HHS Public Access

Author manuscript

Nat Rev Genet. Author manuscript; available in PMC 2020 November 24.

Published in final edited form as:

Nat Rev Genet. 2019 September ; 20(9): 503–519. doi:10.1038/s41576-019-0135-1.

Chromatin-associated RNAs as facilitators of functional genomic interactions

Xiao Li, **Xiang-Dong Fu***

Department of Cellular and Molecular Medicine and Institute for Genomic Medicine, University of California, San Diego, La Jolla, CA, USA.

Abstract

Mammalian genomes are extensively transcribed, which produces a large number of both coding and non-coding transcripts. Various RNAs are physically associated with chromatin, through being either retained in *cis* at their site of transcription or recruited in *trans* to other genomic regions. Driven by recent technological innovations for detecting chromatin-associated RNAs, diverse roles are being revealed for these RNAs and associated RNA-binding proteins (RBPs) in gene regulation and genome function. Such functions include locus-specific roles in gene activation and silencing, as well as emerging roles in higher-order genome organization, such as involvement in long-range enhancer-promoter interactions, transcription hubs, heterochromatin, nuclear bodies and phase transitions.

Mammalian genomes are now known to be prevalently transcribed, with ~70% of the genome active in generating various RNA products[1]. Such a large transcript repertoire consists of both protein-coding mRNAs and non-coding RNAs (ncRNAs), the latter of which encompass traditional classes of ncRNA, such as tRNA, ribosomal RNA (rRNA), small nucleolar RNA (snoRNA) and small nuclear RNA (snRNA), as well as more recently described classes, including microRNA (miRNA), enhancer-associated RNA (eRNA) and long non-coding RNA (lncRNA)[2–5]. Most ncRNA classes interact with specific proteins, some of which may even form relatively stable ribonucleoprotein particles (RNPs) to function in the cell, as exemplified by the small and large sub units of the ribosome (with rRNAs) and the spliceosome (with U1, U2, U4/6 and U5 snRNAs).

Although it is entirely possible that many transcripts of low abundance are simply ‘noise’, perhaps reflecting a degree of promiscuous action of the transcription machinery to sample open chromatin regions, specific and diverse functions have been increasingly ascribed to each class of RNA molecules, including various architectural and/or gene regulatory roles in different cellular compartments, adding to a new dimension of functional genomics. Of course, most RNAs are unlikely to act alone and instead interact with specific RNA-binding proteins (RBPs). Like protein-based interactions, RNAs have specific sequence motifs as

* xdfu@ucsd.edu.

Author contributions

Both authors contributed to all aspects of the article.

Competing interests

The authors declare no competing interests.

well as secondary and higher-order structures that enable them to engage in sequence-mediated and structure-based molecular interactions; furthermore, RNAs have the additional ability to form base-pairing interactions with other nucleic acids.

In this Review, we first introduce the modes that RNAs can associate with chromatin in *cis* and in *trans*. We then focus on recent technological advances in detecting the chromatin-RNA interactome and emerging insights from the analysis of chromatin-associated RNAs and RBPs, including roles in gene activation and repression across diverse biological pathways, the formation of several types of nuclear body as well as potential functions in 3D genome organization. Notable findings are that chromatin-associated RNAs appear to form 'RNA clouds' over specific clusters of active gene promoters and their distal enhancers that are looped into proximity. Furthermore, specific RNA-mediated genomic interactions are linked to various liquid-liquid phase-separation processes[6,7], suggesting that the entire nucleus may be viewed as a membraned compartment of a series of organized yet highly dynamic membrane-less condensates that shape the functional genome.

Distinct mechanisms to retain RNA on chromatin

In principle, RNA may be associated with chromatin via one of two modes. One is that newly transcribed RNAs (nascent RNAs) remain at their sites of synthesis due to various mechanisms. This mode of RNA-chromatin interactions is referred to as *cis*-interactions. The second mode is *trans*-interactions, which result from RNAs being released from their sites of transcription to interact with different genomic loci. These modes may act in combination at a genomic locus: *cis*-acting RNAs may form specific RNPs, which may then recruit additional *trans*-acting RNAs or RNPs, and thus specific genomic loci may contain both *cis*- and *trans*-acting RNAs to achieve specific regulatory purposes.

Cis-acting RNAs.

One of the best-known *cis*-mechanisms for retaining nascent RNA to chromatin is through R-loop formation, which involves annealing nascent RNA back to template DNA during transcription, forming a double-stranded RNA:DNA hybrid opposite a displaced single strand of DNA (FIG. 1a). R-Loop formation and its functional impact on genome integrity have been reviewed[8]; however, where R-loops predominantly form in mammalian genomes remains unclear, except for a consensus of their occurrence on promoters and enhancers associated with GC-skewed sequences[9,10]. Importantly, R-loop formation and dynamics have been linked to transcription activities under physiological conditions[11–14].

Another mechanism to retain a nascent RNA on its site of transcription is through targeting of the nascent RNA by additionally amplified RNA (FIG. 1b). Amplified RNA may result from locally processed nascent RNA or be supplied from a different genomic locus, thus giving rise to genomic loci that contain both *cis*- and *trans*-acting RNAs. This mechanism has been exploited during the formation of telomeric and pericentromeric heterochromatin regions through the action of nuclear RNA interference (RNAi) in fission yeast[15–18] and the repression of retrotransposons by the Piwi-interacting RNA (piRNA) system in *Drosophila melanogaster*[19,20].

