                    | Well-conserved motifs between human and mouse     | Pandey et al., 2004, 2008; Mancini-Dinardo et al., 2006; Mohammad et al., 2008, 2010; Zhang et al., 2014 Mayer et al., 2006, 2008; Santoro et al., 2010; Schmitz et al., 2010; Guetg et al., 2012; Jacob et al., 2013; Savić et al., 2014; Wehner et al., 2014 |
| pRNA     | Mediates silencing by CpG methylation of rRNA genes at nucleolus via DNA:RNA triplex formation | In cis         | NoRC and DNMT3b              | Establishment of nucleolar heterochromatin                                                             | Conserved across eutherians, various levels of sequence conservation, and highly conserved secondary-structure motifs | Pandey et al., 2004, 2008; Mancini-Dinardo et al., 2006; Mohammad et al., 2008, 2010; Zhang et al., 2014 Mayer et al., 2006, 2008; Santoro et al., 2010; Schmitz et al., 2010; Guetg et al., 2012; Jacob et al., 2013; Savić et al., 2014; Wehner et al., 2014 |
| LUNAR1   | IGF1 signaling, promotes cell proliferation in cancers | In cis         | None reported                | A chromatin loop that brings into contact the promoter of LUNAR1 and the enhancer of IGF1R is necessary for the expression of both genes, which reside in the same TAD Kbps1 transcription leads to a transcriptionally active open chromatin state by recruitment of EP300/CBP, transcripton of Kbps1 enhancer and eviction of CTCF | Not reported                                      | Trimparchi et al., 2014; Peng and Feng, 2016 |
| Kbps1    | Activates transcription of its sense proto-oncogene SPHK1, via DNA:RNA triplex formation at SPHK1 enhancer | In cis         | EP300                        | In chemokine TAD, chromosomal looping brings the super-enhancer region harboring UMLILO into contact with chemokine genes, allowing UMLILO RNA to guide WDR5/MLL to the promoters to facilitate H3K4me3 epigenetic priming Eleanors RNA cloud delineate ESR1 TAD and activate transcription | Conservation between humans and rodents          | Imamura et al., 2004; Postepksa-Igielska et al., 2015; Blank-Giwojna et al., 2019 |
| UMLILO   | Trained immune response on chemokine genes    | In cis         | WDR5/MLL                     | Diverse, chromatin looping and modification as scaffold for proteins in both active and inactive domains | Partial conservation between human, chimpanzee, and pig, absent in mouse | Fanucchi et al., 2019 |
| Eleanor  | Activation of ESR1 locus, apoptosis resistance | In cis         | None reported                | Eleanors RNA cloud delineate ESR1 TAD and activate transcription                                      | Varying levels of conservation for each Eleanor   | Tomita et al., 2015; Abdalla et al., 2019; Fujita et al., 2020; Rinn et al., 2007; Gupta et al., 2010; Tsai et al., 2010; Li et al., 2013; Somarowethu et al., 2015; Zhang et al., 2015; Portoso et al., 2017; Lanz et al., 1999; Shi et al., 2001; Zhao et al., 2007; Yao et al., 2010; Novikova et al., 2012; Wongtraoongate et al., 2015; Franke and Baker, 1999; Park et al., 2008; Ilk et al., 2013; Maenner et al., 2013; Ramirez et al., 2015; Quinn et al., 2016 |
| HOTAIR   | Represses expression in HOXD locus and other genes, including imprinted | In trans       | PRC2, RCOR1, and AR          | HOTAIR transcripts demarcate silent and active domains in HOXD locus.                                  | Poorly conserved by sequence, secondary structure motifs conserved between mouse and human            | Tomita et al., 2015; Abdalla et al., 2019; Fujita et al., 2020 |
| ncSRA    | Activation of steroid receptors (isoforms of SRA code for protein) | In trans       | SRC-1, PRC2, TrxG, NANOG, CTCF, SHARP, DDX5, and others | Diverse, chromatin looping and modification as scaffold for proteins in both active and inactive domains | Significant sequence conservation and high structural conservation | Tomita et al., 2015; Abdalla et al., 2019; Fujita et al., 2020 |
| roX1 and roX2 | Dosage compensation in Drosophila | In trans       | MSL Proteins                | The MSL complex (rox + MSL proteins) has high affinity sites on TAD borders                            | There are roX orthologs across drosophilids       | Tomita et al., 2015; Abdalla et al., 2019; Fujita et al., 2020 |

