



# Emerging Roles of Repetitive and Repeat-Containing RNA in Nuclear and Chromatin Organization and Gene Expression

Giuseppe Trigiantef, Nerea Blanes Ruiz† and Andrea Cerase\*

Centre for Genomics and Child Health, Barts and The London School of Medicine and Dentistry, Blizard Institute, Queen Mary University of London, London, United Kingdom

## OPEN ACCESS

### Edited by:

Peter Askjaer,  
Andalusian Center for Development  
Biology (CABD), Spain

### Reviewed by:

Pascal Chartrand,  
Université de Montréal, Canada  
Argyris Papantonis,  
University Medical Center Göttingen,  
Germany

### \*Correspondence:

Andrea Cerase  
a.cerase@qmul.ac.uk

† These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Nuclear Organization and Dynamics,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

Received: 02 July 2021

Accepted: 09 September 2021

Published: 06 October 2021

### Citation:

Trigiantef G, Blanes Ruiz N and  
Cerese A (2021) Emerging Roles  
of Repetitive and Repeat-Containing  
RNA in Nuclear and Chromatin  
Organization and Gene Expression.  
*Front. Cell Dev. Biol.* 9:735527.  
doi: 10.3389/fcell.2021.735527

Genomic repeats have been intensely studied as regulatory elements controlling gene transcription, splicing and genome architecture. Our understanding of the role of the repetitive RNA such as the RNA coming from genomic repeats, or repetitive sequences embedded in mRNA/lncRNAs, in nuclear and cellular functions is instead still limited. In this review we discuss evidence supporting the multifaceted roles of repetitive RNA and RNA binding proteins in nuclear organization, gene regulation, and in the formation of dynamic membrane-less aggregates. We hope that our review will further stimulate research in the consolidating field of repetitive RNA biology.

**Keywords:** epigenetics, nuclear organization, long non-coding RNA, tandem repeats, repetitive RNA, membraneless compartments, X chromosome inactivation (XCI), *Xist* (*X-inactive specific transcript*)

## NON-CODING RNAs IN THE NUCLEUS

The eukaryotic nucleus is a very complex organelle containing large amounts of DNA, RNA, proteins (Boija et al., 2018; Sawyer et al., 2019) and nuclear compartments, such as membrane-less organelles like the nucleoli, speckles, paraspeckles. These organelles are droplet-like particles that exhibit liquid-like features and are believed to be formed via a process called liquid-liquid phase separation (LLPS, discussed below).

The mechanisms behind the differential expression of protein coding and non-coding genes in time, space and in response to external stimuli, is the object of intensive studies. It is well known that, transcriptionally, chromatin exists in one of two main states: euchromatin, actively transcribed and not condensed, and heterochromatin, relatively compacted, gene-poor and transcriptionally silent (Trojer and Reinberg, 2007), and that these share different nuclear compartments (Van Bortle and Corces, 2012). The fraction of the heterochromatin that is permanently silent is called constitutive heterochromatin, whereas the fraction of chromatin that undergoes transcriptional activation or repression depending on cellular, extracellular and developmental clues is called facultative heterochromatin. Constitutive heterochromatin is enriched for repetitive DNA (for example the centromeric and telomeric regions and genomic repeats) (Nishibuchi and Déjardin, 2017).

While the role of repetitive genomic elements has been intensely studied, the role of repetitive or repeat-containing RNA has only relatively recently gained some attention. It is now becoming increasingly clear that mRNAs and nuclear non-coding RNAs are centrally involved in the regulation of chromatin state and therefore gene expression. For example, it has been known for decades that RNA can be intimately associated to chromatin (Holmes and Bonner, 1974). This type of RNA is now collectively labeled as chromatin-associated RNA (caRNA) (Li and Fu, 2019). This term encompasses a wide variety of specific RNAs such as short nuclear/nucleolar (snRNA/snoRNA) and long non-coding RNA (lncRNA) and newly transcribed mRNA (Li and Fu, 2019). While their specific action mechanism is mostly unknown, it is now generally believed that they act in combination with RNA-binding proteins (RBP) and histone modifying proteins such as, for example, the Suppressor of Variegation 3–9 (SUV39), Polycomb Repressor complex (PRC1 and 2) (Davidovich and Cech, 2015; Cerase and Tartaglia, 2020), and hNRNP-family protein (Klimek-Tomczak et al., 2004; West et al., 2019) to regulate gene expression and nuclear dynamics.

Due to recent excellent reviews on the topic of lncRNA and RNA in general nuclear organization (Creamer and Lawrence, 2017; Khosraviani et al., 2019; Michieletto and Gilbert, 2019; Thakur et al., 2019), this review will focus on the emerging evidence of the role of repetitive RNA and repeat-containing repetitive motifs (i.e., genomic, simple/tandem containing repeats) in the organization of chromatin and the regulation of gene expression.

## REPETITIVE caRNA

Repetitive caRNAs are transcribed from repetitive DNA sequences such as DNA-containing genomic repeats, fragments of genomic repeats or repeated motifs such as simple repeats (Biscotti et al., 2015). It is believed that repetitive sequences make up at least half of the human genome, with some estimates placing it at two thirds (de Koning et al., 2011). Such sequences were once labeled the “dark matter” of the genome or “junk DNA,” but it is becoming more and more clear that they instead play critical roles in regulating gene expression at different levels (Statello et al., 2021). They can be classified in two broad categories: low complexity, consisting of adjacent sequences repeated in tandem, and interspersed repeats, complex sequences generally capable of being transcribed and sometimes translated. Examples of tandem repeats are the telomeric and centromeric repeats and satellite DNA (Janssen et al., 2018). These repeats are often associated with a constitutive heterochromatin state even though a basal transcription of these regions has not only been detected but is now considered key to the maintenance of the condensed state, as it will be discussed later.

The vast majority of repetitive DNA consists of interspersed repeats, also called transposons or transposable elements (TEs) (Saleh et al., 2019). These have the capacity to migrate or replicate across the genome like many types of viral DNA sequences (Saleh et al., 2019). TEs are further classified into retrotransposons (comprising Long Interspersed Nuclear Elements (LINE), Short

Interspersed Nuclear Elements (SINE) and long tandem repeats (LTR) retrotransposons, which include endogenous retroviruses) and DNA transposons (Chuong et al., 2017). Retrotransposons replicate via an RNA intermediate transcript which is later reverse transcribed into the target site (“copy and paste” mechanism). DNA transposons rely on a transposase to physically relocate to another site (Pace and Feschotte, 2007). The most abundant and important TEs in the human genome are the LINE L1 and the SINE *Alu*. Together they make up around 30% of the human genome. Most of these TEs are however silenced as a possible cellular mechanism of defense, but a relatively small number of them (labeled as retrotransposition-competent LINE-1 or RC-L1) is responsible for a large amount of nuclear transcription (Chuong et al., 2017; Saleh et al., 2019). Noticeably, while genomic repeats are, *per se*, silenced, a large fraction of TE fragments and low-complexity DNA is embedded and transcribed in mRNAs and ncRNA (Fort et al., 2021), and accounts for the largest share of nuclear transcription (Saleh et al., 2019).

Transposable elements were generally considered a remnant of a viral or parasitic insertion in the genome with no positive function and, if anything, a source of mutation and disease occasioned by their random insertion into promoters or coding sequences. Modern research has instead revealed that the repetitive RNA produced from TE may have been “exapted” or exploited by the cell to carry out some important functions in the regulation of gene expression (Chuong et al., 2017). For example, the *Alu* repeats were identified as a nuclear localization motif for RNA (Lubelsky and Ulitsky, 2018) and two thirds of lncRNAs are reported to contain TE elements (Chuong et al., 2017). Interestingly, the pattern of expression of long intergenic non-coding RNAs (lincRNA) shows a decrease in TE element content during differentiation, possibly highlighting a connection to cell stemness and early embryonal stages (Kelley and Rinn, 2012).

In the next paragraphs we will briefly examine the role of medium repetitive RNAs (such tandem and interspersed) in the establishment and maintenance of the two heterochromatin and the euchromatin states.

## HETEROCHROMATIN AND ASSOCIATED REPETITIVE RNAs

### Constitutive Heterochromatin

Constitutive Heterochromatin is the transcriptionally silent, permanently condensed form of chromatin (Janssen et al., 2018). It is a state set and maintained in regions such as the chromosome centromere and telomeres, which have important structural roles and encode no proteins. Both regions are characterized by tandem repeat DNA, although of a different kind.

Centromeres are the sites of kinetochore attachment during mitosis and have obvious critical importance in cell survival (Aldrup-Macdonald and Sullivan, 2014). They vary in sequence among species but in humans and primates they contain  $\alpha$ -satellite DNA, a long tandem repeat sequence of a unit (the Higher Order Repeat, HOR), which is itself formed of a specific number of 171 bp monomers. The  $\alpha$ -satellite

DNA is surrounded by a pericentromeric region formed by the same monomers but without the HOR organization (Aldrup-Macdonald and Sullivan, 2014).

Centromeric HC is defined by the presence of H3K9 methylation (H3K9Me; Me1, Me2, Me3 : mono-, di- and trimethylation), common to all constitutive HC forms. This marker is deposited by histone methyltransferase SUV39 and recognized by HP1 (Heterochromatin Protein 1); however, RNA, and specifically the repetitive RNA proceeding from the centromeric region, is necessary for histone methylation to take place and silencing to be maintained (Velazquez Camacho et al., 2017). In the fission yeast *S. pombe* the silencing occurs via an RNA-dependent RNA polymerase (RdRP) acting on transcribed repeat RNA and forming dsRNA which on one hand triggers the Argonaute protein binding and subsequent transcriptional silencing and, on the other hand, recruits an RNA-induced transcriptional silencing (RITS) complex able to bind the histone methyltransferase Clr4 (homolog of human SUV39) to deposit the required H3K9Me marks. A feed-forward mechanism sees these methylated H3 recognized by another protein (Sw16, homolog of human HP1) (Li and Fu, 2019).

While this mechanism has been very well studied in yeast, it has been ruled out in mammals by the lack of a mammalian RdRP (Li and Fu, 2019). An alternative mechanism suggested sees the intervention of piwi-associated RNAs (piRNA) in what has been labeled a “ping-pong” mechanism, which uses piRNA transcripts and retrotransposons to amplify the RITS/SUV39 methylation activity without the need for an RdRP (Li and Fu, 2019). It is also apparent that HP1 can directly bind centromeric repetitive RNA (Maison et al., 2011) just like SUV39H1 (Johnson et al., 2017). In summary, the stable silencing of centromeric chromatin—paradoxically—requires a certain baseline amount of centromeric transcription (and repetitive RNA molecules), in what is considered to be a mechanism designed to re-establish the silencing following DNA replication, which removes epigenetic marks and allows HC to be briefly transcribed (Volpe et al., 2002).

Telomeres are found at the terminal sequences of every eukaryotic chromosome and their main function is to preserve chromosome integrity during each round of DNA replication. Critically, they prevent the recognition of chromosome ends as DNA breaks and also limit somatic cellular proliferation by inducing senescence when shortened below a critical length (Palm and de Lange, 2008). Telomeres are constitutively silenced and contain the conserved tandem repeat TTAGGG along several kb (10–15 in humans, 20–50 in mouse). The actual chromosome end is arranged in a complex loop structure made possible by a 50–500 bp 3′ overhang of the + (“G”) strand (McElligott and Wellinger, 1997). The telomeric region is transcribed into a long non-coding RNA called TERRA (TElomeric Repeat containing RNA) whose promoter and transcription start lie in a poorly defined subtelomeric region containing a conserved 61 bp element, followed by 29 bp, and 37 bp repeats (61–29–37 repeats) (Cusanelli and Chartrand, 2014; **Figure 1A**). TERRA transcripts vary in length between a hundred bp and about 9 kb and stem from the – (“C”) telomeric strand, therefore being copies of the + strand and G-rich (Palm and de Lange, 2008). The role of TERRA in telomere maintenance is still the object of active research but it

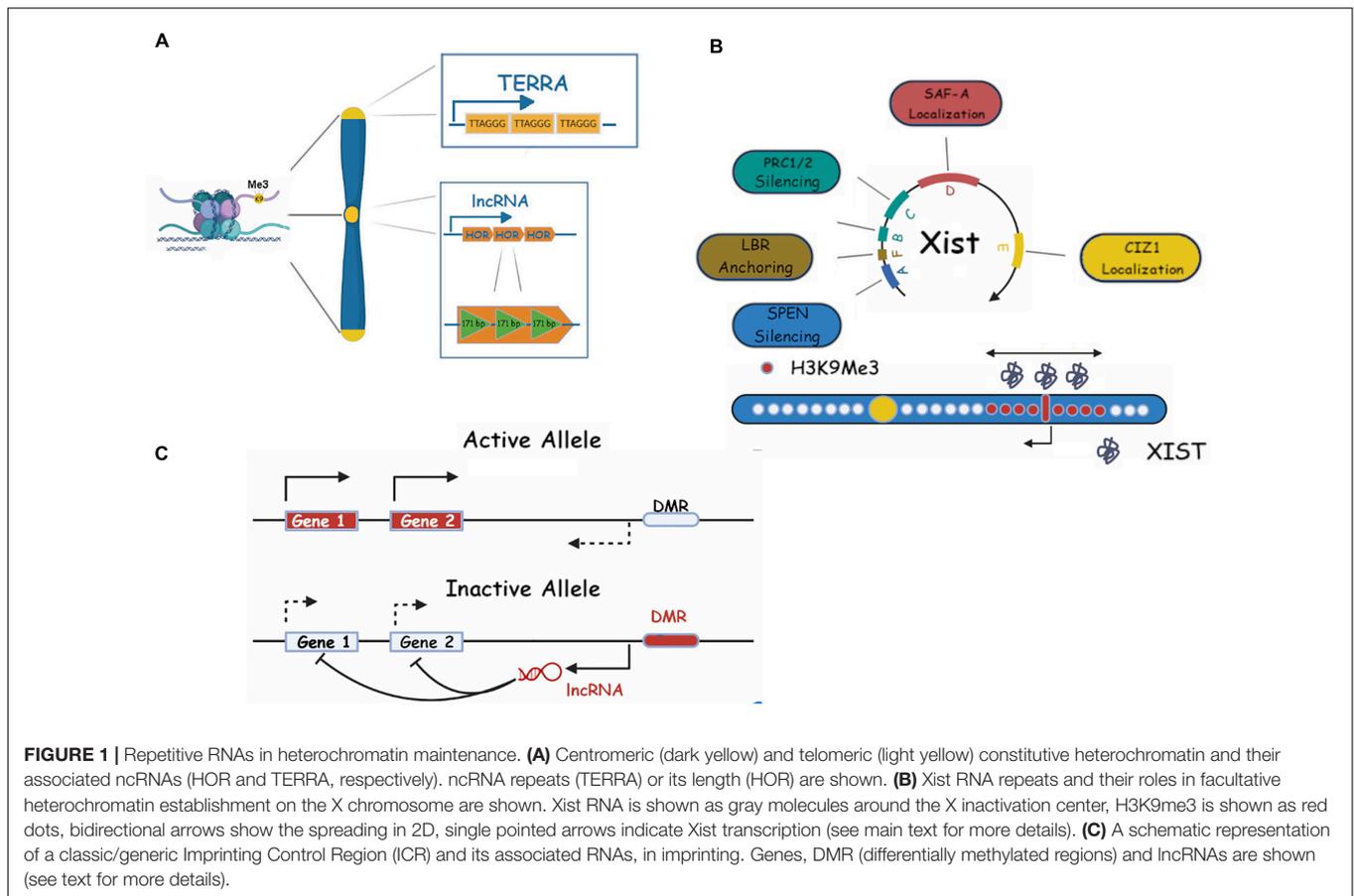
is clear that it is involved in telomere elongation via telomerase regulation and heterochromatin formation and maintenance. Indeed, TERRA has been shown to be bound by HP1 and to bind H3K9Me3 and its expression negatively correlates with this histone PTM. It also binds the Origin Replication Complex (ORC) and subunits of the Shelterin telomere complex (TRF1 and 2). Furthermore, siRNA depletion of TERRA results in H3K9 hypomethylation of the telomeric region (Deng et al., 2009), highlighting its role in stable telomere silencing.