The largest class of nascent RNAs that remain on chromatin is protein-coding pre-mRNAs (FIG. 1c). A key mechanism for tethering nascent RNAs to chromatin is through their association with RNA polymerases during transcription. It is now well established that pre-mRNAs are predominately co-transcriptionally processed into mature mRNAs, although many partially processed mRNAs are also known to continue their processing post-transcriptionally after being released from chromatin. The classic co-transcriptional processing events include 5′ capping, pre-mRNA splicing and polyadenylation, which has been extensively reviewed[21,22]. Recent studies suggest that circular RNAs formed through a non-canonical back-splicing mechanism are another type of *cis*-acting RNA[23,24]. Whereas most circular RNAs are released from chromatin and currently are of unknown function, some appear to accumulate at their site of transcription, associating with the RNA polymerase II (Pol II) machinery and acting as a positive regulator of transcription[23,24]. However, the mechanism for this new regulatory paradigm remains essentially elusive. The functional benefit of co-transcriptional RNA splicing by the spliceosome has been well documented, which enables sequential recognition of functional splice sites in pre-mRNA emerging from a transcribing Pol II, and also allows various splicing enhancing and repressing factors to compete in a timely m anner[21 , 22 ,24], a nd, in addition, nascent RNA can in turn modulate transcription[26], as further detailed below. Certain secondary RNA processing events, such as miRNA processing and N^6-methyladenosine (m^6A) modification, also appear to take place in a co-transcriptional manner[27,28]. In this Review, rather than concentrating on how transcription influences pre-mRNA processing, we focus particularly on the other side of the relationship; that is, the potential impact of chromatin-associated pre-mRNA on transcription.

Trans-acting RNAs.

There are at least two modes for *trans*-acting RNAs to interact with chromatin. One is through the formation of triplex nucleic acid structures, which involves Hoogsteen base-pairing interactions of RNA with the major groove of double-stranded DNA[29,30]. This type of RNA-DNA interaction has a stringent requirement for a polypurine sequence in DNA and a length restriction (FIG. 1d). Triplex structures are thus distinct from R-loops, although both are RNA:DNA hybrids in nature. Multiple lncRNAs appear to use this triplex mechanism to dock on specific genomic loci to exert their regulatory functions[30,31], thus taking full advantage of the molecular recognition strategy that is uniquely suited for nucleic acids through sequence complementarity.

The most common mode of *trans*-interactions is through protein-mediated interactions with RNA, in which the RNA acts as a scaffold to bind an RBP, and at least one component of this RNP is in direct contact with a DNA-bound factor (FIG. 1e). Such a mechanism is exemplified by the role of scaffold-attachment factor A (SAF-A; also known as hnRNP U) in *trans*-localization of the lncRNA *Firre*[32]. Interestingly, besides traditionally studied RBPs with recognizable RNA-binding motifs, a large number of non-canonical RBPs also exist in mammalian cells, which first surfaced from analysis of ultraviolet-crosslinking of proteins to mRNA[33–35], ncRNA[36] or nascent RNA[37–39]. Some of these approaches take advantage of physicochemical separation of RNAs and proteins into aqueous and organic phases, respectively, where RNA-protein complexes are efficiently recovered from the interface of

the layers for mass spectrometry analyses[38,40]. Interestingly, both canonical and non-canonical RBPs tend to form complexes to function together, which may thus be deduced from protein-protein interactomes[41]. Many DNA-binding proteins may also have the capacity to bind RNA. Indeed, hnRNP K was first characterized as a DNA-binding transcription factor in the regulation of MYC expression[42–44], and the typical transcription factor PGC1α, which is critical for mitochondrial biogenesis, contains a classic RNA recognition motif (RRM) type of RNA-binding motif[45]. To date, multiple classic DNA-binding transcription factors and epigenome regulators have been found to have the ability to directly contact RNA, including EZH2 and SUZ12 (both are key components of Polycomb repressive complex 2 (PRC2))[46–50], YY1 (REFs[51,52]), NELF and DSIF[53,54], CCCTC-binding factor (CTCF)[55], DNA methyltransferases[56,57] and, more recently, SUV39 (REFs[58,59]). Therefore, regulatory RNAs may be recruited to specific genomic loci through direct interactions with an array of both canonical and non-canonical RBPs.

Combinatorial cis- and trans-interaction modes for RNAs.

RNA interaction with chromatin in *cis* or in *trans* does not have to be mutually exclusive, but rather can be a combination of both, as discussed above for nascent RNA interacting with amplified small RNAs during the establishment of heterochromatin. Perhaps the best established example for combinatorial RNA actions is the lncRNA *Xist*, which is involved in dosage compensation in mammals, as has been extensively reviewed[60,61]. A fascinating aspect of *Xist* interaction with the inactive X chromosome (Xi) is its *cis*-action initiated by the expression of several ncRNAs from the *Xist* locus followed by spreading of the repressive chromatin signature trimethylation of K27 on histone H3 (H3K27me3) across the entire Xi[62]. This spreading phase is likely to involve *Xist* that acts in *trans*, as depletion of a nuclear matrix protein SAF-A appears to prevent *Xist* from accumulating on the Xi[63,64]. The heterochromatin signature is catalysed by PRC2, and spreading of the silenced state requires cooperative interactions between PRC1 and PRC2. However, the critical RNA elements and the sequence of actions of PRC1 and PRC2 have been the subject of ongoing debate[65–67]. In any case, *Xist*-mediated epigenetic regulation has been a prototype for studying RNA-chromatin interactions in mammalian cells.