(Continued)
TABLE 1 | Continued

| lncRNA | Function | Mode of action | Interacting proteins | Association with chromatin topology | Conservation | References |
|--------|----------|----------------|----------------------|--------------------------------------|-------------|-----------|
| IPW    | Repression of maternally expressed genes. Possibly implicated in Prader-Willi syndrome | in trans | EHMT2 | Allele-specific formation of heterochromatin at DLK1–DIO3 region. | Poorly conserved by sequence between human and mouse | Wewrick et al., 1994; Wewrick and Francke, 1997; Stelzer et al., 2014 |
| Fire   | Role in adipogenesis, nuclear architecture, inflammatory response (in vitro), and hematopoiesis (in vivo) | in trans, but occupies domain in cis | HNRNPU | Fire acts as a scaffold for the formation of an inter-chromosomal structure. Locates at border of TAD in a CTCF binding region. Required for super loop formation of inactive X | Conservation across mammals, high convergence of repeating domain in primates. Local repeats are conserved between species of the same order | Hacisuleyman et al., 2014, 2016; Yang et al., 2015; Lu et al., 2017; Barutcu et al., 2018; Lewandowski et al., 2019 |
| TERRA  | Implicated in telomeric and subtelomeric heterochromatin formation, stability and maintenance | in cis (to telomeres) and in trans | Shelterin components (TERF1 and TERF2), ORC1, GBX5 NoRC, ATRX, POT1, and others | TERRA transcription depends on chromosome looping | Telomere transcription is conserved across vertebrates and Saccharomyces cerevisiae | Azzalin et al., 2007; Luke et al., 2008; Schoeftner and Blasco, 2008; Deng et al., 2009; Postepiska-Igielska et al., 2013; Beishline et al., 2017 |

candidates for experimental functional characterization. Sequence conservation can be identified using traditional computational sequence comparison methods. Recent examples have shown that conserved sequence stretches can be much shorter in lncRNAs than in protein-coding sequences, highlighting the need to look for tiny stretches of sequence conservation (microhomology; Quinn et al., 2016). Positional conservation of lncRNAs can be identified using multiple genome alignments complemented with transcriptomic data that support the existence of non-coding transcripts in multiple taxa. The detection of splice site conservation uses a similar approach but focuses on identifying splice sites via modeling or direct RNA-seq evidence, followed by comparison across taxa (Nitsche et al., 2015). In the case of structural conservation, covariation signatures in multiple sequence alignments may indicate the conservation of a structure (Nawrocki et al., 2009; Gruber et al., 2010; Will et al., 2012). One of the most significant limitations is the difficult problem of distinguishing covariation from sequence conservation. Thus, these methods can better identify conserved structures in highly varying sequences in diverse and multiple taxa (Rivas et al., 2017, 2020).

In the context of studying novel lncRNAs, its unique conservation signatures, albeit more difficult to detect, are excellent ways to identify potentially functional lncRNA candidates and give a first insight on their possible mechanisms of action. They can also help guide the search for homologous mechanisms in other species. Complementing in silico studies with experimental approaches in the context of spatiotemporal gene expression programs is crucial to further assess the impact of these ncRNAs on modulating genome architecture, including their specific contribution to the complexity and evolution of animal gene regulation.

AUTHOR CONTRIBUTIONS

All authors participated in writing and reviewing the manuscript and approved the final version for publication.

FUNDING

AR-C was funded by the Consejo Nacional de Ciencia y Tecnología (CONACYT) M.Sc. fellowship. KO and SF-V were funded by the Newton Advanced Fellowship (No. NAF\R1\180303) awarded to SF-V.