## Facultative Heterochromatin

This term describes chromosomal regions which are silenced in a conditional way, being transcribed or silenced when circumstances, such as a certain stage of embryonic development or tissue differentiation, require it (Cerase et al., 2015; Żylicz and Heard, 2020). Facultative heterochromatin is also characterized by histone PTMs. For instance, during the process of X chromosome inactivation (XCI), the methylation of lysine 27 on Histone 3 (H3K27Me), deposited by the Polycomb repressive complex 2 (PRC2) is common in mouse, and H3K27me3 and H3K9me2–3 are instead common in humans (Chadwick and Willard, 2004; Schuettengruber et al., 2017). The typical fHC scenario is that in which a region of the genome becomes specifically silenced in a developmentally regulated process. In this case, the resulting heterochromatin can remain irreversibly condensed, be temporarily reactivated (Patrat et al., 2009) or even be silenced again (Cerase et al., 2015), via incompletely understood mechanisms.

In this review, we will focus our attention to two of the most studied fHC events: XCI and genomic imprinting (Żylicz and Heard, 2020).

*X chromosome inactivation* is a universal event in female mammalian development. As females have two X chromosomes, the simultaneous expression of both would result in abnormal protein dosage and lethal consequences (Borensztein et al., 2017). As a compensation mechanism, one of the X chromosomes is inactivated in the early stages of embryo development. Interestingly, in mouse, the event unfolds in two waves: in an early silencing step, at the four-cell stage, the paternal X chromosome is inactivated, after which reactivation occurs and at the implantation stage a second XCI wave, this time occurring on a random chromosome, takes place (Patrat et al., 2009). XCI is a complex process articulated in several stages, all orchestrated by the master regulator lncRNA *Xist* (*X inactive specific transcript*) (Brown et al., 1991). This transcript, around 15–17 kb long, contains six tandem repeats, labeled A to F, which are essential for *Xist* activities (Pintacuda et al., 2017b; **Figure 1B**). Repeat A consists of 8.5 copies (7.5 in mouse) of a 24 base pair sequence. It is necessary for XCI to occur and it has been shown to directly bind SPEN (SPLIT ENds homolog) (Monfort and Wutz, 2020). This binding recruits the histone deacetylase HDAC3 to deacetylate histone H3 thus clearing the way for repressive post-translational modifications (PMTs) (Monfort and Wutz, 2020). The *Xist* A repeat has been suggested, by modeling studies and experimental data (Lu et al., 2020a) to provide a scaffold for RNA-RNA interaction and hence multimerization of the transcript (Duszczak et al., 2011). It is tempting to somehow link this multimerization process to the mechanism of the



coating event, however, direct experimental evidence is currently not available. A recent paper by Rodermund et al. (2021) has shed light on the temporal mechanism of Xist spreading via time resolved structural illumination microscopy (3D-SIM). This study highlighted the role of the A repeat and related binding proteins such as Spen in mouse ES cells and neuronal progenitor cells (NPCs), in the spreading process. Repeat B, which consists of 32 copies of a cytidine rich hexamer, and C, which contains 14 copies of a 120-nucleotide unit, directly binding hnRNPK, recruit the Polycomb factor PRC1 which results in ubiquitination of H2AK119 (Pintacuda et al., 2017a). This modification is in turn recognized by the PRC2 complex which deposits the final H3K27 silencing modification (Blackledge et al., 2015). Repeat D contains multiple copies of a 290 bp unit and has been shown to bind the nuclear scaffold organizer SAF-A (Creamer and Lawrence, 2017) (3D-SIM), although SAF-A has been shown to bind Xist RNA broadly (Cirillo et al., 2016). Repeat E, a high number (50+) repetition of a 25 bp unit, binds the CIZ1 protein which does not appear to be essential for development (Khan et al., 2018). However, its deletion triggers a lymphocyte proliferative disorder in turn linked to X dosage compensation defects, and correspondingly a delocalization of Xist and partial derepression of Xi-silenced genes (Ridings-Figueroa et al., 2017), probably through its spreading role in differentiated cells (Pinter, 2016). The short repeat F is composed of two copies of a 10-mer repeat and is involved, together with

repeat A/E, in the binding to the Lamin B Receptor (LBR), a mediator of the anchoring of the Xi chromosome to the heterochromatin rich nuclear periphery. Disruption of this bond results in Xi mislocalization and various degrees of silencing defects of the Xi genes (Chen et al., 2016). The spatial kinetics of XCI is also of great interest. While the end result is complete X inactivation via extensive chromosome coating by Xist, this process follows a spatial radiation from the Xist locus itself to the Xist entry site (XES) with genes in the proximity of these loci being silenced first. It is of note that the XES lies in 3D proximity to the Xist locus (Engreitz et al., 2013), suggesting a diffusion mediated mechanism for the Xist spread onto the X chromosome (Rodermund et al., 2021). A remarkable characteristic of XCI is its irreversibility following just a few days of differentiation. In fact, Xist is no longer required for long term maintenance of the Xi state after this period (Wutz and Jaenisch, 2000). It is believed that the state is maintained either by histone PTM (such as histone deacetylation, H3K27me3/H3K9me2-3) or by CpG island (CGI) promoter methylation (Żylicz and Heard, 2020), or stable protein accumulation (Cerese et al., 2019; Pandya-Jones et al., 2020) (discussed below).

The mechanism of Xist spreading and chromosome inactivation is not completely clear. It is known that the YY1, Ciz1 and SAF-A RNA binding proteins are required for Xist RNA spreading and silencing. Furthermore, the deletions of the Xist D repeats or the RGG (RNA binding) repeats of

SAF-A, in some cell lines but not others, are sufficient to disrupt the bond (Kolpa et al., 2016). This has led to the suggestion that these repeat sequences act as multivalent scaffolds to promote multimerization of the RNA/protein complex (Pinter, 2016). There is conflicting data regarding the redundancy of the SAF-A RGG repeats and the indispensability of its DNA binding domain (Kolpa et al., 2016), therefore the matter is presently not yet settled. Noticeably, deletion of Ciz-1 also leads to loss of Xist localization and its diffusion in the nucleoplasm, a feature which can be rescued by the reexpression of Ciz-1 (Ridings-Figueroa et al., 2017). It is possible that Xist RNA uses different chromatin anchors in different tissues and/or developmental stages.

## Genomic Imprinting

Imprinting is the process by which some genes (in mammalian genomes thought to represent a small percentage) are exclusively expressed in a monoallelic fashion, i.e., by only one of the parental alleles, the other one being permanently silenced (Tucci et al., 2019; **Figure 1C**). One example is the human *Igf2* gene, only the paternal copy of which is ever expressed (DeChiara et al., 1991). About 80 such genes have been identified so far and all their loci follow a similar expression pattern depending on the parent of origin (PoO): the region contains a demethylated CpG region called DMR (Differentially Methylated Region) also variously called IC (Imprinting Center), ICE (Imprinting Control Element), ICR (Imprinting Control Region) (Royo and Cavallé, 2008).

These regions always express at least one lncRNA, whose expression PoO generally inversely correlates with the parent of origin of the expressed allele of the neighboring genes, i.e., expression of the paternal lncRNA will result in the expression of the maternal cluster protein coding genes (Royo and Cavallé, 2008). The expression of the regulatory RNA is dependent on the methylation status of the DMR so that the methylated cluster will not express the ncRNA and will consequently express the protein coding genes. The direct mechanism for this activity has not been clarified but it is speculated that it may involve antisense binding (some lncRNAs are transcribed from the antisense strand of the protein gene) or small non-coding /microRNA activities. The end result is the *cis* silencing of the cluster protein coding genes of the PoO and consequent exclusive expression of the opposite parent's.

A well-studied case is that of *kcnq1ot1*, a lncRNA located at the distal position of chromosome 7 (Mohammad et al., 2008). Its cluster controls a large number (at least eight) of genes that are always maternally expressed, *kcnq1ot1* being always expressed paternally. The lncRNA itself is long (90 kb) but truncation studies have shown that the minimal length required for silencing is 1.7 kb with a crucial internal 860 bp 3' stretch (Mohammad et al., 2008). Interestingly, *Kcnq1ot1* contains five repeat sequences of 30 bp (MD1 repeats) which, however, do not seem to be strictly required for silencing (Mohammad et al., 2008). The silencing activity of *Kcnq1ot1* seems to be mediated, like that of *Xist*, by the deposition of H3K27Me3 PTM (Andresini et al., 2019), however, its action is limited to its genomic neighborhood.

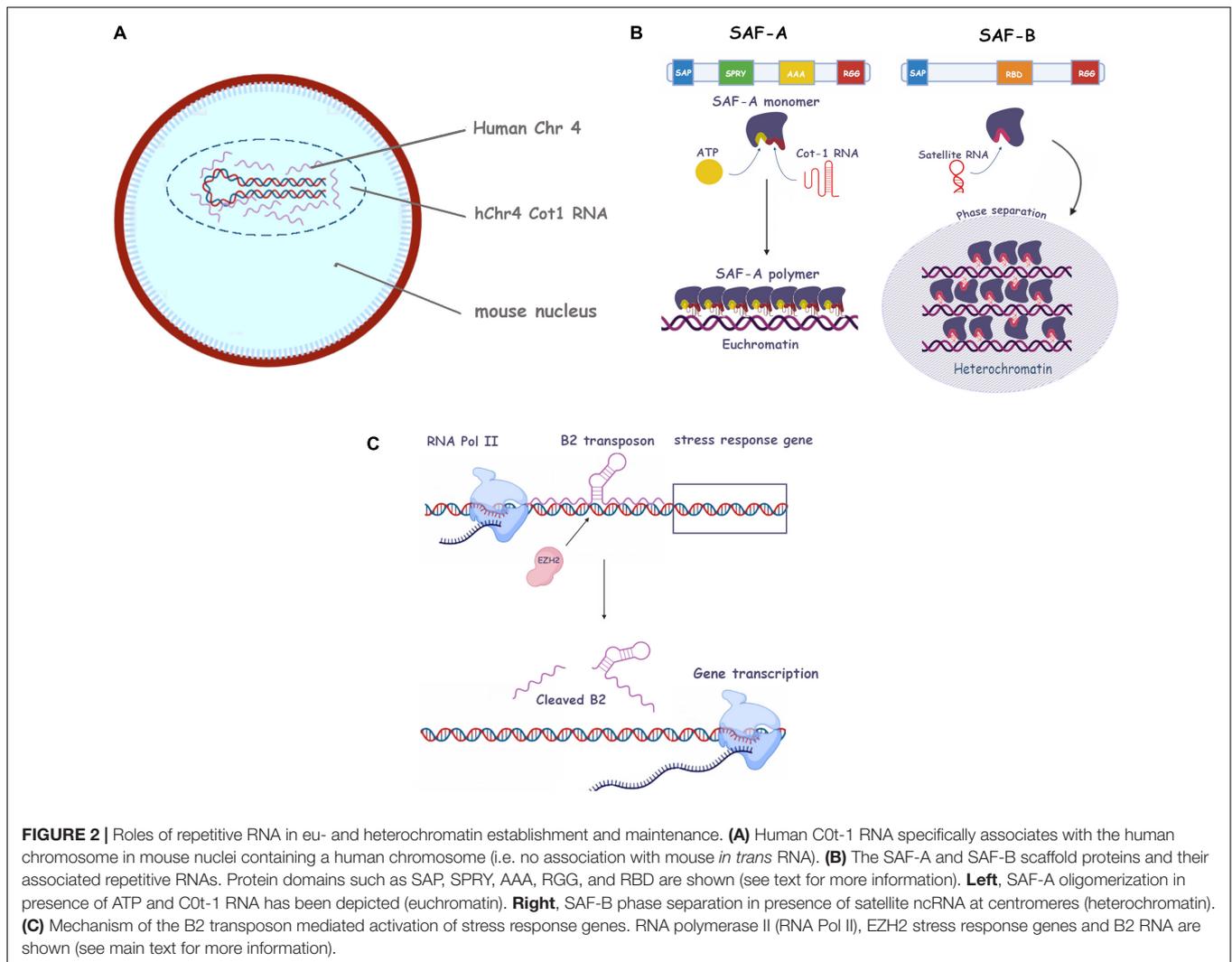
## EUCHROMATIN AND ITS REPETITIVE RNAs

Euchromatin, defined as the non-condensed, open, and transcriptionally active form of chromatin, lies at the opposite end of the spectrum to heterochromatin. As such, it is characterized by different histone PMTs and DNA methylation profiles which is quite distinct to that of HC, namely acetylation of histones or methylation on histone H3 Lys 4 and 36 among the others (Bannister and Kouzarides, 2011). However different the two chromatin states may be in appearance and behavior, they both share the feature of interacting with caRNA and, more specifically, with repetitive caRNA. A turning point in this research was the observation by Hall et al. (2014) that euchromatin is particularly enriched in C0T-1 repetitive RNA (**Figure 2A**). This RNA is the transcription product of the homonymous C0T-1 DNA, mostly known by researchers for its background suppression role in *in situ* DNA hybridization techniques and microarray screening. In fact, this DNA suppresses the spurious binding of ubiquitous repetitive sequences. Its name derives from a time parameter of re-hybridization experiments which originally led to its identification in the 70s (Holmes and Bonner, 1974). Hall's remarkable findings imply that the nuclear RNA hybridizing to C0T-1 DNA is almost exclusively associated with euchromatin, that is species-specific (no cross reactivity between mouse and human euchromatin caRNA) and it "spreads" onto chromosomes much like Xist (Hall et al., 2014). This RNA appeared enriched in L1 and to a lesser extent, *Alu* repeats. The L1 component was revealed to be 5' truncated, consistent with the knowledge that full-length L1 transposons are silenced in mammalian genomes (Rangwala et al., 2009).