Strategies to detect RNA on chromatin

Chromatin-associated RNAs can be profiled using fractionated chromatin followed by deep sequencing[68]. In a standard cell fractionation procedure, nuclei are isolated and sedimentation is used to separate nuclear particles from the nucleoplasmic fraction. Such nuclear particles have been considered the chromatin fraction and thus used to extract RNA for sequencing. However, caution must be taken in interpreting the results because the isolated particle fraction consists of not just chromatin but also a wider ‘nuclear matrix’, and thus the standard nuclear fractionation protocol cannot fully differentiate between chromatin-associated versus nuclear-matrix-attached RNAs. To identify true chromatin-associated RNAs, several experimental approaches have been developed to capture RNA on chromatin, which can be roughly divided into ‘one-to-many’ versus ‘all-to-all’ strategies.

One-to-many strategies.

The one-to-many strategies typically use a specific synthetic RNA or DNA to capture chromatin associated with a single RNA species, followed by deep sequencing to identify the constituent DNA loci within the isolated chromatin (FIG. 2a). These include chromatin isolation by RNA purification (ChIRP)[69], capture hybridization analysis of RNA targets (CHART)[70] and RNA antisense purification (RAP)[71]. For capturing RNA-containing chromatin, all three methods use biotin-labelled probes to hybridize to a specific target RNA. To ensure specificity, multiple probes are divided into two separate pools (the so-called odd and even pools of probes) to identify common targets[69]. However, this approach is ineffective in dealing with the background problem, as certain RNAs or RNA-containing particles may stick to beads or probes in a non-specific manner. One way to overcome this problem is to design a series of probes and use RNase H to test which probes are accessible to their intended RNA target and which probes are not, as shown earlier[70,72]. This would generate two separate sets of probes for affinity capture in parallel so that the signals captured by the accessible probe set can be considered the foreground and those non-specifically pulled down with the inaccessible probe set provide the background. The RAP procedure partially alleviates the specificity problem by using a longer biotin-labelled probe for RNA capture[71]. In any case, all capture-based approaches would benefit from background correction after knockdown of the intended RNA.

All capture-based approaches have an intrinsic limitation as the procedures can only study one specific RNA at a time, but the advantage is their ability to study potential chromatin-associated RNA regardless of their abundance in the cell, so long as non-specific signals can be efficiently managed by both experimental and computational approaches. Another major advantage of such capture-based approaches is the ability to couple with multiple other strategies to detect all potential interactions in a single set of experiments, as illustrated with domain-specific ChIRP (dChIRP)[73] to study specific RNA-associated proteins by mass spectrometry[74,75] and RNA-RNA interactions via psoralen or ultraviolet-crosslinking as in RAP-RNA and RNA hybrid and individual-nucleotide resolution ultraviolet-crosslinking and immunoprecipitation (hiCLIP), respectively[76,77]. In principle, RNA capture can also be coupled with high-throughput chromosome conformation capture (Hi-C)[78] to detect RNA-linked DNA-DNA interactions.

All-to-all strategies.

Recently, three groups developed related all-to-all strategies to detect RNA-chromatin interactions globally across RNAs, all of which are based on the use of a bivalent linker that is able to ligate to RNA on one end and to restriction-digested DNA on the other end; the methods are applicable to fixed and isolated nuclei (FIG. 2b). One strategy is ‘mapping RNA-genome interactions’ (MARGI), which uses a bivalent linker to ligate to RNA and DNA followed by circularization for library construction[79]. The other two independently developed strategies, global RNA interaction with DNA sequencing (GRID-seq)[80] and chromatin-associated RNA sequencing (ChAR-seq)[81], are nearly identical but with one distinction: GRID-seq uses a linker that carries two restriction sites for a type IIS restriction enzyme (MmeI) and subsequent digestion to generate size-specific products[80], whereas ChAR-seq sonicates ligated products to obtain smaller fragments before the addition of

adaptors for library construction[81]. The restriction-based approach of GRID-seq ensures that all reads contain RNA and DNA sequences, which is not certain with either MARGI or ChAR-seq, but the trade-off is the decreased mapping power of short reads with GRID-seq, whereas MARGI and ChAR-seq require more sequencing to obtain a sufficient number of mated RNA and DNA reads. It is also important to emphasize that all of these strategies are based on RNA proximity to nearby DNA, thus they do not necessarily capture only direct RNA-chromatin interactions.

Although all of these global mapping strategies provide options for the research community to use and further improve, the most important need is to ensure specificity. The first pass in evaluating the library quality before further bioinformatics analysis is to demonstrate the strand specificity, as the sequences from the RNA side are supposed to selectively align with the transcribed DNA strand with evidence for producing detectable transcripts, whereas the sequences from the DNA side are without strand specificity. A second important specificity issue is to differentiate between specific and non-specific RNA-chromatin interactions. It is conceivable that a specific RNA may interact with a single or multiple specific genomic regions (either through *cis*- or *trans*-interaction modes), but the same RNA may also travel around in the nucleus once released from its site of synthesis, thus sticking to all accessible chromatin regions to contribute to a general RNA-chromatin interaction background. GRID-seq addresses this problem by using interspecies RNAs (for example, mixing human and *D. melanogaster* cells during library construction) so that human RNAs ligated to fly DNA and vice versa are clearly indicative of background interactions. This background turns out to be nearly identical to that based on normalized total *trans*-acting mRNAs (that is, mRNA ligated to all chromosomes other than the chromosome that transcribes the mRNA), thus allowing the use of the collection of internal *trans*-acting mRNAs as a control to compute specific RNA-chromatin interactions. ChAR-seq uses spike-in RNA for such non-specific control during library construction, thus providing an estimate for the degree of non-specific RNA-chromatin interactions. However, unlike GRID-seq, published work with either MARGI or ChAR-seq did not develop a null model to infer specific interactions at the individual genomic loci.