REFERENCES

Abdalla, M. O. A., Yamamoto, T., Maehara, K., Yogami, J., Ohkawa, Y., Miura, H., et al. (2019). The Eleanor ncRNAs activate the topological domain of the ESRI locus to balance against apoptosis. Nat. Commun. 10:3778. doi: 10.1038/s41467-019-11378-4

Andersson, D., Muckenhuber, M., Bammer, P. C., Kulinski, T. M., Theusl, H.-C., Shimizu, T., et al. (2019). The Airn IncRNA does not require any DNA elements within its locus to silence distant imprinted genes. PLoS Genet. 15:e1008268. doi: 10.1371/journal.pgen.1008268

Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E., and Lingner, J. (2007). Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science 318, 798–801. doi: 10.1126/science.1147182

Babak, T., Blencowe, B. J., and Hughes, T. R. (2005). A systematic search for new mammalian noncoding RNAs indicates little conserved intergenic transcription. BMC Genom. 6:104. doi: 10.1186/1471-2164-6-104

Barutcu, A. R., Maass, P. G., Lewandowski, J. P., Weiner, C. L., and Rinn, J. L. (2018). A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. Nat. Commun. 9:1444. doi: 10.1038/s41467-018-03614-0

Barutcu, A. R., Maass, P. G., Lewandowski, J. P., Weiner, C. L., and Rinn, J. L. (2017). Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science 318, 798–801. doi: 10.1126/science.1147182

Beishline, S. F., Winckler, M., and Pachovsky, J. E. (2015). Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science 318, 798–801. doi: 10.1126/science.1147182

Babak, T., Blencowe, B. J., and Hughes, T. R. (2005). A systematic search for new mammalian noncoding RNAs indicates little conserved intergenic transcription. BMC Genom. 6:104. doi: 10.1186/1471-2164-6-104

Barutcu, A. R., Maass, P. G., Lewandowski, J. P., Weiner, C. L., and Rinn, J. L. (2018). A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. Nat. Commun. 9:1444. doi: 10.1038/s41467-018-03614-0

Babak, T., Blencowe, B. J., and Hughes, T. R. (2005). A systematic search for new mammalian noncoding RNAs indicates little conserved intergenic transcription. BMC Genom. 6:104. doi: 10.1186/1471-2164-6-104

Barutcu, A. R., Maass, P. G., Lewandowski, J. P., Weiner, C. L., and Rinn, J. L. (2018). A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. Nat. Commun. 9:1444. doi: 10.1038/s41467-018-03614-0

Babak, T., Blencowe, B. J., and Hughes, T. R. (2005). A systematic search for new mammalian noncoding RNAs indicates little conserved intergenic transcription. BMC Genom. 6:104. doi: 10.1186/1471-2164-6-104

Barutcu, A. R., Maass, P. G., Lewandowski, J. P., Weiner, C. L., and Rinn, J. L. (2018). A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. Nat. Commun. 9:1444. doi: 10.1038/s41467-018-03614-0

Babak, T., Blencowe, B. J., and Hughes, T. R. (2005). A systematic search for new mammalian noncoding RNAs indicates little conserved intergenic transcription. BMC Genom. 6:104. doi: 10.1186/1471-2164-6-104

Barutcu, A. R., Maass, P. G., Lewandowski, J. P., Weiner, C. L., and Rinn, J. L. (2018). A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. Nat. Commun. 9:1444. doi: 10.1038/s41467-018-03614-0

Babak, T., Blencowe, B. J., and Hughes, T. R. (2005). A systematic search for new mammalian noncoding RNAs indicates little conserved intergenic transcription. BMC Genom. 6:104. doi: 10.1186/1471-2164-6-104

Barutcu, A. R., Maass, P. G., Lewandowski, J. P., Weiner, C. L., and Rinn, J. L. (2018). A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. Nat. Commun. 9:1444. doi: 10.1038/s41467-018-03614-0

Babak, T., Blencowe, B. J., and Hughes, T. R. (2005). A systematic search for new mammalian noncoding RNAs indicates little conserved intergenic transcription. BMC Genom. 6:104. doi: 10.1186/1471-2164-6-104

Barutcu, A. R., Maass, P. G., Lewandowski, J. P., Weiner, C. L., and Rinn, J. L. (2018). A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. Nat. Commun. 9:1444. doi: 10.1038/s41467-018-03614-0