The significance of finding a non-functional fragment of a transposable element as the most abundant component of euchromatin associated RNA is unclear (Hall et al., 2014). It is tempting to speculate that the 3' L1 fragment retains the parent transposon's capacity to somehow decondense DNA in order to effect its migration, but without the actual ability to integrate, and that this feature may have been exploited by the cell to maintain the euchromatin state (Jachowicz et al., 2017). In terms of mechanism, it has been shown that L1 RNA achieves its DNA decondensing effect by direct interaction with histone H2B, pointing at the electrostatics interactions of caRNA to histone binding as the rationale for the eu- to heterochromatin transition (Dueva et al., 2019).

## MECHANISMS OF CHROMATIN REGULATION BY REPETITIVE RNAs AND RNA BINDING PROTEINS

Having seen how caRNAs in general, and repetitive caRNAs in particular, play a decisive role in the regulation of chromatin states and, by extension, in gene expression, the question arises of what their mechanism of this action is. Given the high complexity of the nuclear microenvironment it can be predicted



that the condensation state of chromatin will be a result of several interplaying components, mainly RNA and proteins, but also higher order structures such as Topological Associated Domains (TAD), Lamin Associated domains (LADs) (Collas et al., 2019) and phase separated spatial compartments within the nucleoplasm (Nozawa and Gilbert, 2019) (discussed below). In the following paragraphs, we will give a brief overview of the current state of research in the field, with a particular focus on critical selected examples of RNA binding proteins. For some extensive reviews on the roles of RNA binding proteins in nuclear and cellular functions, we refer the reader to the following articles (Díaz-Muñoz and Turner, 2018; Hentze et al., 2018; Gebauer et al., 2021).

## Scaffold Attachment Factors (SAF-A and B)

Two of the major protein partners of RNA in the regulation of chromatin state are the SAF (Scaffold Associated Factor) proteins, SAF-A and B (Nozawa et al., 2017; **Figure 2B**). SAF-A is a monomeric protein able to polymerize in an

ATP-dependent fashion and to bind RNA through its RGG domain, in what appears to be a sequence independent way. Its association with RNA yields what has been defined as a “mesh” which contributes to maintaining the euchromatic state (Nozawa et al., 2017). Moreover, it can also bind A/T-rich double-stranded DNA sequences, known as scaffold attachment regions (SARs) (Xiao et al., 2012). This feature probably explains its ability to “bridge” the two nucleic acids to form the mesh. SAF-A (also known as HNRNPU) depletion in mouse hepatocytes results in an enhanced condensation and aberrant lamin association of DNA (Fan et al., 2018). The authors also observe its association with euchromatin. However, it would be inaccurate to draw a direct correlation between SAF-A and euchromatic DNA as the protein is also abundantly present on the inactivated X chromosome and a necessary partner of *Xist* RNA (McHugh et al., 2015). It appears that SAF-A is thus an RNA co-factor able to carry out different functions depending on the RNA (or RNA repeats) it associates with. Of note, the Xi associated SAF-A may be post translationally modified as it is not recognized by some antibodies (Nakagawa and Prasanth, 2011).

SAF-B is a cognate protein of SAF-A, sharing with its SAP (serum amyloid P domain, a DNA-binding motif allowing direct binding to the matrix scaffold attachment regions) and RGG domains. It likewise binds RNA to modulate chromatin condensation. However, its role appears to be distinctly repressive (Huo et al., 2020; **Figure 2B**). It is able to bind repetitive pericentromeric ncRNA and its depletion causes the disappearance of H3K9Me3 foci, while at the same time leaving the total amount of H3K9Me3 unchanged. The authors hence suggest the protein may act as a trigger for phase separation, which we will cover later in our review.

## lncRNA Control of Nuclear Localization: Matrin 3 and CIZ1

The functions of proteins binding to the *Xist* E repeats have been recently studied in order to understand lncRNA spatial localization and function within the nucleus (Ridings-Figueroa et al., 2017; Pandya-Jones et al., 2020). Ridings-Figueroa et al. (2017) showed that CIZ1 interaction with *Xist* RNA allows proper anchoring of the Xi to the nuclear matrix thanks to its nuclear matrix bound C-terminus. Pandya-Jones et al. (2020) have identified several binding partners of *Xist* (Matrin 3, PTBP1, TDP43 and CELF1) and pinpointed their binding site to the E repeats, which would then act as a multimerization scaffold and as a seed for protein aggregation and condensation. Once formed, the protein condensate can then survive *Xist* removal and could explain the eventual dispensability of *Xist* in maintaining the Xi state (Csankovszki et al., 2001). Noticeably, Matrin 3 has also been associated to the nuclear retention and localization of *Charme* lncRNA, potentially suggesting a wider role of this protein in lncRNA/RNA nuclear retention (Desideri et al., 2020).

## RNA Degradation Machinery and Alu and B2 in Stress Response

The nucleus sees a constant turnover of RNA, mostly because of mRNA splicing and intron generation. Introns have a much shorter half-life than lncRNAs, and their degradation is mediated by ribonucleases such as XRN2 and the exosome complex (Nozawa and Gilbert, 2019). Other lncRNAs have a longer half-life of several hours, comparable to that of mRNAs, possibly because of their polyadenylation. The possible exceptions are the repetitive RNAs arising from constitutive heterochromatin (see above), whose transcription at a basal level is necessary to promote local HC condensation but whose overexpression is actually detrimental to the maintenance of the HC state, as shown in *Drosophila melanogaster* by Eberle et al. (2015). RNA degradation is therefore required at least for the heterochromatic silencing of the centromeric regions, which bind exosome subunits (Oya et al., 2013).

As another example highlighting the complexity of the repetitive RNA/chromatin relationship, we can quote the intriguing mechanism of the mammalian genetic stress response, as elucidated by Hernandez et al. (2020). To orchestrate the response to cellular stress, thermal or Estrogen Receptor (ER), a number of previously silenced genes are rapidly reactivated. The authors demonstrate that this silencing is actually mediated by

B2 retrotransposons binding to target genes, and that the stress signal recruits EZH2, a subunit of the PRC2 complex, to the relevant loci (**Figure 2C**). This subunit promotes cleavage of the SINE B2 RNA, its release from chromatin and the transactivation of the stress genes. A similar behavior was observed with the other SINE, *Alu*. The phenomenon is interesting because it does not only show a clear exaptation of the “parasitic” SINEs, but also uncovers a novel activity (RNase enhancer) of EZH2, previously encountered as a histone methyltransferase. Moreover, in this context, its activity is not a transcriptional repressor but an activating protein (Hernandez et al., 2020).

## RNA-DNA Triplex

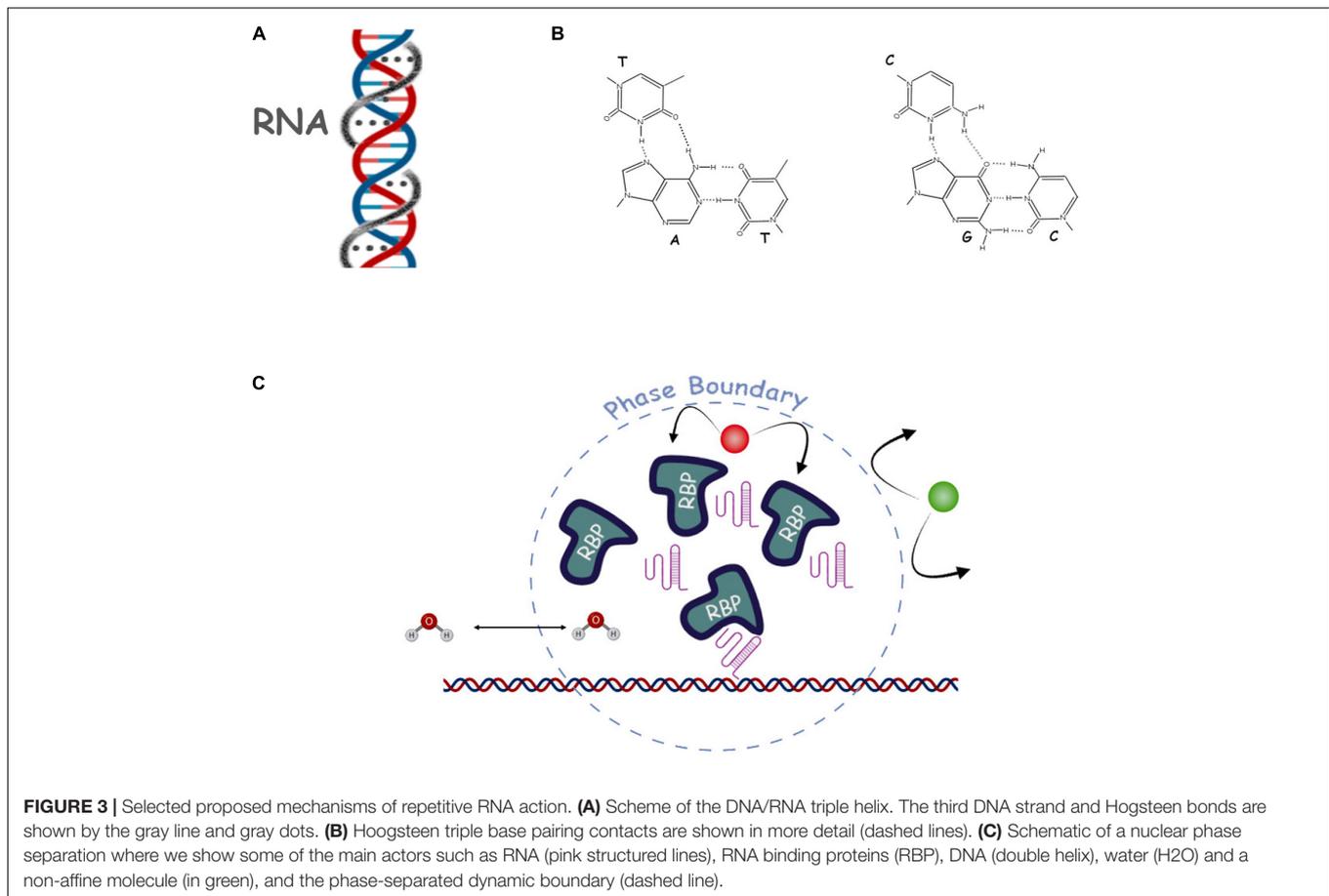
Soon after the classical DNA double helix structure was elucidated (Watson and Crick, 1953, 1974), speculation arose that it may not be the only possible stable structure and that, in particular, a triple helix with a single strand of DNA or RNA may be formed with the third strand inserting into the major groove and forming additional hydrogen bonds with DNA base pairs called Hoogsteen base pairs (Ghosal and Muniyappa, 2006). These triplex structures were eventually found *in vitro* (Morgan and Wells, 1968; Li et al., 2016) and later *in vivo* with RNA as the most stable third strand (**Figures 3A,B**). Moreover, the triplexes can have a silencing effect on transcription even *in trans*, possibly by recruiting PRC2, but also activating effects by binding p300/CBP (Li et al., 2016). The RNAs involved in these events are lncRNAs and particular repeating motifs (e.g., AG and TC) or palindromic sequences are thought to be needed to stabilize the interaction. The helix model would elegantly explain the role of some classes of lncRNAs, and especially the role of their sequence, in targeting the regions of DNA to activate or silence chromatin, however, the evidence in this regard is still preliminary (Ghosal and Muniyappa, 2006). One noticeable example, however, is the regulation of TGF- $\beta$  pathway genes by the lncRNA MEG3 (Mondal et al., 2015), which is mediated by triple helix structures. By using ChIP with anti-triplex antibodies, the authors were able to determine enrichment of the triple helix structures at the TGFBR1, TGFB2 and SMAD2 genes which was lost upon removal of the MEG3 RNA.

## REPETITIVE RNAs IN NUCLEAR ARCHITECTURE

### Repetitive RNAs and Membrane-Less Organelles

One of the nucleus' most remarkable features is the presence of membrane-less organelles, that are regions of distinct composition to the surrounding nucleoplasm but devoid of material barriers, such as phospholipid bilayers, to enforce the separation (Hyman et al., 2014). Examples are the nucleoli, the Cajal bodies, and paraspeckles. In this review, we will focus on the role of RNA, and in particular repetitive RNA in the formation of membrane-less nuclear compartments (**Figure 3C**).

It is thought that membrane-less compartments can exist thanks to a physico-chemical phenomenon known as phase



separation, by which regions of a solution spontaneously split from the rest to form a distinct phase with different physical properties and chemical composition (Hyman et al., 2014). While such events are commonly observed in the presence of different solvents, they are a rarer occurrence in biological solutions (Titus et al., 2020). There are three main classes of phase separation: Liquid-Liquid (LLPS), Liquid-Gel (LGPT) and Polymer-Polymer (PPPT), depending on the type of solutes included in the different phases (Frank and Rippe, 2020). It is believed that the presence of high local concentrations of specific RNA, acting as a seed for condensation, and proteins can cause phase separation and the formation of membrane-less compartments. Different local chemical composition of these condensates would allow for differential chromatin transcriptional states and different rates of transcription, RNA processing and splicing (Huo et al., 2020). It is also thought that repetitive RNAs are key to the occurrence of this phenomenon in the nucleus, as their repeated motifs could work as multivalent ligands for RBP which in turn could result in a non-covalent RNA-protein “polymer” scaffold, initiating, either a LLPS or a LGPS process. The hallmarks of LLPS are: (i) a roughly spherical conformation (droplet); (ii) the ability to fuse and split (fusion and fission); (iii) concentration dependent aggregation and disaggregation and (iv) the possibility of reversible disruption by specific chemicals (i.e., hexanediol) (Sawyer et al., 2019).

The most studied membrane-less compartment, the nucleolus, is a large organelle that also has a role in genome organization and clusters heterochromatin at its periphery in the form of nucleolus-associated domains (NADs) in all somatic cells (Németh et al., 2010; van Koningsbruggen et al., 2010). NADs are, therefore, heterochromatic regions of low gene density and low gene expression (Vertii et al., 2019), and centromeres and telomeres often associate to nucleoli (Carvalho et al., 2001; Weierich et al., 2003; Zhang et al., 2004). For example, the inactive X chromosome was found to contact the perinucleolar compartment during mid/late-S-phase and it was suggested that this location could be important for faithful duplication of silent chromatin (Zhang et al., 2007). The nucleolus is formed by ribosomal DNA (rDNA) and RNA (rRNA) and RBPs (e.g., nucleophosmin, fibrillarin etc.). It collapses into an irregular structure upon inhibition of transcription or by depletion of RNA, suggesting that its integrity is not only dependent on the presence of RNA/repetitive RNA but also on RNA production (Boisvert et al., 2007).