It is conceivable that additional strategies can be developed to detect RNA-chromatin interactions either globally or at specific genomic loci by coupling with other detection methods. For example, the recently developed split-pool recognition of interactions by tag extension (SPRITE) technology used a barcoding strategy to detect DNA-DNA interactions from a sonicated nuclear fraction instead of relying on proximity ligation on fixed nuclei as in Hi-C, which in principle can also be used to detect global RNA-RNA and RNA-DNA interactions[82] (FIG. 2c).

A common drawback of all-to-all approaches is modest sensitivity, and, as a result, many low-abundance chromatin-associated RNAs may escape detection. One way to overcome this limitation for a genomic locus of interest is to use methods targeted to the genomic locus itself (rather than targeting RNAs), for locus-specific analysis of RNA-protein, DNA-protein and RNA-DNA interactions. A fluorescence in situ hybridization (FISH)-based strategy has been developed to capture specific genomic loci for multiomic analysis[83], and a more recent study reported a strategy to use guide RNA coupled with biotinylated nuclease-

dead Cas9 (dCas9) to detect locus-specific interactions[84] (FIG. 2d). Therefore, different strategies may be combined to address specific biological questions. It is likely that many new inventions are on the horizon to study RNA-chromatin interactions.

Nascent RNAs in gene regulation

Chromatin-associated RNAs (see BOX 1 for different classes of chromatin-associated RNAs documented in the literature), whether acting in *cis* or in *trans*, may have regulatory functions in gene expression. Such regulation may activate or silence gene expression in an RNA-dependent and/or locus-specific manner. As a consequence of these regulatory functions, genomic regions that are active in transcription remain as euchromatin, whereas those silenced can eventually become heterochromatin. Interestingly, evidence suggests that these processes are likely to involve various RNA-mediated feedback and feedforward mechanisms. In the following, we discuss different types of RNA detected on chromatin and biological insights into their regulatory functions.

The concept of transcription hubs in the 3D genome.

Increasing evidence suggests that nascent RNAs are not only products but may also function as regulators of transcription. One class of such nascent RNAs consists of promoter upstream transcripts (PROMPTs) and eRNAs, both of which are capped but not polyadenylated[85,86]. Interestingly, these RNAs were not efficiently captured by GRID-seq, which may be due to blockage of their ligation to the linker, potentially because of their involvement in R-loop formation and their relatively low abundance. PROMPT-induced R-loops may enhance the recruitment of specific transcription factors and chromatin remodellers[87]. The production of eRNAs is clearly associated with enhancer activities[88–90], but how eRNAs might contribute to transcriptional activation remains a subject of many ongoing investigations. A recent study suggests that eRNA is able to directly bind to key transcription co-activators, such as CBP-p300, which may represent an RNA-dependent feedforward mechanism to amplify enhancer activities[91]. Various lines of evidence suggest that eRNAs may facilitate specific DNA looping of enhancers to target promoters to enhance transcription[92–94], but this has not become a consensus in the field because other studies failed to detect significant changes in DNA looping after blocking eRNA production[90,95]. Interestingly, deletion of certain target promoters also diminishes eRNA expression[88]. Therefore, the 3D proximity and transcriptional activity at promoter-enhancer interactions might be mutually reinforcing, thus creating a local environment for recycling transcriptional factors[96]. However, the extent to which 3D genomic interactions are a cause versus a consequence of transcriptional activation remains entirely unclear.

A puzzling observation from recent genomic studies is that genetic inactivation of a given enhancer does not always lead to detectable functional consequence on transcription[95,97]. This might be due to redundant functions of multiple enhancers that target a given promoter and/or the activation of a ‘shadow enhancer’[98]. In general, it has been unclear which enhancers are redundant to one another and how many enhancers may work together, either redundantly or synergistically, to activate a given promoter. Interestingly, analysis of the GRID-seq data reveals that nascent pre-mRNAs are not only retained on their sites of

transcription but also selectively interact with distal genomic sequences (including enhancers) in the spatial proximity, as if nascent transcripts form individual ‘clouds’ over regulatory DNA elements[99] (FIG. 3a,b). This observation, coupled with Pol II-associated DNA-DNA interactions detected by chromatin interaction analysis by paired-end tag sequencing (ChIA-PET)[100], introduces the concept of nascent transcript-decorated ‘transcription hubs’ in the 3D space of the nucleus. On average, each hub contains ~4 active promoters, one or two super-enhancers and up to ~20 typical enhancers[80]. This hub concept, which has been similarly proposed based on the interactions of key transcription co-activators with super-enhancers[101] and high-resolution imaging analysis of transcription factors and Pol II[102,103], provides a potential explanation for the lack of functional impact when a given enhancer is genetically inactivated, as it is conceivable that deletion of one enhancer in a transcription hub may not be sufficiently detrimental to transcription. Based on this hub concept, it also becomes conceivable that removal of one enhancer or promoter might cause reorganization of an existing hub, thus inducing a degree of reassignment of promoters and enhancers in the hub, as recently reported[104]. As a result of such modulation of transcription hubs, a gene may become associated with or excluded from an existing hub during its activation or repression.