Babak, T., Blencowe, B. J., and Hughes, T. R. (2005). A systematic search for new mammalian noncoding RNAs indicates little conserved intergenic transcription. BMC Genom. 6:104. doi: 10.1186/1471-2164-6-104

Barutcu, A. R., Maass, P. G., Lewandowski, J. P., Weiner, C. L., and Rinn, J. L. (2018). A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. Nat. Commun. 9:1444. doi: 10.1038/s41467-018-03614-0
Narendra, V., Rocha, P. P. A., D’Aviram, R., Skok, J. A., Mazzoni, E. O., et al. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science 347, 1017–1021. doi: 10.1126/science.1258888

Nawrocki, E. P., Kolbe, D. L., and Eddy, S. R. (2009). Infernal 1.0: inference of RNA alignments. Bioinformatics 25, 1335–1337. doi: 10.1093/bioinformatics/btp326

Necsulea, A., Soumillon, M., Warnefors, M., Liechti, A., Daish, T., Zeller, U., et al. (2014). The evolution of IncRNA repertoires and expression patterns in tetrapods. Nature 505, 635–640. doi: 10.1038/nature12943

Nesterova, T. B., Slobodyanyuk, S. Y., Elisaphenko, E. A., Shevchenko, A. I., Johnston, C., Pavlova, M. E., et al. (2001). Characterization of the genomic Xist locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence. Genomes Res. 11, 833–849. doi: 10.1101/gr.174901

Nitsche, A., Rose, D., Fasold, M., Reiche, K., and Stadler, P. F. (2015). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129, 1311–1323. doi: 10.1016/j.cell.2007.05.022

Rivas, E., Clements, J., and Eddy, S. R. (2017). A statistical test for conserved RNA structure shows lack of evidence for structure in IncRNAs. Nat. Methods 14, 45–48. doi: 10.1038/nmeth.4066

Rivas, E., Clements, J., and Eddy, S. R. (2020). Estimating the power of sequence covariance for detecting conserved RNA structure. Bioinformatics 36, 3072–3076. doi: 10.1093/bioinformatics/btaa780

Santoro, F., Mayer, D., Klement, R. M., Waczko, K. E., Stukalov, A., Barlow, D. P., et al. (2013). Imprinted Ig2r silencing depends on continuous Airn IncRNA expression and is not restricted to a developmental window. Development 140, 1184–1195. doi: 10.1242/dev.088849

Santoro, R., Schmitz, K. -M., Sandoval, J., and Grummt, I. (2010). Intergenic transcripts originating from a subclass of ribosomal DNA repeats silence ribosomal RNA genes in trans. EMBO Rep. 11, 52–58. doi: 10.1038/embor.2009.254

Sarma, K., Cifuentes-Rojas, C., Ergun, A., Del Rosario, A., Jeon, Y., White, F., et al. (2014). ATRX directs binding of PRC2 to XIST RNA and Polycomb targets. Cell 159, 869–883. doi: 10.1016/j.cell.2014.10.019

Savic, N., Bär, D., Leone, S., Frommel, S. C., Weber, F. A., Vollenweider, E., et al. (2014). IncRNA saturation to initiate heterochromatin formation in the nucleolus is required for exit from pluripotency in ESCs. Cell Stem Cell 15, 720–734. doi: 10.1016/j.stem.2014.10.005

Saxena, A., and Carninci, P. (2011). Long non-coding RNA modifies chromatin: epigenetic silencing by long non-coding RNAs. Bioessays 33, 830–839. doi: 10.1002/bies.201000884

Schmitt, A. M., and Chang, H. Y. (2016). Long noncoding RNAs in cancer pathways. Cancer Cell 29, 452–463. doi: 10.1016/j.ccell.2016.03.010

Schmitz, K. -M., Mayer, C., Postepska, A., and Grummt, I. (2010). Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. Genes Dev. 24, 2264–2269. doi: 10.1101/gad.590910

Schoefnater, S., and Blasco, M. A. (2008). Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. Nat. Cell Biol. 10, 228–236. doi: 10.1038/ncb1685