Under heat shock and acidosis conditions, the expression of ribosomal intergenic spacer long non-coding RNA (IGS lncRNA) is induced and these transcripts are essential and sufficient for the immobilization of proteins that contain a nucleolar detention signal (NoDS) within the nucleolus. The mature IGS transcripts then tether the molecular complexes to

their sites of expression on the rDNA cassette. These ncRNAs contribute to dissolving the characteristic tripartite organization of the nucleolus (Granular Compartment, Fibrillar Centers and Dense Fibrillar Components), giving place to the “protein detention center” (DC), which is spatially, dynamically, and biochemically distinct. Upon removal of the environmental stressor, the ncRNAs are repressed, the DC is dissolved, and tripartite nucleolar organization is re-established (Jacob et al., 2013). This form of nucleolar detention could also be considered a regulated posttranslational regulatory mechanism.

As mentioned before, the nucleolus loses its integrity when Pol II is inhibited, which is remarkable given the high rRNA content of the organelle. It is thought that Pol II provides stabilizing caRNA in the form of *Alu* repeats proceeding from intron splicing events in transcription hubs (Caudron-Herger et al., 2015). Transposon-associated ncRNAs represent one of the best examples of how transposable elements-derived ncRNAs (TE) modulate the spatial organization of nucleolus/genome. *Alu* RNA repeats interact with the nucleolin protein and thus contribute to the maintenance of nucleolar structure and function. Interestingly, *Alu* RNAs can target other genomic loci to the nucleolus suggesting that these ncRNAs may impact spatial genome organization by establishing physical links within and outside of the nucleolus (Caudron-Herger et al., 2015).

A recent study (Singh et al., 2018) shows how ncRNAs can help organize the nucleolus when associating with other forms of RNA and proteins. MiCEE (Mirlet7d, C1D, EXOSC10 EXOSC5 complex) was described as a ribonucleoprotein complex that mediates epigenetic silencing of bidirectionally expressed genes and is required for proper nucleolar organization. MiCEE acts by tethering the regulated genes to the perinucleolar region, inducing ncRNA degradation and transcriptional silencing. Specifically, the microRNA Mirlet7d forms a duplex with ncRNA/repeat-containing ncRNA expressed from bidirectionally transcribed genes and associates to C1D protein. C1D, in turn, targets the RNA exosome complex and the polycomb repressive complex 2 (PRC2) to the bidirectionally active loci. The exosome degrades the ncRNAs, whereas PRC2 induces heterochromatin and transcriptional silencing through EZH2 (Singh et al., 2018).

## Other Examples of Membrane-Less Sub-Nuclear Compartments

Early studies identified a role for molecular crowding in the formation of some membrane-less nuclear compartments (Richter et al., 2007; Cho and Kim, 2012), through liquid-liquid phase separation (Zhu and Brangwynne, 2015; Hall et al., 2019).

Nuclear RNAs and in particular, lncRNAs, have been shown to be involved in the formation of these sub-nuclear structures (Khosraviani et al., 2019). This is because lncRNAs exhibit properties (such as secondary structures) and repetitive elements that make them potential candidates for acting as architectural elements for chromatin organization and in this role are labeled architectural RNA (arcRNA). RNA forms secondary structures which interact with specific proteins and other RNA molecules. A single lncRNA can act as an RNA scaffold either by interacting with multiple copies of the same protein or

several different proteins at once, representing the ideal seed molecule for condensation seeds. *Neat1* and *Malat1* and *Xist* (discussed above) are remarkable examples and are amongst the most conserved lncRNAs during vertebrate evolution although containing minimal repetitive elements.

The large isoform of *NEAT1* lncRNA (NEAT1.2) appears to play a critical role in organizing a type of nuclear compartment called paraspeckles, containing various mRNAs and RBPs. Targeted degradation of *NEAT1* disrupts the structure of these clusters (Chen and Carmichael, 2010). Moreover, many repeat-containing RNAs have been shown to associate with paraspeckles, suggesting that the domain might arise from clustering some specific classes of ncRNAs along with their RBPs (Prasanth et al., 2005). From a functional point of view, paraspeckles are dynamically “designed” to retain inside the nucleus certain mRNAs that had been subjected to high levels of adenosine-to-inosine editing and to concentrate certain RBPs to limit their functions in the nucleus.

*MALAT1* lncRNA localizes to compartments called nuclear speckles containing various splicing, RNA-processing and transcription factors, and that are thought to function as a storage for RNA-processing proteins when they are not actively engaged. Interestingly, *MALAT1* associates with actively transcribed genes in the periphery of nuclear speckles and dozens of RBPs, and these findings suggest that it could act as a scaffold mediating those interactions.

The pericentromeric-derived ncRNA, *HSATIII* arcRNA, leads to the formation of membrane-less nuclear compartments known as nuclear stress bodies. *HSATIII* arcRNA consists mainly of highly repetitive (GGAAU)<sub>n</sub> sequences (Valgardsdottir et al., 2008) and is transcribed from the primate-specific pericentromeric satellite III regions under thermal stress conditions (Denegri et al., 2002; Jolly et al., 2004; Valgardsdottir et al., 2008). *HSATIII* arcRNAs remain stable in nuclei, but form membrane-less nuclear stress bodies (nSBs) upon recruitment of specific RNA-binding proteins such as Scaffold attachment factor B (SAFB), specific sets of SRSFs (SRSF1 and 9 during thermal stress), transcription factors HSF1 and CREBBP, bromodomain protein BRD4 (Jolly et al., 2004; Metz et al., 2004; Kawaguchi et al., 2015; Hussong et al., 2017) and many nuclear RBPs involved in pre-mRNA splicing and processing (Ninomiya et al., 2020). Similarly, *Hsr omega* (heat-shock RNA-omega) arcRNA from *Drosophila melanogaster* contains tandem repeats of 280 nt in a stretch of ~10 kb that contribute to the recruitment of various RNA-binding proteins to omega speckles (thermal stress-induced nuclear bodies) (Prasanth et al., 2005; Singh and Lakhota, 2015).

An excellent recent review summarizing the current knowledge about cell organization and membrane-less compartments was published recently by Quinodoz and Guttman (2021). We refer the author to this review for more details on these topics.

## Repetitive RNA in the Formation and Maintenance of Nuclear Domains

It has been suggested that nuclear RNAs are an essential component of interphase chromosomes (Hall and Lawrence,

2016) but it remains unclear which of these may have roles in shaping large-scale chromatin structure and regulating genome function. An example suggesting RNA-based roles in nuclear architecture is how the digestion of RNA, but not of proteins, resulted in a highly disorganized nucleus, as assessed by electron microscopy, and also mislocalization of chromatin regulatory complexes (Davidson and Britten, 1979; Nickerson et al., 1997; Bernstein et al., 2006; Britten, 2010). Also, the removal of RNA leads to the collapse of nuclear bodies providing clear evidence for the architectural role of RNA in the formation of these structures (Bond and Fox, 2009; Mao et al., 2011; Shevtsov and Dundr, 2011).

Regarding spatial genome organization, non-coding RNAs (ncRNAs) have particularly emerged as major regulators and can regulate transcription at the same locus (*in cis*) from where they are transcribed or elsewhere in the genome (*in trans*) (Khosraviani et al., 2019). Collectively, ncRNAs can impact genome organization by modulating perinuclear chromosome tethering, the formation of major nuclear compartments, chromatin looping, and various other chromosomal structures (Khosraviani et al., 2019). Remarkably, several ncRNAs from repetitive DNA loci have emerged as major players that mediate crosstalk between spatial genome organization, expression, and stability (Caudron-Herger et al., 2015). Considering that highly repetitive sequences (i.e., fragments of TEs) are found in most transcripts (including unspliced mRNA) and that they comprise up to 2/3 of the human genome, they might contribute significantly to the abovementioned nuclear roles. We also refer the reader to other reviews on ncRNAs, and particularly repetitive nuclear ncRNAs, mostly emerging from rDNA repeats, telomeric regions, transposable elements, and centromeres (Hall and Lawrence, 2016; van Steensel and Furlong, 2019).

## Role of ncRNAs in Lamin Associated Domains Generation and Maintenance

In mammalian cells, the nuclear lamina is thought to be the key organizer of the radial arrangement of chromatin in interphase nuclei, by creating a large nuclear compartment where most of the inactive chromatin clusters in the form of lamina-associated domains (LADs) (Peric-Hupkes et al., 2010; Kind et al., 2015; **Figure 4A**). LADs are typically 0.1–10 megabases, gene-poor, enriched in heterochromatin, and display low gene activity (Guelen et al., 2008; Lund et al., 2014; Khosraviani et al., 2019). The association of chromatin with the nuclear lamina through LADs aids functional organization of the genome and enables a spatio-temporal regulation of replication and transcription (Buchwalter et al., 2019).

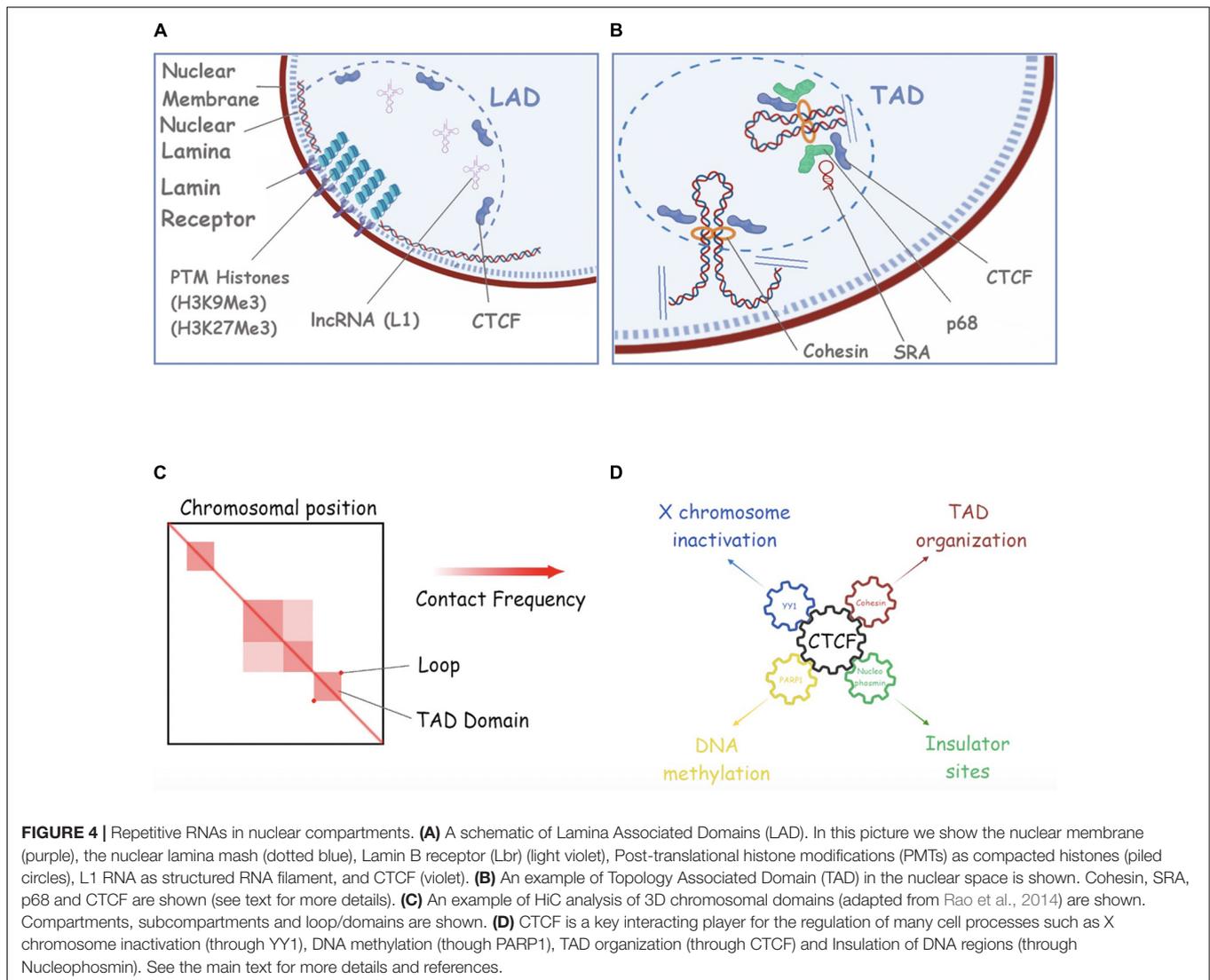
Several studies might suggest that ncRNAs and especially, repetitive ncRNAs play a critical role in anchoring specific loci to the nuclear lamina, organizing genome architecture and regulating gene expression. *Xist* lncRNA is the clearest example, and its interaction with lamin B receptor (LBR) is required for tethering the entire X-chromosome to the nuclear lamina during XCI (Chen et al., 2017; Young et al., 2021). Apart from *Xist*

RNA, there are a few examples of ncRNAs possibly regulating the localization of specific genetic loci to the nuclear lamina, and hence affecting gene expression. As an example, the *L1 repeat* RNA plays a role in sequestering L1-rich sequences and associates genes in inactive domains for silencing. Depletion of *L1* RNA in embryonic stem cells (ESCs) leads to the relocation of L1-enriched chromosomal segments from inactive domains to the nuclear interior and derepression of L1-associated genes. Moreover, *L1* RNA seems to facilitate the interaction of L1 DNA to proteins such as nucleolin to target it to NADs/LADs in mouse ESCs (Lu et al., 2020b).

Recent studies have highlighted the role of telomeric-derived ncRNAs in the crosstalk between telomeric heterochromatin and the subnuclear positioning of telomeres. This is the case for PIWI-interacting RNAs (piRNAs), which are typically transcribed from telomeric regions and play a pivotal role in the establishment and maintenance of telomeric chromatin in the germline, facilitating loading of HP1 and H3K9me3 at different telomeric regions. In the fly germline, the loss of some piRNAs decreased perinuclear telomere positioning and lowered the local enrichment of HP1 and H3K9me3, resulting in telomere dysfunction (Radion et al., 2018).

Similarly, transposable elements (TEs)-derived ncRNAs might play a role in chromatin compartmentalization. Enhancer lncRNAs (elncRNAs), often originating from repetitive sequences/fragments derived from transposable elements (Rebollo et al., 2012; Su et al., 2014) may reinforce the interactions between enhancers and promoters (Hou et al., 2019) as reported during developmental progression and tumor suppression. For example, in developing T cells, *ThymoD* ncRNA transcription promoted demethylation at CTCF bound sites and activated Cohesin-dependent looping to reposition the *Bcl11b* enhancer from the lamina to the nuclear interior and to juxtapose the *Bcl11b* enhancer and promoter into a single-loop domain. As expected, these large-scale changes in nuclear architecture were associated with the deposition of activating epigenetic marks across the loop domain (Isoda et al., 2017).