A key unsolved question is whether a given hub-associated RNA simply reflects its transient interaction with nearby DNA elements or actually has an organizational role in creating or maintaining such a hub for coordinated transcription of gene clusters. Various lncRNAs have been documented to have *cis*-effects on the regulation of neighbouring genes[105,106]. A dramatic case is the lncRNA *roX2*, which is involved in dosage compensation in *D. melanogaster*, where it appears to cover most DNA-DNA interactions on the *roX2*-bound X chromosome[73]. However, it remains unclear whether such potential function in organizing genomic interactions applies only to particular lncRNAs or to hub-associated nascent RNAs in general. This question is also pertinent for the debate on whether an eRNA is required for mediating a specific promoter-enhancer interaction or is actually involved in an interaction network within a hub. Given the fact that deletion of a particular enhancer may or may not have a functional consequence, as it may or may not disrupt a transcription hub, it is possible that blocking a specific eRNA may not have a measurable functional consequence. Thus, the concept of transcription hubs may help to explain various conflicting results in the literature with respect to the functional requirement of eRNAs for transcription control[96]. Additionally, although individual transcription hubs appear to be mainly associated with nascent RNAs transcribed from the genes associated with the hub, many lncRNAs are also thought to function in *trans*, thereby influencing the transcription of other genes. Whether or not the deduced transcription hubs correspond to the proposed ‘transcription factories’ in the nucleus[107] is another intriguing question to address in future studies. A benefit of transcription factories is to allow transcription factors to recycle locally, as demonstrated on an induced heat shock gene in *D. melanogaster*[108].

Nascent RNA for dynamic control of gene expression.

When a nascent pre-mRNA first emerges from transcribing Pol II at the promoter-proximal region, it might form an R-loop, which was recently shown to contribute to transcriptional pausing[9]. Conversely, R-loop formation may also create a favourable chromatin environment

for the recruitment of transcription activators, such as Tip60 and p400, thus facilitating transcription pause release[87] (FIG. 4a). A more recent study shows that R-loops are able to activate gene promoters by recruiting the DNA methylcytosine dioxygenase TET1 to demethylate CpG islands[109]. The 5′ end of certain pre-mRNAs also carries specific sequence elements that may not only enhance co-transcriptional splicing but also promote transcriptional pause release (FIG. 4b). This has been demonstrated with SRSF2, a member of the SR family member of splicing factors, which undergoes translocation from the promoter-associated *7SK* complex to nascent RNA when a high-affinity binding site for this RBP newly emerges from Pol II[110]. Importantly, this switch from the *7SK* ncRNA to nascent RNA is accompanied by the release of the pTEFb enzyme from the *7SK* complex to catalyse multiple phosphorylation events, including those on the Ser2 positions in the C-terminal domain (CTD) of Pol II and on transcription elongation factors DISF and NELF, leading to transcriptional pause release[110]. Interestingly, this cascade of events was first elucidated with the HIV Tat protein[111,112], and, as both Tat and SRSF2 contain related arginine-rich sequences, both proteins appear to activate transcription using a similar mechanism. More recently, as support for the physiological relevance of this process, mutations disrupting SRSF2 function were found to be causal for the blood progenitor disorder myelodysplastic syndromes[13,113,114].

Compared to transcriptional activation via RBPs binding to discrete sequences in nascent RNAs, PRC2, the enzyme responsible for depositing the H3K27me3 mark on facultative heterochromatin[115], has been found to interact with newly transcribed RNAs in a highly promiscuous fashion[49,116]. Despite such a broad RNA binding profile, PRC2 clearly has a degree of specificity[117], for example, GC-rich sequences as recently elucidated[50]. It is also remarkable to note that multiple subunits of PRC2 appear to have the capacity to directly contact RNA based on in vitro binding studies, either with individual purified subunits[46–48,118,119] or using pre-assembled complexes[50,117,120]. Importantly, such interactions between PRC2 and pre-mRNAs appear to offer a dynamic mechanism for transcription control[121–123], as highlighted by two seemingly competing models (FIG. 4c). In one, PRC2 appears to bind to DNA and RNA in a mutually exclusive manner[49,120,124–126], which may help to explain why PRC2 does not affect many active genes, as those genes may maintain their highly active state of transcription by using nascent RNAs to prevent PRC2 from gaining access to chromatin (FIG. 4c, left).

Interestingly, however, mammalian genomes are also populated with a large number of so-called bivalent promoters that harbour both trimethylation of K4 on histone H3 (H3K4me3), a positive chromatin mark for gene activation, and H3K27me3, a negative chromatin mark for gene silencing[123,127], which is dynamically regulated during the transition from naive to primed embryonic stem cells[128,129]. The question is how such bivalency is maintained on a subset of gene promoters in a given cell type. As established with *Xist*[47], nascent RNA may guide and facilitate the initial recruitment of PRC2, and additional *Xist*-interacting proteins, such as SHARP and SPEN, may act in synergy with PRC2 to enhance heterochromatin spreading and transcriptional silencing[74,130]. This principle may apply to bivalent genes in general (FIG. 4c, right), which is consistent with a broad role of the RBP RBFox2 in partnership with PRC2 facilitating H3K27me3 deposition in a nascent-RNA-dependent manner[131], taking advantage of nascent-RNA-recruited PRC2 in a poised state[125].

Importantly, a more recent study revealed that RNA, although competing with chromatin for PRC2 binding, does not block the histone methyltransferase activity per se[120]. Therefore, after the initial nascent-RNA-mediated recruitment of PRC2, any additional mechanism to anchor and/or stabilize PRC2 near bivalent promoters — including the existing H3K27me3 that has been shown to enhance PRC2 binding through its EED subunit[132] — may contribute to the maintenance of the bivalency of the promoters to limit their transcriptional output. Such a mechanism thus represents a form of nascent-RNA-dependent feedback control in regulated gene expression. In reconciling the two competing models for PRC2 targeting and regulation by nascent RNAs, it is also important to keep in mind that various PRC2 components may interact with chromatin independently of the core complex, as shown with EZH2 (REF.[133]) and SUZ12 (REFs[49,134]), suggesting that PRC2 might function on different sets of genes through partnership with different protein binding partners in the cell.