Schoenfelder, S., and Fraser, P. (2019). Long-range enhancer-promoter contacts in gene expression control. Nat. Rev. Genet. 20, 457–455. doi: 10.1038/s41576-019-0128-0

Schorderet, P., and Duboule, D. (2011). Structural and functional differences in the long non-coding RNA hotair in mouse and human. PLoS Genet. 7:e1002071. doi: 10.1371/journal.pgen.1002071

Seidl, C. I. M., Stricker, S. H., and Barlow, D. P. (2006). The imprinted air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. EMBO J. 25, 3565–3575. doi: 10.1038/sj.emboj.7601245

Shi, Y., Downes, M., Xie, W., Kao, H. Y., Ordentlich, P., Tsai, C. C., et al. (2001). Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. Genes Dev. 15, 1140–1151. doi: 10.1101/gad.871201

Simon, M. D., Pinter, S. F., Fang, R., Sarma, K., Rutenberg-Schoenberg, M., Bowman, S. K., et al. (2013). High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. Nature 504, 465–469. doi: 10.1038/nature12719

Sletiefs, F., Zwart, R., and Barlow, D. P. (2002). The non-coding air RNA is required for silencing autosomal imprinted genes. Nature 415, 810–813. doi: 10.1038/nature015810a

Smith, M. A., Gesell, T., Stadler, P. F., and Mattick, J. S. (2013). Widespread purifying selection on RNA structure in mammals. Nucleic Acids Res. 41, 8220–8236. doi: 10.1093/nar/gkt596

Somarouthu, S., Legiewicz, M., Chillon, I., Marcia, M., Liu, F., and Pyle, A. M. (2015). HOTAIR forms an intricate and modular secondary structure. Mol. Cell 58, 353–361. doi: 10.1016/j.molcel.2015.03.006

Sridhar, B., Rivas-Astroza, M., Nguyen, T. C., Chen, W., Yan, Z., Cao, X., et al. (2017). Systematic mapping of RNA-chromatin interactions in vivo. Curr. Biol. 27, 602–609. doi: 10.1016/j.cub.2017.01.011
Stelzer, Y., Sagi, I., Yamuka, O., Eiges, R., and Benvenisty, N. (2014). The noncoding RNA IPW regulates the imprinted DLK1-DIO3 locus in an induced pluripotent stem cell model of Prader-Willi syndrome. *Nat. Genet.* 46, 551–557. doi: 10.1038/ng.2968

Tavares, R. C. A., Pyle, A. M., and Somarowithu, S. (2019). Phylogenetic analysis with improved parameters reveals conservation in IncRNA structures. *J. Mol. Biol.* 431, 1592–1603. doi: 10.1016/j.jmb.2019.03.012

Tomita, S., Abdalla, M. O. A., Fujiwara, S., Matsumori, H., Maehara, K., Ohkawa, Y., et al. (2015). A cluster of noncoding RNAs activates the ESR1 locus during breast cancer adaptation. *Nat. Commun.* 6:6966. doi: 10.1038/ncomms7966

Trimarchi, T., Bilal, E., Ntzirimochristos, P., Fabbri, G., Dalla-Favera, R., Tisigros, A., et al. (2014). Genome-wide mapping and characterization of notochord-regulated long noncoding RNAs in acute leukemia. *Cell* 158, 593–606. doi: 10.1016/j.cell.2014.05.049

Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., et al. (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* 39, 925–938. doi: 10.1016/j.molcel.2010.08.011

Tsai, M.-C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., et al. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329, 689–693. doi: 10.1126/science.1192002

Ulitsky, I. (2016). Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. *Nat. Rev. Genet.* 17, 601–614. doi: 10.1038/nrg.2016.85

Ulitsky, I., Lin, F., He, S., Seidman, J. G., and Bartel, D. P. (2011). Conserved function of IncRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* 147, 1537–1550. doi: 10.1016/j.cell.2011.11.055

van Bakel, H., Nislow, C., Blencowe, B. J., and Hughes, T. R. (2010). Most “dark matter” transcripts are associated with known genes. *PLoS Biol.* 8:e1000371. doi: 10.1371/journal.pbio.1000371

van Bemmelen, J. G., Galupa, R., Gard, C., Servant, N., Picard, C., Davies, J., et al. (2019). The bipartite TAD organization of the X-inactivation center ensures opposing developmental regulation of Tsix and Xist. *Nat. Genet.* 51, 1024–1034. doi: 10.1038/s41588-019-0412-0