Sophisticated techniques are being developed to analyze nuclear DNA architecture with increasing accuracy and minimal interference. For example, in 2020, Girelli et al. (2020) developed a method called GPseq that allowed generating the first high-resolution map of radial chromatin organization in human cells. To infer radial locations throughout the genome, GPseq is used in combination with a special FISH assay called YFISH, which allows monitoring the pattern of *in situ* digestion before sequencing the GPseq sample. Specifically, YFISH uses a Y-shaped adapter that will ligate the cuts introduced *in situ* by a restriction enzyme, and then the use of complementary fluorescently labeled oligos will help with the detection. In 2016 Brant et al. (2016) developed i3C as a chromosome conformation capture technique aimed at minimizing the *in vivo* interference of formaldehyde crosslinking (Brant et al., 2016). In order to stimulate future work in this direction, we suggest the use of these or similar techniques, in conjunction with existing LAD mapping techniques at population or single cell level (Kind et al., 2015) in different conditions to investigate the contribution of repetitive RNAs to LADs formation. For example, comparing



the chromatin distribution in the presence and in the absence of a specific ncRNA or a specific class of ncRNA.

## Role of ncRNAs in Topologically Associated Domains Generation and Maintenance

At a large scale, chromosomes segregate into regions that form two different types of chromatin, referred to as “A-type” (euchromatin) and “B-type” (heterochromatin) genomic compartments (Lieberman-Aiden et al., 2009). At the sub-megabase scale, chromosomes form a series of architectural chromatin units termed topologically associated domains (TADs), each of which includes hundreds of kilobases of DNA (Dekker and Heard, 2015; **Figures 4B,C**). TADs are separated by regions known as TAD boundaries and exhibit a higher frequency of intradomain interactions compared to interdomain interactions (Rao et al., 2014; Dixon et al., 2015). The organization of the genome into TADs is critical

for coordinated transcriptional regulation, chromatin states, and DNA replication. These structures dynamically change during differentiation (Dixon et al., 2015) and are perturbed in disease (Barutcu et al., 2015). Over the years, multiple studies have described the role of different types of ncRNAs in TADs formation and maintenance (Fudenberg et al., 2017; Amaral et al., 2018). Generally, the mammalian genome is arranged into compartments of active and inactive chromatin (Lieberman-Aiden et al., 2009). Moreover, linearly non-contiguous TADs can contact each other, defining long-range interactions that can vary between cell types and during differentiation (Szabo et al., 2018; Connelly et al., 2019; Paulsen et al., 2019).

Traditionally, Cohesin, loop extrusion, and CCCTC-binding factor (CTCF) have been proposed to create TADs to regulate gene expression. TADs boundaries are enriched for the insulator binding protein CTCF, housekeeping genes, transfer RNAs, and short interspersed element (SINE) retrotransposons, indicating that these factors may have a role in establishing the topological domain structure of the genome (Liyakat Ali et al., 2021). Cohesin

can utilize its ATPase activity to extrude loops of chromatin and this is limited by CTCF-occupied insulator DNA elements (Sanborn et al., 2015; Fudenberg et al., 2017). This process is thought to structure and insulate chromosomes, limiting the effects of distal gene regulatory elements to genes within a given TAD. Importantly, altering TAD boundaries can lead to ectopic contacts between cis-regulating elements and gene promoters, and thus gene misexpression, which can contribute to developmental defects and cancer (Akdemir et al., 2020). However, there is recent evidence suggesting that the weakening of TAD boundaries also appears to be independent of CTCF binding and, transcription would have supporting roles in the formation of TADs and regulation of *inter-TAD* interactions (Barutcu et al., 2019). In this study, Barutcu et al. (2019) showed that the strength of TAD boundaries, measured by the degree of interactions that occur across a TAD boundary, is significantly decreased upon transcriptional inhibition, suggesting a role for total, steady-state single-stranded RNA on genome architecture. This finding is consistent with others where transcriptional inhibition, as well as transcriptional elongation can displace Cohesin from CTCF sites and disrupt chromatin interactions (Li et al., 2015; Heinz et al., 2018; Vian et al., 2018; Rowley et al., 2019), a phenomenon that correlates with the weakening of TAD boundaries. This evidence do not, however, exclude other possibilities (i.e., transcription inhibition weakens transcriptional condensates). Another supporting fact for the potential role of active transcription in the topological organization of the genome is the enrichment of transcription-associated RNAs with TAD boundaries (Bell et al., 2018) as well as the observation that active transcription is a stronger predictor for TAD partitioning in flies (Ulianov et al., 2016; Hou et al., 2019) than CTCF and Cohesin accumulation, the prototypical TAD boundary markers in mice and humans (Merkenschlager and Nora, 2016). As already speculated, pre-existing and newly transcribed RNA (which contains repeats) might play a role in genomic compartmentalization (Erdel and Rippe, 2018) and as Erdel and Rippe described, after RNase A treatment and before cross-linking, there is a subtle perturbation of compartmental interactions, especially in B-type compartments (Erdel and Rippe, 2018).

There are also a few studies supporting the role of ncRNAs in facilitating long-distance chromatin interactions through Cohesin-binding. In mammalian cells, for instance, not all the CTCF sites are co-occupied by Cohesin, suggesting that additional factors could dictate Cohesin binding at CTCF sites (Zlatanova and Caiafa, 2009). In fact, several partners of CTCF have been identified (e.g., YY1, Nucleophosmin, PARP), each associated with a particular and distinct function of the protein (Figure 4D). The interaction between Cohesin and CTCF is modulated by the DEAD-box RNA binding protein p68, together with its associated ncRNA called steroid receptor RNA activator (SRA), and promotes insulator function, for example, at the *Igf2/H19* locus. Additionally, Cohesin was reported to bind to the ncRNAs transcribed on enhancer regions, termed enhancer RNAs (eRNAs) (Racko et al., 2018). For instance, the eRNAs bind to Cohesin and increase its recruitment to the enhancer regions in response to the ER ligand estradiol, stimulating the

enhancer-promoter interactions in MCF7 breast cancer cells. Furthermore, key components of the Cohesin complex, SA1 and SA2, bind to various RNA containing substrates, including ssRNA, dsRNA, RNA: DNA hybrids and R-loops and it has been shown that both SA1 and SA2 Cohesin subunits localize to regions on dsDNA that contain RNA (Pan et al., 2020). Another example in this category is the *blncRNA1* (boundary associated long non-coding RNA-1), generated from the CBS5 boundary element (*HOXA* locus) promoter activity. CBS5 employs both Cohesin and *blncRNA1* to establish and maintain TADs at the *HOXA* locus and the transcript promotes proper expression of *HOXA* genes (Nwigwe et al., 2015).

Likewise, there are examples of ncRNAs facilitating long-distance chromatin interactions through CTCF binding. CTCF is found to interact with a multitude of transcripts genome-wide, both protein-coding mRNA and non-coding transcript, mRNAs as well as many long-non-coding RNA (lncRNA) such as well-characterized species from imprinted loci and previously unannotated transcripts from intergenic space (Kung et al., 2015). CTCF is recruited in a locus-specific manner and implicates CTCF-RNA interactions in long-range chromosomal interactions. For example, *Tsix* and *Xist* RNAs target CTCF to the X-inactivation centre, thereby facilitating homologous X-chromosome pairing (Kung et al., 2015). In accordance with this, it is not surprising that mutation of the RNA-binding regions in CTCF (ZF1 and ZF10) disrupts gene expression, chromatin binding, and the formation of chromatin loops (Saldaña-Meyer et al., 2019). Moreover, transcription inhibition disrupts CTCF binding to chromatin (Saldaña-Meyer et al., 2019). Interestingly, there are positionally conserved RNAs linked to chromatin organization structures called topological anchor point RNAs (tapRNAs). These tapRNAs overlap binding sites for the CTCF chromatin organizer and localize at chromatin loop anchor points and borders of TADs. Characterization of these ncRNAs and their associated coding genes shows that they regulate each other's expression and influence the metastatic phenotype of cancer cells *in vitro* in a similar fashion (Amaral et al., 2018). Remarkably, and in relation to CTCF, a very recent study based on the CTCF CUT and RUN technique revealed that intact RNA, of unknown nature, is required for maintaining the chromatin environment around CTCF likely by facilitating local chromatin compaction (Thakur et al., 2019).

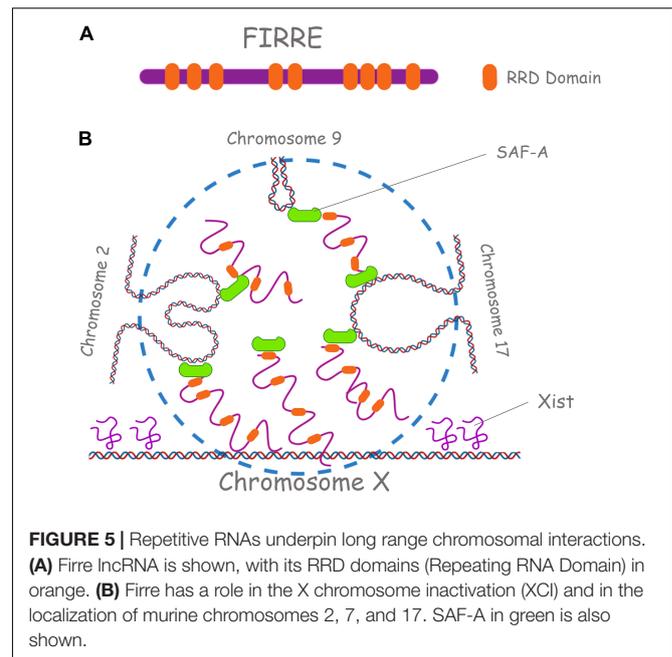
TEs-derived ncRNAs have been frequently described as involved in the establishment and maintenance of insulator boundaries between TADs. Transposable elements (TEs) are responsible for genomic instability, epigenetic silencing and are intrinsically linked to 3D organization as several studies described how they shape genome organization from demarcating TAD boundaries to harboring binding sites for architectural proteins (Diehl et al., 2020). Mammalian-wide interspersed repeats (MIRs) are a conserved family of TEs that have a substantial regulatory capacity and share sequence characteristics with tRNA-related insulators. MIR insulators appear to be CTCF independent and show a distinct local chromatin environment with marked peaks for RNA Pol III and several histone modifications. This suggests that MIR insulators recruit transcriptional complexes and chromatin

modifying enzymes *in situ* to help establish chromatin and regulatory domains in the human genome (Wang et al., 2015). The primate-specific endogenous retrotransposon human endogenous retrovirus subfamily H (HERV-H) RNA represents another example and has a role in creating TADs in hPSCs and PSCs from other species and this ability depends on abundant transcription, as transcriptional repression of HERV-H elements prevents the formation of boundaries (Zhang et al., 2019). The Murine Endogenous Retroviral Element (MuERV-L/MERVL) family of transposable elements drives the 3D reorganization of the genome in the early mouse embryo by promoting the formation of insulating domain boundaries throughout the genome. The formation of these boundaries is coupled to the upregulation of directional transcription from MERVL, which results in the activation of a subset of the gene expression program of the 2-cell stage embryo (Kruse et al., 2019).

In spite of the previously mentioned findings, recent studies suggest that the pool of RNA in the cell appears to be largely dispensable for the maintenance of TADs (Fudenberg et al., 2017; Tan et al., 2017; Amaral et al., 2018). As the authors stated, while pre-existing transcribed RNA may play a role at small local scales or mediate inter-chromosomal interactions (Maass et al., 2012; Hacisuleyman et al., 2014), overall it does not appear to significantly influence TAD boundary formation. The fact that TAD boundaries remain intact in cells treated with RNase A, either before or after formaldehyde crosslinking, is consistent with a model explaining that TAD formation is primarily driven by DNA-protein and protein-protein interactions rather than by RNA (Fudenberg et al., 2017). Considering these contradictory results, more studies and accurate techniques need to be developed, in order to discriminate direct vs. secondary effects of RNA depletion.

## Repetitive RNAs Involved in Trans-Chromosomal Interactions

Linearly non-contiguous TADs can also contact each other, defining long-range interactions that can vary between cell types and during differentiation (Quinodoz et al., 2018; Szabo et al., 2018; Paulsen et al., 2019). For instance, repetitive and repeat-containing lncRNAs such as *XIST* and *FIRRE* (*functional intergenic repeating RNA element*), colocalize with Xi and determine trans-chromosomal interactions (Figure 5). It is suggested that during XCI, *Xist* would facilitate the atypical TAD structure of the Xi into two “mega-domains” around the DZX4 locus (Bonora et al., 2018). These mega-domains, differently from those on the active X chromosome (Xa), exhibit random/semi-random pattern of interactions compared to the punctate interactions between specific loci seen on most chromosomes. With regards to *FIRRE lncRNA*, it interacts with SAF-A through its RRD (Repeating RNA Domain), a 156-bp repeating sequence (Hacisuleyman et al., 2014), contributing to the organization of higher-order chromosome architecture to spatially coordinate the regulation of genes involved in the same biological process (e.g., adipogenesis). This conserved and unique motif is necessary to localize *FIRRE* around its site of transcription in the nucleus but it can also localize any RNA containing it (Hacisuleyman, 2015). In mouse, *FIRRE* forms a punctate compartment in the



nucleus where its locus on the X chromosome and several specific loci on mouse chromosomes 2, 9, 15, and 17 colocalize with it. *FIRRE* is then required for these inter-chromosomal interactions. Both genetic deletion of the *Firre* locus and knockdown of SAF-A resulted in loss of colocalization of these trans-chromosomal interacting loci (Hacisuleyman et al., 2014).

Similarly, TEs may be involved in the establishment of conserved long-range chromosomal interactions in different organisms and some of these interactions appear to be important in gene regulation. For example, in the fungus *Epichloe festucae*, these repeat-rich blocks mediate genome folding within the nucleus and help to divide the genome into distinct regions that have similar gene expression profiles, modulating in this way the expression of genes that are strongly differentially expressed (Winter et al., 2018).

Within the nuclear matrix, novel classes of repetitive RNAs have been identified as major players of the nuclear architecture and associated genome regulation. A class of repeat-containing lncRNA from the AAGAG satellite DNA repeat have emerged as a crucial component of the nuclear architecture in *Drosophila melanogaster* (Lohe and Brutlag, 1986; Lohe et al., 1993; Smith et al., 2007). Reduction of these repetitive RNAs results in disruption of the nucleoskeleton and, consequently, the assembly and stability of the chromosome compartments are disturbed (Pathak et al., 2013). In 2010, Zheng et al. (2010) identified a heterogeneous population of GAA-repeat-containing RNAs (GRC RNAs) that primarily consist of polypurine repeats, ranging from 1.5 to 4 kb. These RNAs are distributed throughout the nucleus in a micropunctate pattern in both primary and transformed human and mouse cell lines. GRC-RNAs associate with the nuclear matrix and interact with several *bona fide* nuclear matrix proteins and have been proposed to play important structural roles in the maintenance of the nuclear and nuclear architecture and regulate gene expression.