Repeat-derived RNAs for gene silencing

Compared to facultative heterochromatin, constitutive heterochromatin is more stable. It is decorated by di/trimethylation of K9 on histone H3 (H3K9me2/3) and occurs on retrotransposons, telomeres and centromeric and pericentromeric regions of autosomes. Our current understanding of the formation of constitutive heterochromatin, which is mainly based on studies of fission yeast and *D. melanogaster*, has been extensively reviewed[135–137]. In general, the process takes place in three phases: first, transcription of repeat-containing DNA; second, spreading of the H3K9me2/3 heterochromatin mark to neighbouring sequences; and third, inheritance of such a silenced state after cell division. An intriguing aspect of this epigenetic control mechanism is the fact that transcription is required for silencing transcription.

In fission yeast, initial repeat-derived transcripts are amplified by an RNA-dependent RNA polymerase (RdRP); such amplified double-stranded RNAs are then processed by Dicer to produce small interfering RNAs (siRNAs); and upon loading onto Ago1, these guide RNAs target nascent RNAs to induce transcriptional silencing[138]. Both initiation and subsequent spreading events are enforced by two positive feedforward loops (FIG. 5a). The first loop is RNA-mediated and involves the interaction of the RNA-induced transcriptional silencing (RITS) complex with chromatin-associated repeat transcripts, where the RITS complex recruits specific H3K9me2/3 methyltransferases (Clr4 in fission yeast; SUV39H1 and SUV39H2 in humans) in a large complex. Once some initial H3K9me2/3 marks are deposited, the signals are next amplified by the second, RNA-independent loop, as the methyltransferase itself is able to bind to K9-methylated histone 3 and this histone mark is also recognized by an additional methylated K9 binding protein (Swi6 in fission yeast; HP1 in humans). Multiple protein-protein interactions reinforce this process, thereby facilitating H3K9me2/3 spreading to adjacent sequences in a sequence-context-dependent manner[139,140]. Interestingly, a more recent study showed that both SUV39H1 and SUV39H2 are also able to directly bind to RNA[58,59]. Both of these H3K9 methyltransferases contain a chromodomain, which has been long known to bind to RNA, but without much specificity[141]. SUV39H2 carries an additional basic amino-terminal sequence, which has recently been shown to bind to RNA as well[59]. These findings account for their RNA-

dependent binding to heterochromatin[59,142], together suggesting additional molecular connections between the first and second feedforward loops.

The dilemma with the general applicability of this elegantly elucidated pathway is that *D. melanogaster* and mammalian genomes do not encode RdRP. In flies, a related RNAi pathway is involved[19] (FIG. 5b). Initially, primary piRNAs are generated from retrotransposons, which are loaded onto the Piwi complex to target retrotransposon transcripts with the aid of multiple Piwi complex components, including Asterix[143] and Panoramix[144]. To fully establish the silent state both locally and on other loci, repeat-derived RNAs are consecutively processed by the ping-pong mechanism[145]. As in fission yeast, the spreading phase involves a network of interactions that involves Piwi, SUV39, HP1 and several other factors[146,147]. However, although the piRNA system provides a mechanism to amplify transposon-derived small RNAs to silence retrotransposons in both flies and mammals[20], it remains unclear whether such a piRNA-dependent pathway is also responsible for establishing constitutive heterochromatin in centromeric and pericentromeric regions. As the piRNA pathway mainly operates in germ cells in both flies and mammals, there is a possibility that the heterochromatic state is maintained throughout development once established in the germline. Alternatively, certain reinforcing mechanisms might exist, which would be consistent with the expression of major satellite repeats in somatic cells in mammals[148]. In this regard, it is interesting to note that an intron-targeted siRNA has been shown to induce H3K9me2/3, which modulates alternative splicing probably through reduced transcription elongation[149], and multiple Argonaute proteins, including Ago1 and Ago2, have been detected on chromatin[150–152]. These findings suggest that whereas the piRNA system is devoted to silencing retrotransposons in germ cells, the siRNA pathway might be responsible for the formation and/or maintenance of constitutive heterochromatin in centromeric and pericentromeric regions in somatic cells.

Heterochromatin is heritable after cell division and even through multiple generations of some organisms. In *Caenorhabditis elegans*, for example, this process appears to involve multiple proteins that organize different types of granules in the cytoplasm[153]. A general question is whether both maintenance and inheritance of constitutive heterochromatin require not only locally amplified RNAs but also repeat-derived RNAs supplied in *trans*, because constitutive heterochromatin still undergoes a detectable degree of turnover[154]. A recent study in fission yeast suggests that siRNAs supplied in *trans* can indeed induce heterochromatin formation on the endogenous allele in the euchromatin region[155]. Therefore, it appears that after initial RNA-dependent induction of heterochromatin formation, additional RNAs may be required to maintain the status of the silenced chromatin state. This complex pathway may thus be connected to additional RNA processing steps, as suggested by the roles of various factors involved in pre-mRNA splicing, transport and RNA decay for the inheritance of heterochromatin[156,157]. For example, a variant of the RNA export factor NXF1 has a role in maintaining silencing (and thus limiting transposition activity) of intracisternal A particle (IAP) retrotransposons in mice[158,159]. Much remains to be explored to understand the functional interplay between regulated RNA metabolism and epigenetic control of gene expression.