Wang, K. C., and Chang, H. Y. (2011). Molecular mechanisms of long noncoding RNAs. *Cell* 43, 904–914. doi: 10.1016/j.molcel.2011.08.018

Wang, F., Tang, Z., Shao, H., Guo, J., Tan, T., Dong, Y., et al. (2018). Long noncoding RNA HOTTIP cooperates with CCCTC-binding factor to coordinate HOXA gene expression. *Biochem. Biophys. Res. Commun.* 500, 852–859. doi: 10.1016/j.bbrc.2018.04.173

Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., et al. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472, 120–124. doi: 10.1038/nature09819

Wehner, S., Dörrich, A. K., Ciba, P., Wilde, A., and Marz, M. (2014). pRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. *Cell* 157, 773–788. doi: 10.1016/j.cell.2014.05.054

Will, S., Joshi, T., Hofacker, I. L., Stadler, P. F., and Backofen, R. (2012). LocARNA: accurate boundary prediction and improved detection of structural RNAs. *RNA* 18, 900–914. doi: 10.1261/rna.029041.111

Wongtrakoonpate, P., Riddick, G., Fucharoen, S., and Felsenfeld, G. (2015). Association of the long non-coding RNA steroid receptor RNA activator (SRA) with TraG and PRC2 complexes. *PLoS Genet.* 11:e1005615. doi: 10.1371/journal.pgen.1005615

Wutz, A., and Jaenisch, R. (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell* 5, 695–705. doi: 10.1016/S1097-2765(00)00248-8

Yang, F., Deng, X., Ma, W., Berletch, J. B., Rabaia, N., Wei, G., et al. (2015). The IncRNA Fire1 anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. *Genome Biol.* 16:52. doi: 10.1186/s13059-015-0618-0

Yang, L., Lin, C., Liu, W., Zhang, J., Ohgi, K. A., Grinstein, J. D., et al. (2011). ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. *Cell* 147, 773–788. doi: 10.1016/j.cell.2011.08.054

Yao, H., Brick, K., Evrard, Y., Xiao, T., Camerini-Otero, R. D., and Felsenfeld, G. (2010). Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes Dev.* 24, 2543–2555. doi: 10.1101/gad.1967810

Yotova, I. Y., Vlatkovic, I. M., Pauler, F. M., Warzczok, K. E., Ambros, P. F., Oshimura, M., et al. (2008). Identification of the human homolog of the imprinted mouse air non-coding RNA. *Genomics* 92, 464–473. doi: 10.1016/j.ygeno.2008.08.004

Zhang, B., Arun, G., Mao, Y. S., Lazar, Z., Hung, G., Bhattacharjee, G., et al. (2012). The IncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Rep.* 2, 111–123. doi: 10.1016/j.celrep.2012.06.003

Zhang, H., Zeitz, M. J., Wang, H., Niu, B., Ge, S., Li, W., et al. (2014). Long noncoding RNA-mediated intrachromosomal interactions promote imprinting at the Kcnq1 locus. *J. Cell Biol.* 204, 61–75. doi: 10.1083/jcb.201304152

Zhang, A., Zhao, J. C., Kim, J., Feng, K. W., Yang, Y. A., Chakravarti, D., et al. (2015). IncRNA HOTAIR enhances the androgen-receptor-mediated transcriptional program and drives castration-resistant prostate cancer. *Cell Rep.* 13, 209–221. doi: 10.1016/j.celrep.2015.08.069

Zhao, X., Patton, J. R., Ghosh, S. K., Fischel-Ghodsian, N., Shen, L., and Spanjaard, R. A. (2007). Pus3p- and Pus1p-dependent pseudouridylation of proteins targeted by a short repeat RNA to the mouse X chromosome. *Cell* 129, 1727–1739. doi: 10.1016/j.cell.2007.09.017

Copyright © 2020 Ramirez-Colmenero, Oktaba and Fernandez-Valverde. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.