## FINAL REMARKS

Early in the human genome project it became clear that most our DNA is made of non-coding, repetitive sequences. Lacking any obvious function and carrying no protein coding information, these vast stretches of the genome were soon labeled “junk DNA” and were thought to be remnants of viruses that had lost replicative competence and merely multiplied inside our DNA, or even aberrant products of DNA replication errors.

It is now becoming more and more clear instead that, far from being genetic “deadwood” these repetitive expanses are actively and deliberately transcribed into non-coding RNAs which play a major role in regulating gene expression and silencing, organizing nuclear architecture, compartmentalizing the nucleus, and modulating protein function. We can now state with confidence that the study of repetitive RNA role and mechanism of action will open a new frontier in cell biology. We hope that our review will further stimulate research in the consolidating field of repetitive RNA biology.

## REFERENCES

- Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhi, R., et al. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nat. Genet.* 52, 294–305. doi: 10.1038/s41588-019-0564-y
- Aldrup-Macdonald, M. E., and Sullivan, B. A. (2014). The past, present, and future of human centromere genomics. *Genes* 5, 33–50. doi: 10.3390/genes5010033
- Amaral, P. P., Leonardi, T., Han, N., Gascoigne, D. K., Zhang, A., Pluchino, S., et al. (2018). Genomic positional conservation identifies topological anchor point (tap) RNAs linked to developmental loci. *Genome Biol.* 19:32. doi: 10.1186/s13059-018-1405-5
- Andresini, O., Rossi, M. N., Matteini, F., Petrai, S., Santini, T., and Maione, R. (2019). The long non-coding RNA Kcnq1ot1 controls maternal p57 expression in muscle cells by promoting H3K27me3 accumulation to an intragenic MyoD-binding region. *Epigenet. Chromatin* 12:8. doi: 10.1186/s13072-019-0253-1
- Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395. doi: 10.1038/cr.2011.22
- Barutcu, A. R., Blencowe, B. J., and Rinn, J. L. (2019). Differential contribution of steady-state RNA and active transcription in chromatin organization. *EMBO Rep.* 20:e48068. doi: 10.15252/embr.201948068
- Barutcu, A. R., Lajoie, B. R., McCord, R. P., Tye, C. E., Hong, D., Messier, T. L., et al. (2015). Chromatin interaction analysis reveals changes in small chromosome and telomere clustering between epithelial and breast cancer cells. *Genome Biol.* 16:214. doi: 10.1186/s13059-015-0768-0
- Bell, J. C., Jukam, D., Teran, N. A., Risca, V. L., Smith, O. K., Johnson, W. L., et al. (2018). Chromatin-associated RNA sequencing (ChAR-seq) maps genome-wide RNA-to-DNA contacts. *eLife* 7:e27024. doi: 10.7554/eLife.27024.037
- Bernstein, E., Duncan, E. M., Masui, O., Gil, J., Heard, E., and Allis, C. D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol. Cell Biol.* 26, 2560–2569. doi: 10.1128/MCB.26.7.2560-2569.2006
- Biscotti, M. A., Olmo, E., and Heslop-Harrison, J. S. P. (2015). Repetitive DNA in eukaryotic genomes. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* 23, 415–420. doi: 10.1007/s10577-015-9499-z
- Blackledge, N. P., Rose, N. R., and Klose, R. J. (2015). Targeting polycomb systems to regulate gene expression: modifications to a complex story. *Nat. Rev. Mol. Cell Biol.* 16, 643–649. doi: 10.1038/nrm4067

## AUTHOR CONTRIBUTIONS

AC conceived the idea of this review. AC, NB, and GT wrote the manuscript and selected representative examples. All authors contributed to the article and approved the submitted version.

## FUNDING

AC is funded by a Rett Syndrome Research Trust (RSRT), BARTS CHARITY grants, and intramural QMUL support.

## ACKNOWLEDGMENTS

We thank Joanna Jachowicz, Hellen Rowe, Remo Sanges, and Liane Fernandes for critical reading and useful suggestions to improve our review. We thank Miss Federica Capraro for her initial help with this review. Figures were realized with the help of the Biorender software. We apologize to all colleagues whose work could not have been cited due to space and focus constraints.

- Boija, A., Klein, I. A., Sabari, B. R., Dall’Agnese, A., Coffey, E. L., Zamudio, A. V., et al. (2018). Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* 175, 1842.e–1855.e. doi: 10.1016/j.cell.2018.10.042
- Boisvert, F.-M., van Koningsbruggen, S., Navascués, J., and Lamond, A. I. (2007). The multifunctional nucleolus. *Nat. Rev. Mol. Cell Biol.* 8, 574–585. doi: 10.1038/nrm2184
- Bond, C. S., and Fox, A. H. (2009). Paraspeckles: nuclear bodies built on long noncoding RNA. *J. Cell Biol.* 186, 637–644. doi: 10.1083/jcb.200906113
- Bonora, G., Deng, X., Fang, H., Ramani, V., Qiu, R., Berletch, J. B., et al. (2018). Orientation-dependent Dxx4 contacts shape the 3D structure of the inactive X chromosome. *Nat. Commun.* 9:1445. doi: 10.1038/s41467-018-03694-y
- Borensztein, M., Syx, L., Ancelin, K., Diabangouaya, P., Picard, C., Liu, T., et al. (2017). Xist-dependent imprinted X inactivation and the early developmental consequences of its failure. *Nat. Struct. Mol. Biol.* 24, 226–233. doi: 10.1038/nsmb.3365
- Brant, L., Georgomanolis, T., Nikolic, M., Brackley, C. A., Kolovos, P., van Ijcken, W., et al. (2016). Exploiting native forces to capture chromosome conformation in mammalian cell nuclei. *Mol. Syst. Biol.* 12:891. doi: 10.15252/msb.20167311
- Britten, R. J. (2010). Transposable element insertions have strongly affected human evolution. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19945–19948. doi: 10.1073/pnas.1014330107
- Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R., et al. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 349, 38–44. doi: 10.1038/349038a0
- Buchwalter, A., Kaneshiro, J. M., and Hetzer, M. W. (2019). Coaching from the sidelines: the nuclear periphery in genome regulation. *Nat. Rev. Genet.* 20, 39–50. doi: 10.1038/s41576-018-0063-5
- Carvalho, C., Pereira, H. M., Ferreira, J., Pina, C., Mendonça, D., Rosa, A. C., et al. (2001). Chromosomal G-dark bands determine the spatial organization of centromeric heterochromatin in the nucleus. Gall J, editor. *Mol. Biol. Cell* 12, 3563–3572. doi: 10.1091/mbc.12.11.3563
- Caudron-Herger, M., Pankert, T., Seiler, J., Németh, A., Voit, R., Grummt, I., et al. (2015). Alu element-containing RNA s maintain nucleolar structure and function. *EMBO J.* 34, 2758–2774. doi: 10.15252/embj.201591458
- Cerase, A., Armaos, A., Neumayer, C., Avner, P., Guttman, M., and Tartaglia, G. G. (2019). Phase separation drives X-chromosome inactivation: a hypothesis. *Nat. Struct. Mol. Biol.* 26, 331–334. doi: 10.1038/s41594-019-0223-0

- Cerese, A., Pintacuda, G., Tattermusch, A., and Avner, P. (2015). Xist localization and function: new insights from multiple levels. *Genome Biol.* 16:166. doi: 10.1186/s13059-015-0733-y
- Cerese, A., and Tartaglia, G. G. (2020). Long non-coding RNA-polycomb intimate rendezvous. *Open Biol.* 10:200126. doi: 10.1098/rsob.200126
- Chadwick, B. P., and Willard, H. F. (2004). Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. *Proc. Natl. Acad. Sci. U.S.A.* 101, 17450–17455. doi: 10.1073/pnas.0408021101
- Chen, C.-K., Blanco, M., Jackson, C., Aznauryan, E., Ollikainen, N., Surka, C., et al. (2016). Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science* 354, 468–472. doi: 10.1126/science.aae0047
- Chen, C.-K., Chow, A., Lai, M., and Guttman, M. (2017). Response to Comment on “Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing”. *Science* 356:eaam5439. doi: 10.1126/science.aam5439
- Chen, L., and Carmichael, G. G. (2010). Long noncoding RNAs in mammalian cells: what, where, and why? *Wiley Interdiscip. Rev. RNA* 1, 2–21. doi: 10.1002/wrna.5
- Cho, E. J., and Kim, J. S. (2012). Crowding effects on the formation and maintenance of nuclear bodies: insights from molecular-dynamics simulations of simple spherical model particles. *Biophys. J.* 103, 424–433. doi: 10.1016/j.bpj.2012.07.007
- Chuong, E. B., Elde, N. C., and Feschotte, C. (2017). Regulatory activities of transposable elements: from conflicts to benefits. *Nat. Rev. Genet.* 18, 71–86. doi: 10.1038/nrg.2016.139
- Cirillo, D., Blanco, M., Armaos, A., Buñes, A., Avner, P., Guttman, M., et al. (2016). Quantitative predictions of protein interactions with long noncoding RNAs. *Nat. Methods* 14, 5–6. doi: 10.1038/nmeth.4100
- Collas, P., Liyakat Ali, T. M., Brunet, A., and Germier, T. (2019). Finding friends in the crowd: three-dimensional cliques of topological genomic domains. *Front. Genet.* 10:602. doi: 10.3389/fgene.2019.00602
- Connelly, K. E., Weaver, T. M., Alpsoy, A., Gu, B. X., Musselman, C. A., and Dykhuizen, E. C. (2019). Engagement of DNA and H3K27me3 by the CBX8 chromodomain drives chromatin association. *Nucleic Acids Res.* 47, 2289–2305. doi: 10.1093/nar/gky1290
- Creamer, K. M., and Lawrence, J. B. (2017). XIST RNA: a window into the broader role of RNA in nuclear chromosome architecture. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 372:20160360. doi: 10.1098/rstb.2016.0360
- Csankovszki, G., Nagy, A., and Jaenisch, R. (2001). Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. *J. Cell Biol.* 153, 773–784. doi: 10.1083/jcb.153.4.773
- Cusanelli, E., and Chartrand, P. (2014). Telomeric noncoding RNA: telomeric repeat-containing RNA in telomere biology. *Wiley Interdiscip. Rev. RNA* 5, 407–419. doi: 10.1002/wrna.1220
- Davidovich, C., and Cech, T. R. (2015). The recruitment of chromatin modifiers by long noncoding RNAs: lessons from PRC2. *RNA (N. Y.)* 21, 2007–2022. doi: 10.1261/rna.053918.115
- Davidson, E., and Britten, R. (1979). Regulation of gene expression: possible role of repetitive sequences. *Science* 204, 1052–1059. doi: 10.1126/science.451548
- de Koning, A. P. J., Gu, W., Castoe, T. A., Batzer, M. A., and Pollock, D. D. (2011). Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* 7:e1002384. doi: 10.1371/journal.pgen.1002384
- DeChiara, T. M., Robertson, E. J., and Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* 64, 849–859. doi: 10.1016/0092-8674(91)90513-X
- Dekker, J., and Heard, E. (2015). Structural and functional diversity of topologically associating domains. *FEBS Lett.* 589(20 Pt A), 2877–2884. doi: 10.1016/j.febslet.2015.08.044
- Denegri, M., Moralli, D., Rocchi, M., Biggiogera, M., Raimondi, E., Cobianchi, F., et al. (2002). Human chromosomes 9, 12, and 15 contain the nucleation sites of stress-induced nuclear bodies. *Mol. Biol. Cell* 13, 2069–2079. doi: 10.1091/mbc.01-12-0569
- Deng, Z., Nørseen, J., Wiedmer, A., Riethman, H., and Lieberman, P. M. (2009). TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Mol. Cell* 35, 403–413. doi: 10.1016/j.molcel.2009.06.025
- Desideri, F., Cipriano, A., Petrezselyova, S., Buonaiuto, G., Santini, T., Kasperek, P., et al. (2020). Intronic determinants coordinate lncRNA nuclear activity through the interaction with MATR3 and PTBP1. *Cell Rep.* 33:108548. doi: 10.1016/j.celrep.2020.108548
- Díaz-Muñoz, M. D., and Turner, M. (2018). Uncovering the role of RNA-binding proteins in gene expression in the immune system. *Front. Immunol.* 9:1094. doi: 10.3389/fimmu.2018.01094
- Diehl, A. G., Ouyang, N., and Boyle, A. P. (2020). Transposable elements contribute to cell and species-specific chromatin looping and gene regulation in mammalian genomes. *Nat. Commun.* 11:1796. doi: 10.1038/s41467-020-15520-5
- Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., He, J., Lee, A. Y., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature* 518, 331–336. doi: 10.1038/nature14222
- Dueva, R., Akopyan, K., Pederiva, C., Trevisan, D., Dhanjal, S., Lindqvist, A., et al. (2019). Neutralization of the positive charges on histone tails by RNA promotes an open chromatin structure. *Cell Chem. Biol.* 26, 1436.e–1449.e. doi: 10.1016/j.chembiol.2019.08.002
- Duszczak, M. M., Wutz, A., Rybin, V., and Sattler, M. (2011). The Xist RNA A-repeat comprises a novel AUCG tetraloop fold and a platform for multimerization. *RNA* 17, 1973–1982. doi: 10.1261/rna.2747411
- Eberle, A. B., Jordán-Pla, A., Gañez-Zapater, A., Hesse, V., Silberberg, G., von Euler, A., et al. (2015). An interaction between RRP6 and SU(VAR)3-9 targets RRP6 to heterochromatin and contributes to heterochromatin maintenance in *Drosophila melanogaster*. *PLoS Genet.* 11:e1005523. doi: 10.1371/journal.pgen.1005523
- Engreitz, J. M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., et al. (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 341:1237973. doi: 10.1126/science.1237973
- Erdel, F., and Rippe, K. (2018). Formation of chromatin subcompartments by phase separation. *Biophys. J.* 114, 2262–2270. doi: 10.1016/j.bpj.2018.03.011
- Fan, H., Lv, P., Huo, X., Wu, J., Wang, Q., Cheng, L., et al. (2018). The nuclear matrix protein HNRNPU maintains 3D genome architecture globally in mouse hepatocytes. *Genome Res.* 28, 192–202. doi: 10.1101/gr.224576.117
- Fort, V., Khelifi, G., and Hussein, S. M. I. (2021). Long non-coding RNAs and transposable elements: a functional relationship. *Biochim. Biophys. Acta BBA Mol. Cell Res.* 1868:118837. doi: 10.1016/j.bbamcr.2020.118837
- Frank, L., and Rippe, K. (2020). Repetitive RNAs as regulators of chromatin-associated subcompartment formation by phase separation. *J. Mol. Biol.* 432, 4270–4286. doi: 10.1016/j.jmb.2020.04.015
- Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A., and Mirny, L. A. (2017). Emerging evidence of chromosome folding by loop extrusion. *Cold Spring Harb. Symp. Quant. Biol.* 82, 45–55. doi: 10.1101/sqb.2017.82.034710
- Gebauer, F., Schwarzl, T., Valcárcel, J., and Hentze, M. W. (2021). RNA-binding proteins in human genetic disease. *Nat. Rev. Genet.* 22, 185–198. doi: 10.1038/s41576-020-00302-y
- Ghosal, G., and Muniyappa, K. (2006). Hoogsteen base-pairing revisited: resolving a role in normal biological processes and human diseases. *Biochem. Biophys. Res. Commun.* 343, 1–7. doi: 10.1016/j.bbrc.2006.02.148
- Girelli, G., Custodio, J., Kallas, T., Agostini, F., Wernersson, E., Spanjaard, B., et al. (2020). GPSeq reveals the radial organization of chromatin in the cell nucleus. *Nat. Biotechnol.* 38, 1184–1193. doi: 10.1038/s41587-020-0519-y
- Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M. B., Talhout, W., et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453, 948–951. doi: 10.1038/nature06947
- Hacisuleyman, E., Goff, L. A., Trapnell, C., Williams, A., Henao-Mejia, J., Sun, L., et al. (2014). Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* 21, 198–206. doi: 10.1038/nsmb.2764
- Hacisuleyman, F. E. (2015). *The Discovery and Characterization of the lncRNA Firre*. Ph.D thesis. Cambridge, MA: Harvard University.
- Hall, A. C., Ostrowski, L. A., and Mekhail, K. (2019). Phase separation as a melting pot for DNA repeats. *Trends Genet.* 35, 589–600. doi: 10.1016/j.tig.2019.05.001
- Hall, L. L., Carone, D. M., Gomez, A. V., Kolpa, H. J., Byron, M., Mehta, N., et al. (2014). Stable COT-1 repeat RNA is abundant and is associated with euchromatic interphase chromosomes. *Cell* 156, 907–919. doi: 10.1016/j.cell.2014.01.042