RNAs in nuclear body formation

A newly emerged concept is the liquid-liquid phase separation associated with various membrane-less subcellular domains[6,7]. It is striking to note that the vast majority of the examples to date involve RNAs. Because RBPs are the major class of proteins that contain so-called low-complexity domains (LCDs), which are intrinsically disordered and have the tendency to form liquid-like droplets in vitro[160,161], it has become evident that RNAs may drive various phase-transition processes, thereby contributing to the organization and dynamics of the 3D genome (FIG. 6a). Here, we discuss the function of chromatin-associated RNAs in this context.

The nucleolus: a paradigm for conceptualizing the formation of nuclear bodies.

The nucleolus is a highly specialized nuclear body for ribosome biogenesis[162] which has been suggested to provide a general conceptual framework for the formation of other nuclear bodies, such as Cajal bodies and nuclear speckles[163–165]. The nucleolus is organized around ribosomal DNA (rDNA) clusters distributed on five different chromosomes in mammalian cells, in which transcription of rDNA is coupled with rRNA processing and ribosome assembly. Such concerted activities are subdivided into at least three cytologically distinct zones within the nucleolus (FIG. 6b): the fibrillar centre (FC) for rRNA transcription, the dense fibrillar compartment (DFC) for rRNA processing and the granule compartment (GC) for rRNA modification and ribosome assembly[162]. Each of these compartments is associated with a unique set of proteins, such as Pol I for the FC, fibrillarin (FIB1) for the DFC and nucleoplasmin (NPM1) for the GC. Remarkably, a recent study showed that each of these compartments has the property of liquid droplets, which undergo highly dynamic homotypic fusion as well as coalescence between different compartments, and disruption of the nuclear actin network causes coalescing of all three compartments into a single liquid droplet[166].

Although this elegant study demonstrated that FIB1 and NPM1 are each able to form separate droplets due to their LCDs, it leaves open the questions of what connects these droplets and what is the mechanism (or mechanisms) to keep distinct droplets apart? Co-transcriptional rRNA processing[167] and ribosome assembly[168] that take place in a highly ordered fashion are likely to be responsible for anchoring these compartments to the sites of rRNA transcription, as well as for connecting all three compartments together. Interestingly, a recent study revealed that the RNA helicase DDX21 forms a ring around multiple Pol I complexes, which is regulated by the snoRNA-containing lncRNA *SLEAT* to regulate rDNA transcription[169]. It is thus possible that DDX21 rings may help keep the FC from fusing to the DFC and GC. Another lncRNA, *LoNA* was recently found to bind via its 5′ sequence to the RBP nucleolin (NCL) to modulate rDNA transcription and through its 3′ sequence with FIB1 to regulate rRNA modification[170]. *LoNA* may thus help to bridge the FC to both the DFC and GC. Based on these recent discoveries, it is conceivable that although compartment-specific proteins form separate liquid droplets, distinct sets of RNAs and RBPs may help to keep them connected as well as to prevent them from coalescing, suggesting that RNPs may play central roles in facilitating liquid-liquid phase separation to coordinate ribosome biogenesis in the nucleus.

Cajal body: layered RNP liquid droplets?

Cajal bodies are nuclear structures associated with the biogenesis and recycling of multiple snRNAs required for pre-mRNA splicing[171,172]. Interestingly, several histone genes are also linked to this nuclear subdomain[173,174]. Multiple observations suggest that Cajal body formation and dynamic regulation follow almost identical principles to the nucleolus (FIG. 6c). First, the engineered expression of the replication-dependent histone gene *H2B* or an exogenous array of U2 snRNA expression units was found to be sufficient for de novo formation of a Cajal body[165,175]. A recent study using the circularized chromosome conformation capture with deep sequencing (4C-seq) method also showed that multiple U-type snRNA genes as well as Cajal body-associated histone genes are clustered around Cajal bodies, which involves not only intra-chromosomal but also inter-chromosomal interactions[176]. This initiation mechanism thus resembles a nucleolus assembly resulting from rDNA expression from different chromosomes. Second, snRNP maturation from Cajal bodies requires the survival of motor neuron (SMN) complex and related Germin proteins[177,178], and, interestingly, these proteins show intimate association but not direct overlap with Coilin, a scaffold protein for Cajal bodies[179,180]. The integrity of Cajal bodies also depends on nuclear actin[181], and live-cell imaging suggests constant fission and fusion during the cell cycle[182]. These observations are highly reminiscent of distinct subcompartments in the nucleolus, suggesting that individual Germin-decorated Cajal bodies may correspond to separate liquid droplets that are connected by nascent snRNAs as well as a highly intricate protein-protein interaction network during their co-transcriptional processing and assembly[183]. Of note, the cytoplasmic P-bodies and stress granules appear to mirror the behaviour of Cajal bodies in the nucleus via multivalent RNA-protein interactions[184,185], especially in light of the recent identification of multiple spatially stacked granules that are important for transgenerational RNAi inheritance in *C. elegans*[153].

Nuclear speckles and paraspeckles: storage sites or organizers of gene clusters?