- Hall, L. L., and Lawrence, J. B. (2016). RNA as a fundamental component of interphase chromosomes: could repeats prove key? *Lisa. Physiol. Behav.* 37, 137–147. doi: 10.1016/j.gde.2016.04.005
- Heinz, S., Texari, L., Hayes, M. G. B., Urbanowski, M., Chang, M. W., Givarkes, N., et al. (2018). Transcription elongation can affect genome 3D structure. *Cell* 174, 1522.e–1536.e. doi: 10.1016/j.cell.2018.07.047
- Hentze, M. W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* 19, 327–341. doi: 10.1038/nrm.2017.130
- Hernandez, A. J., Zovoilis, A., Cifuentes-Rojas, C., Han, L., Bujisic, B., and Lee, J. T. (2020). B2 and ALU retrotransposons are self-cleaving ribozymes whose activity is enhanced by EZH2. *Proc. Natl. Acad. Sci. U.S.A.* 117, 415–425. doi: 10.1073/pnas.1917190117
- Holmes, D. S., and Bonner, J. (1974). Interspersion of repetitive and single-copy sequences in nuclear ribonucleic acid of high molecular weight. *Proc. Natl. Acad. Sci. U.S.A.* 71, 1108–1112. doi: 10.1073/pnas.71.4.1108
- Hou, Y., Zhang, R., and Sun, X. (2019). Enhancer lncRNAs influence chromatin interactions in different ways. *Front. Genet.* 10:936. doi: 10.3389/fgene.2019.00936
- Huo, X., Ji, L., Zhang, Y., Lv, P., Cao, X., Wang, Q., et al. (2020). The nuclear matrix protein SAFB cooperates with major satellite RNAs to stabilize heterochromatin architecture partially through phase separation. *Mol. Cell* 77, 368.e–383.e. doi: 10.1016/j.molcel.2019.10.001
- Hussong, M., Kaehler, C., Kerick, M., Grimm, C., Franz, A., Timmermann, B., et al. (2017). The bromodomain protein BRD4 regulates splicing during heat shock. *Nucleic Acids Res.* 45, 382–394. doi: 10.1093/nar/gkw729
- Hyman, A. A., Weber, C. A., and Jülicher, F. (2014). Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* 30, 39–58. doi: 10.1146/annurev-cellbio-100913-013325
- Isoda, T., Moore, A. J., He, Z., Chandra, V., Aida, M., Denholtz, M., et al. (2017). Non-coding transcription instructs cohesin-dependent chromatin folding and compartmentalization to dictate enhancer-promoter communication and T cell fate. *Cell* 171, 103–119. doi: 10.1016/j.cell.2017.09.001
- Jachowicz, J. W., Bing, X., Pontabry, J., Bošković, A., Rando, O. J., and Torres-Padilla, M.-E. (2017). LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat. Genet.* 49, 1502–1510. doi: 10.1038/ng.3945
- Jacob, M. D., Audas, T. E., Uniacke, J., Trinkle-Mulcahy, L., and Lee, S. (2013). Environmental cues induce a long noncoding RNA-dependent remodeling of the nucleolus. *Matera AG, editor. Mol. Biol. Cell* 24, 2943–2953. doi: 10.1091/mbc.e13-04-0223
- Janssen, A., Colmenares, S. U., and Karpen, G. H. (2018). Heterochromatin: guardian of the genome. *Annu. Rev. Cell Dev. Biol.* 34, 265–288. doi: 10.1146/annurev-cellbio-100617-062653
- Johnson, W. L., Yewdell, W. T., Bell, J. C., McNulty, S. M., Duda, Z., O'Neill, R. J., et al. (2017). RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. *Elife* 6:e25299.
- Jolly, C., Metz, A., Govin, J., Vigneron, M., Turner, B. M., Khochbin, S., et al. (2004). Stress-induced transcription of satellite III repeats. *J. Cell Biol.* 164, 25–33. doi: 10.1083/jcb.200306104
- Kawaguchi, T., Tanigawa, A., Naganuma, T., Ohkawa, Y., Souquere, S., Pierron, G., et al. (2015). SWI/SNF chromatin-remodeling complexes function in noncoding RNA-dependent assembly of nuclear bodies. *Proc. Natl. Acad. Sci. U.S.A.* 112, 4304–4309. doi: 10.1073/pnas.1423819112
- Kelley, D., and Rinn, J. (2012). Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol.* 13:R107. doi: 10.1186/gb-2012-13-11-r107
- Khan, M. M., Xiao, J., Patel, D., and LeDoux, M. S. (2018). DNA damage and neurodegenerative phenotypes in aged Ciz1 null mice. *Neurobiol. Aging* 62, 180–190. doi: 10.1016/j.neurobiolaging.2017.10.014
- Khosraviani, N., Ostrowski, L. A., and Mekhail, K. (2019). Roles for non-coding RNAs in spatial genome organization. *Front. Cell Dev. Biol.* 7:336. doi: 10.3389/fcell.2019.00336
- Kind, J., Pagie, L., de Vries, S. S., Nahidiazar, L., Dey, S. S., Bienko, M., et al. (2015). Genome-wide maps of nuclear lamina interactions in single human cells. *Cell* 163, 134–147. doi: 10.1016/j.cell.2015.08.040
- Klimek-Tomczak, K., Wyrwicz, L. S., Jain, S., Bomsztyk, K., and Ostrowski, J. (2004). Characterization of hnRNP K protein-RNA interactions. *J. Mol. Biol.* 342, 1131–1141. doi: 10.1016/j.jmb.2004.07.099
- Kolpa, H. J., Fackelmayer, F. O., and Lawrence, J. B. (2016). SAF-A requirement in anchoring XIST RNA to chromatin varies in transformed and primary cells. *Dev. Cell* 39, 9–10. doi: 10.1016/j.devcel.2016.09.021
- Kruse, K., Diaz, N., Enriquez-Gasca, R., Gaume, X., Torres-Padilla, M.-E., and Vaquerizas, J. M. (2019). Transposable elements drive reorganization of 3D chromatin during early embryogenesis. *bioRxiv* [Preprint] doi: 10.1101/523712
- Kung, J. T., Kesner, B., An, J. Y., Ahn, J. Y., Cifuentes-Rojas, C., Colognori, D., et al. (2015). Locus-specific targeting to the X-chromosome revealed by the RNA interactome of CTCF recruited in a locus-specific manner and implicates CTCF-RNA interactions in long-range chromosomal interactions. *Mol. Cell* 57, 361–375. doi: 10.1016/j.molcel.2014.12.006
- Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, H. Q., Ong, C.-T., et al. (2015). Widespread rearrangement of 3D chromatin organization underlies Polycomb-mediated stress-induced silencing. *Physiol. Behav.* 58, 216–231. doi: 10.1016/j.molcel.2015.02.023
- Li, X., and Fu, X.-D. (2019). Chromatin-associated RNAs as facilitators of functional genomic interactions. *Nat. Rev. Genet.* 20, 503–519. doi: 10.1038/s41576-019-0135-1
- Li, Y., Syed, J., and Sugiyama, H. (2016). RNA-DNA triplex formation by long noncoding RNAs. *Cell Chem. Biol.* 23, 1325–1333. doi: 10.1016/j.chembiol.2016.09.011
- Lieberman-Aiden, E., Van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., et al. (2009). Comprehensive mapping of long range interactions reveals folding principles of the human genome. *Science* 326, 289–393. doi: 10.1126/science.1181369
- Liyakat Ali, T. M., Brunet, A., Collas, P., and Paulsen, J. (2021). TAD cliques predict key features of chromatin organization. *BMC Genomics* 22:499. doi: 10.1186/s12864-021-07815-8
- Lohe, A. R., and Brutlag, D. L. (1986). Multiplicity of satellite DNA sequences in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 83, 696–700. doi: 10.1073/pnas.83.3.696
- Lohe, A. R., Hilliker, A. J., and Roberts, P. A. (1993). Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. *Trends Genet.* 9:379. doi: 10.1016/0168-9525(93)90135-5
- Lu, Z., Guo, J. K., Wei, Y., Dou, D. R., Zarnegar, B., Ma, Q., et al. (2020a). Structural modularity of the XIST ribonucleoprotein complex. *Nat. Commun.* 11:6163.
- Lu, J. Y., Shao, W., Chang, L., Yin, Y., Li, T., Zhang, H., et al. (2020b). Genomic repeats categorize genes with distinct functions for orchestrated regulation. *Cell Rep.* 30, 3296–3311. doi: 10.1016/j.celrep.2020.02.048
- Lubelsky, Y., and Ulitsky, I. (2018). Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature* 555, 107–111. doi: 10.1038/nature25757
- Lund, E., Oldenburg, A. R., and Collas, P. (2014). Enriched domain detector: a program for detection of wide genomic enrichment domains robust against local variations. *Nucleic Acids Res.* 42:e92. doi: 10.1093/nar/gku324
- Maass, P. G., Rump, A., Schulz, H., Stricker, S., Schulze, L., Platzer, K., et al. (2012). A misplaced lncRNA causes brachydactyly in humans. *J. Clin. Invest.* 122, 3990–4002. doi: 10.1172/JCI65508
- Maison, C., Bailly, D., Roche, D., de Oca, R. M., Probst, A. V., Vassias, I., et al. (2011). SUMOylation promotes de novo targeting of HP1 $\alpha$  to pericentric heterochromatin. *Nat. Genet.* 43, 220–227. doi: 10.1038/ng.765
- Mao, Y. S., Zhang, B., and Spector, D. L. (2011). Biogenesis and function of nuclear bodies. *Trends Genet.* 27, 295–306. doi: 10.1016/j.tig.2011.05.006
- McElligott, R., and Wellinger, R. J. (1997). The terminal DNA structure of mammalian chromosomes. *EMBO J.* 16, 3705–3714. doi: 10.1093/emboj/16.12.3705
- McHugh, C. A., Chen, C.-K., Chow, A., Surka, C. F., Tran, C., McDonel, P., et al. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* 521, 232–236. doi: 10.1038/nature14443
- Merkenschlager, M., and Nora, E. P. (2016). CTCF and cohesin in genome folding and transcriptional gene regulation. *Annu. Rev. Genomics Hum. Genet.* 17, 17–43. doi: 10.1146/annurev-genom-083115-022339

- Metz, A., Soret, J., Vourc'h, C., Tazi, J., and Jolly, C. (2004). A key role for stress-induced satellite III transcripts in the relocalization of splicing factors into nuclear stress granules. *J. Cell Sci.* 117, 4551–4558. doi: 10.1242/jcs.01329
- Michieletto, D., and Gilbert, N. (2019). Role of nuclear RNA in regulating chromatin structure and transcription. *Curr. Opin. Cell Biol.* 58, 120–125. doi: 10.1016/j.ceb.2019.03.007
- Mohammad, F., Pandey, R. R., Nagano, T., Chakalova, L., Mondal, T., Fraser, P., et al. (2008). Kcnq1ot1/Lit1 noncoding RNA mediates transcriptional silencing by targeting to the perinucleolar region. *Mol. Cell Biol.* 28, 3713–3728. doi: 10.1128/MCB.02263-07
- Mondal, T., Subhash, S., Vaid, R., Enroth, S., Uday, S., Reinius, B., et al. (2015). MEG3 long noncoding RNA regulates the TGF- $\beta$  pathway genes through formation of RNA-DNA triplex structures. *Nat. Commun.* 6:7743. doi: 10.1038/ncomms8743
- Monfort, A., and Wutz, A. (2020). The B-side of Xist. *F1000Res.* 9:F1000. doi: 10.12688/f1000research.21362.1
- Morgan, A. R., and Wells, R. D. (1968). Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. *J. Mol. Biol.* 37, 63–80.
- Nakagawa, S., and Prasanth, K. V. (2011). eXIST with matrix-associated proteins. *Trends Cell Biol.* 21, 321–327. doi: 10.1016/j.tcb.2011.02.001
- Németh, A., Conesa, A., Santoyo-Lopez, J., Medina, I., Montaner, D., Péterfia, B., et al. (2010). Initial Genomics of the Human Nucleolus. Akhtar A, editor. *PLoS Genet.* 6:e1000889. doi: 10.1371/journal.pgen.1000889
- Nickerson, J. A., Krockmalnic, G., Wan, K. M., and Penman, S. (1997). The nuclear matrix revealed by eluting chromatin from a cross-linked nucleus. *Proc. Natl. Acad. Sci. U.S.A.* 94, 4446–4450. doi: 10.1073/pnas.94.9.4446
- Ninomiya, K., Adachi, S., Natsume, T., Iwakiri, J., Terai, G., Asai, K., et al. (2020). lncRNA-dependent nuclear stress bodies promote intron retention through SR protein phosphorylation. *EMBO J.* 39:e102729. doi: 10.15252/embj.2019102729
- Nishibuchi, G., and Déjardin, J. (2017). The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* 25, 77–87. doi: 10.1007/s10577-016-9547-3
- Nozawa, R.-S., Boteva, L., Soares, D. C., Naughton, C., Dun, A. R., Buckle, A., et al. (2017). SAF-A regulates interphase chromosome structure through oligomerization with chromatin-associated RNAs. *Cell* 169, 1214.e–1227.e. doi: 10.1016/j.cell.2017.05.029
- Nozawa, R.-S., and Gilbert, N. (2019). RNA: nuclear glue for folding the genome. *Trends Cell Biol.* 29, 201–211. doi: 10.1016/j.tcb.2018.12.003
- Nwigwe, I. J., Kim, Y. J., Wacker, D. A., Kim, T. H., and Chadwick, B. P. (2015). Boundary associated long noncoding RNA mediates long-range chromosomal interactions. *PLoS One* 10:e0136104. doi: 10.1371/journal.pone.0136104
- Oya, E., Kato, H., Chikashige, Y., Tsutsumi, C., Hiraoka, Y., and Murakami, Y. (2013). Mediator directs co-transcriptional heterochromatin assembly by RNA interference-dependent and -independent pathways. *PLoS Genet.* 9:e1003677. doi: 10.1371/journal.pgen.1003677
- Pace, J. K., and Feschotte, C. (2007). The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage. *Genome Res.* 17, 422–432. doi: 10.1101/gr.5826307
- Palm, W., and de Lange, T. (2008). How shelterin protects mammalian telomeres. *Annu. Rev. Genet.* 42, 301–334. doi: 10.1146/annurev.genet.41.110306.130350
- Pan, H., Jin, M., Ghadiyaram, A., Kaur, P., Miller, H. E., Ta, H. M., et al. (2020). Cohesin SA1 and SA2 are RNA binding proteins that localize to RNA containing regions on DNA. *Nucleic Acids Res.* 48, 5639–5655. doi: 10.1093/nar/gkaa284
- Pandya-Jones, A., Markaki, Y., Serizay, J., Chitashvili, T., Mancía Leon, W. R., Damianov, A., et al. (2020). A protein assembly mediates Xist localization and gene silencing. *Nature* 587, 145–151. doi: 10.1038/s41586-020-2703-0
- Pathak, R., Mamillapalli, A., Rangaraj, N., Kumar, R., Vasanthi, D., Mishra, K., et al. (2013). AAGAG repeat RNA is an essential component of nuclear matrix in *Drosophila*. *RNA Biol.* 10, 564–571. doi: 10.4161/rna.24326
- Patrat, C., Okamoto, I., Diabangouaya, P., Vialon, V., Le Baccon, P., Chow, J., et al. (2009). Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5198–5203. doi: 10.1073/pnas.0810683106
- Paulsen, J., Liyakat Ali, T. M., Nekrasov, M., Delbarre, E., Baudement, M.-O., Kurscheid, S., et al. (2019). Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation. *Nat. Genet.* 51, 835–843. doi: 10.1038/s41588-019-0392-0
- Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S. W. M., Solovei, I., Brugman, W., et al. (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* 38, 603–613. doi: 10.1016/j.molcel.2010.03.016
- Pintacuda, G., Young, A. N., and Cerase, A. (2017b). Function by structure: spotlights on Xist long non-coding RNA. *Front. Mol. Biosci.* 4:90. doi: 10.3389/fmolb.2017.00090
- Pintacuda, G., Wei, G., Roustan, C., Kirmizitas, B. A., Solcan, N., Cerase, A., et al. (2017a). hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-repeat to establish polycomb-mediated chromosomal silencing. *Mol. Cell* 68, 955.e–969.e. doi: 10.1016/j.molcel.2017.11.013
- Pinter, S. F. A. (2016). Tale of two cities: how Xist and its partners localize to and silence the bicompartmental X. *Semin. Cell Dev. Biol.* 56, 19–34. doi: 10.1016/j.semcdb.2016.03.023
- Prasanth, K. V., Prasanth, S. G., Xuan, Z., Hearn, S., Freier, S. M., Bennett, C. F., et al. (2005). Regulating gene expression through RNA nuclear retention. *Cell* 123, 249–263. doi: 10.1016/j.cell.2005.08.033
- Quinodoz, S. A., and Guttman, M. (2021). Essential roles for RNA in shaping nuclear organization. *Cold Spring Harb. Perspect. Biol.* a039719.
- Quinodoz, S. A., Ollikainen, N., Tabak, B., Palla, A., Schmidt, J. M., Detmar, E., et al. (2018). Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* 174, 744.e–757.e. doi: 10.1016/j.cell.2018.05.024
- Racko, D., Benedetti, F., Dorier, J., and Stasiak, A. (2018). Transcription-induced supercoiling as the driving force of chromatin loop extrusion during formation of TADs in interphase chromosomes. *Nucleic Acids Res.* 46, 1648–1660. doi: 10.1093/nar/gkx1123
- Radion, E., Morgunova, V., Ryazansky, S., Akulenko, N., Lavrov, S., Abramov, Y., et al. (2018). Key role of piRNAs in telomeric chromatin maintenance and telomere nuclear positioning in *Drosophila* germline. *Epigenet. Chromatin* 11:40.
- Rangwala, S. H., Zhang, L., and Kazazian, H. H. (2009). Many LINE1 elements contribute to the transcriptome of human somatic cells. *Genome Biol.* 10:R100.
- Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680.
- Rebollo, R., Romanish, M. T., and Mager, D. L. (2012). Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu. Rev. Genet.* 46, 21–42.
- Richter, K., Nessler, M., and Lichter, P. (2007). Experimental evidence for the influence of molecular crowding on nuclear architecture. *J. Cell Sci.* 120, 1673–1680.
- Ridings-Figueroa, R., Stewart, E. R., Nesterova, T. B., Coker, H., Pintacuda, G., Godwin, J., et al. (2017). The nuclear matrix protein CIZ1 facilitates localization of Xist RNA to the inactive X-chromosome territory. *Genes Dev.* 31, 876–888.
- Rodermund, L., Coker, H., Oldenkamp, R., Wei, G., Bowness, J., Rajkumar, B., et al. (2021). Time-resolved structured illumination microscopy reveals key principles of Xist RNA spreading. *Science* 372:eabe7500.
- Rowley, M. J., Lyu, X., Rana, V., Ando-Kuri, M., Karns, R., Bosco, G., et al. (2019). Condensin II counteracts cohesin and RNA polymerase II in the establishment of 3D chromatin organization. *Cell Rep.* 26, 2890.e–2903.e.
- Royo, H., and Cavallé, J. (2008). Non-coding RNAs in imprinted gene clusters. *Biol. Cell* 100, 149–166.
- Saldaña-Meyer, R., Rodríguez-Hernaez, J., Escobar, T., Nishana, M., Jácome-López, K., Nora, E. P., et al. (2019). RNA interactions are essential for CTCF-mediated genome organization. *Mol. Cell* 76, 412.e–422.e.
- Saleh, A., Macia, A., and Muotri, A. R. (2019). Transposable elements, inflammation, and neurological disease. *Front. Neurol.* 10:894. doi: 10.3389/fneur.2019.00894
- Sanborn, A. L., Rao, S. S. P., Huang, S.-C., Durand, N. C., Huntley, M. H., Jewett, A. I., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U.S.A.* 112, E6456–E6465.
- Sawyer, I. A., Sturgill, D., and Dundr, M. (2019). Membraneless nuclear organelles and the search for phases within phases. *Wiley Interdiscip. Rev. RNA* 10:e1514.

- Schuettengruber, B., Bourbon, H.-M., Di Croce, L., and Cavalli, G. (2017). Genome regulation by polycomb and trithorax: 70 years and counting. *Cell* 171, 34–57.
- Shevtsov, S. P., and Dundr, M. (2011). Nucleation of nuclear bodies by RNA. *Nat. Cell Biol.* 13, 167–173.
- Singh, A. K., and Lakhota, S. C. (2015). Dynamics of hnRNPs and omega speckles in normal and heat shocked live cell nuclei of *Drosophila melanogaster*. *Chromosoma* 124, 367–383.
- Singh, I., Contreras, A., Cordero, J., Rubio, K., Dobersch, S., Günther, S., et al. (2018). MiCEE is a ncRNA-protein complex that mediates epigenetic silencing and nucleolar organization. *Nat. Genet.* 50, 990–1001.
- Smith, C. D., Shu, S., Mungall, C. J., and Karpen, G. H. (2007). The release 5.1 annotation of *Drosophila melanogaster* heterochromatin. *Science* 316, 1586–1591.
- Stalio, L., Guo, C.-J., Chen, L.-L., and Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* 22, 96–118.
- Su, M., Han, D., Boyd-Kirkup, J., Yu, X., and Han, J.-D. J. (2014). Evolution of alu elements toward enhancers. *Cell Rep.* 7, 376–385.
- Szabo, Q., Jost, D., Chang, J.-M., Cattoni, D. I., Papadopoulos, G. L., Bonev, B., et al. (2018). TADs are 3D structural units of higher-order chromosome organization in *Drosophila*. *Sci. Adv.* 4:eaar8082.
- Tan, J. Y., Smith, A. A. T., Ferreira da Silva, M., Matthey-Doret, C., Rueedi, R., Sönmez, R., et al. (2017). cis-Acting complex-trait-associated lincRNA expression correlates with modulation of chromosomal architecture. *Cell Rep.* 18, 2280–2288.
- Thakur, J., Fang, H., Llagas, T., Distèche, C. M., and Henikoff, S. (2019). Architectural RNA is required for heterochromatin organization. *bioRxiv* [Preprint] Available Online at: <http://biorxiv.org/lookup/doi/10.1101/784835>. (Accessed May 01, 2021).
- Titus, A. R., Ferreira, L. A., Belgovskiy, A. I., Kooyman, E. E., Mann, E. K., Mann, J. A., et al. (2020). Interfacial tension and mechanism of liquid-liquid phase separation in aqueous media. *Phys. Chem. Chem. Phys.* 22, 4574–4580.
- Trojer, P., and Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature? *Mol. Cell* 28, 1–13.
- Tucci, V., Isles, A. R., Kelsey, G., and Ferguson-Smith, A. C. (2019). Erice imprinting group: genomic imprinting and physiological processes in mammals. *Cell* 176, 952–965.
- Ulianov, S. V., Khrameeva, E. E., Gavrillov, A. A., Flyamer, I. M., Kos, P., Mikhaleva, E. A., et al. (2016). Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. *Genome Res.* 26, 70–84.
- Valgardsdottir, R., Chiodi, I., Giordano, M., Rossi, A., Bazzini, S., Ghigna, C., et al. (2008). Transcription of Satellite III non-coding RNAs is a general stress response in human cells. *Nucleic Acids Res.* 36, 423–434.
- Van Bortle, K., and Corces, V. G. (2012). Nuclear organization and genome function. *Annu. Rev. Cell Dev. Biol.* 28, 163–187.
- van Koningsbruggen, S., Gierlinski, M., Schofield, P., Martin, D., Barton, G. J., Ariyurek, Y., et al. (2010). High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. *Mol. Biol. Cell.* 21, 3735–3748.
- van Steensel, B., and Furlong, E. E. M. (2019). The role of transcription in shaping the spatial organisation of the genome. *Nat. Rev. Mol. Cell Biol.* 20, 327–337.
- Velazquez Camacho, O., Galan, C., Swist-Rosowska, K., Ching, R., Gamalinda, M., Karabiber, F., et al. (2017). Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. *eLife* 6:e25293.
- Vertii, A., Ou, J., Yu, J., Yan, A., Pagès, H., Liu, H., et al. (2019). Two contrasting classes of nucleolus-associated domains in mouse fibroblast heterochromatin. *Genome Res.* 29, 1235–1249.
- Vian, L., Pékowska, A., Rao, S. S. P., Kieffer-Kwon, K.-R., Jung, S., Baranello, L., et al. (2018). The energetics and physiological impact of cohesin extrusion. *Cell.* 175, 292–294.
- Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. S., and Martienssen, R. A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* 297, 1833–1837.
- Wang, T., Birsoy, K., Hughes, N. W., Krupczak, K. M., Post, Y., Wei, J. J., et al. (2015). Identification and characterization of essential genes in the human genome. *Science* 350, 1096–1101.
- Watson, J. D., and Crick, F. H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171, 737–738.
- Watson, J. D., and Crick, F. H. (1974). Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. J.D. Watson and F.H.C. Crick. Published in Nature, number 4356 April 25, 1953. *Nature* 248:765.
- Weierich, C., Brero, A., Stein, S., von Hase, J., Cremer, C., Cremer, T., et al. (2003). [Three-dimensional arrangements of centromeres and telomeres in nuclei of human and murine lymphocytes]. *Chromosome Res.* 11, 485–502.
- West, K. O., Scott, H. M., Torres-Odio, S., West, A. P., Patrick, K. L., and Watson, R. O. (2019). The splicing factor hnRNP M is a critical regulator of innate immune gene expression in macrophages. *Cell Rep.* 29, 1594.e–1609.e.
- Winter, D. J., Ganley, A. R. D., Young, C. A., Liachko, I., Schardl, C. L., Dupont, P. Y., et al. (2018). Repeat elements organise 3D genome structure and mediate transcription in the filamentous fungus *Epichloë festucae*. *PLoS Genet.* 14:e1007467. doi: 10.1371/journal.pgen.1007467
- 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)80248-8
- Xiao, R., Tang, P., Yang, B., Huang, J., Zhou, Y., Shao, C., et al. (2012). Nuclear matrix factor hnRNP U/SAF-A exerts a global control of alternative splicing by regulating U2 snRNP maturation. *Mol. Cell* 45, 656–668.
- Young, A. N., Perlas, E., Ruiz-Blanes, N., Hierholzer, A., Pomella, N., Martin-Martin, B., et al. (2021). Deletion of LBR N-terminal domains recapitulates Pelger-Huet anomaly phenotypes in mouse without disrupting X chromosome inactivation. *Commun. Biol.* 4:478.
- Zhang, L. F., Huynh, K. D., and Lee, J. T. (2007). Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing. *Cell* 129, 693–706.
- Zhang, S., Hemmerich, P., and Grosse, F. (2004). Nucleolar localization of the human telomeric repeat binding factor 2 (TRF2). *J. Cell Sci.* 117, 3935–3945.
- Zhang, Y., Li, T., Preissl, S., Amaral, M. L., Grinstein, J. D., Farah, E. N., et al. (2019). Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. *Nat. Genet.* 51, 1380–1388.
- Zheng, R., Shen, Z., Tripathi, V., Xuan, Z., Freier, S. M., Bennett, C. F., et al. (2010). Polypurine-repeat-containing RNAs: a novel class of long non-coding RNA in mammalian cells. *J. Cell Sci.* 123, 3734–3744.
- Zhu, L., and Brangwynne, C. P. (2015). Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr. Opin. Cell Biol.* 34, 23–30.
- Zlatanova, J., and Caiafa, P. (2009). CTCF and its protein partners: divide and rule? *J Cell Sci.* 122(Pt 9), 1275–1284.
- Żylicz, J. J., and Heard, E. (2020). Molecular mechanisms of facultative heterochromatin formation: an X-chromosome perspective. *Annu. Rev. Biochem.* 89, 255–282.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Trigiante, Blanes Ruiz and Cerase. 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.