Nuclear speckles are a major type of nuclear subdomain of mammalian cells that harbour the majority of splicing factors[163] (FIG. 6d). Nuclear speckles colocalize with one of the most abundant lncRNAs, *MALAT1*, although this lncRNA does not appear to be required for nuclear speckle assembly or maintenance[186]. Nuclear speckles are thought by many in the splicing field to be storage sites for splicing factors, as indicated by the observed recruitment of splicing factors from nuclear speckles to a nearby chromosomal locus expressing an induced pre-mRNA[187]. However, by RNA capture, *MALAT1* was found to associate with active gene promoters[75], which has been confirmed more recently with SPRITE[82], showing extensive inter-chromosomal interactions around nuclear speckles. Thus, it appears that, although the core of nuclear speckles is mainly composed of splicing complexes ready for recycling, highly active genes are organized around individual nuclear speckles, as if chromatin-tethered RNPs are drawn towards a nearby nuclear speckle through surface-tension-driven coalescence, as recently demonstrated with a dCas9-based optogenetic technology[188]. In response to inhibition of transcription, nuclear speckles become coalesced and this process is readily reversible upon resuming transcription[189], indicating that nascent RNAs serve to connect chromatin to nuclear speckles.

It is interesting that recent studies have unveiled a series of phase transition events from transcription to co- and post-transcriptional RNA processing, which may underlie the spatial relationship between chromatin and nuclear speckles. As recently proposed[190], enhancers (especially super-enhancers) involved in transcriptional activation may undergo sequential phase transitions, which involves LCD-containing transcription co-activators, such as MED1 and BRD4 (REFs[101,191]), as well as Pol II itself[192,193]. RNA-binding splicing factors and regulators are a major class of LCD-containing proteins. For example, the splicing regulator RBFox2 has been found to be part of the large assembly of splicing regulators whose assembly depends on critical tyrosine residues in its LCD[194], and multiple hnRNP splicing regulators also showed the propensity to form liquid droplets that are important for their splicing regulation functions[195]. One may thus envision a cascade of liquid-liquid phase transitions from the sites of transcription to the core of nuclear speckles.

Paraspeckles represent another interesting liquid-liquid phase transition paradigm. This nuclear subdomain is always adjacent to a nuclear speckle (hence the name, FIG. 6d), but unlike nuclear speckles the associated lncRNA *NEAT1* has been shown to be required for the assembly and maintenance of this nuclear subdomain in mammalian cells[196,197]. Paraspeckles harbour >30 RBPs, many of which are known to undergo phase separation, as exemplified by the analysis of paraspeckle-associated RBM14 and FUS (also known as TLS)[198,199]. It has been unclear why a paraspeckle is always next to a nuclear speckle, but a recent study revealed that *NEAT1* nucleates a large number of RBPs that collectively bind to almost the entire ensemble of primary miRNAs (pri-miRNAs) as well as the microprocessor for their processing[200]. Because ~80% of pri-miRNAs reside in introns of various pre-mRNAs[201], which are likely to be processed at or near nuclear speckles, this strongly suggests a concerted interplay between primary (intron removal) and secondary (pri-miRNA processing) RNA processing events that contributes to the spatial relationship between paraspeckles and nuclear speckles, whereby separate liquid droplets are connected via nascent pre-mRNAs.

Heterochromatin also forms liquid droplets.

Compared to nascent-RNA-nucleated nuclear bodies, recent studies reveal that HP1, a major component of heterochromatin, is able to form liquid droplets in vitro and to nucleate focus formation in vivo[202,203]. This functional property of HP1 is likely to account for the interaction among different telomeres, which are major genomic regions largely composed of constitutive heterochromatin, as if each telomere end may form a liquid droplet and certain groups of telomere droplets may fuse to connect different chromosomes in the nucleus[204]. As discussed above, heterochromatin is also nucleated by chromatin-associated, repeat-derived RNAs, thus hinting at the involvement of RNPs in organizing and regulating dynamic fusion and fission of HP1-containing liquid droplets, especially when *trans*-acting repeat-derived RNAs are joining the process for maintenance and inheritance of heterochromatin. Therefore, the regulatory principle appears similar to that for other nuclear bodies that involve both nascent chromatin-associated RNAs as well as additional *trans*-acting RNAs (that is, snRNAs and histone mRNAs in Cajal bodies as well as partially processed pre-mRNAs in nuclear speckles and paraspeckles). The principle of RNA-nucleated formation of these major nuclear bodies may also apply to the formation of other

types of cellular bodies, such as PML and Sam68 bodies in the nucleus[205,206] and Yb body for primary piRNA biogenesis in the cytoplasm[207,208].

RNAs in functional genome organization

Despite intensive studies of genome organization in the past decade, a fundamental issue remains regarding genomic interactions and genome organization as a cause or a consequence of gene expression. This problem is also pertinent to RNAs, which may have regulatory functions in transcription rather than being simply products of transcription. Analysis of chromatin-associated RNAs suggests that these RNAs, whether produced in *cis* or joined in *trans*, may provide the proposed multiple valencies for gene activation by transcription enhancers, especially super-enhancers[190]. Because nascent RNAs associated with super-enhancers are the major class of chromatin-associated RNAs revealed by GRID-seq[80], such nascent-RNA-nucleated transcription hubs are likely to contribute to the establishment of chromosome territories[209,210]. Because active genes tend to locate at the periphery of individual chromosome territories and nuclear bodies (particularly nuclear speckles) are known to fill the spaces between chromosome territories[211], coordinated transcriptional and co-transcriptional RNA processing events may thus provide driving forces for both intra- and inter-chromosomal interactions in the nucleus.

Many RBPs are now known to have intrinsically disordered domains for driving liquid droplet formation and mediating liquid-liquid phase separation. Interestingly, recent studies show that even certain RNAs, such as PolyQ-encoding triplet repeats, contain specific secondary structures that help facilitate phase separation[212,213]. Importantly, RNAs not only facilitate liquid-liquid phase transition but also prevent the collapse of liquid droplets, and in certain cases they also inhibit the formation of pathological aggregation of specific RBPs[214–217], as highlighted by their functional implications in the development of neurodegenerative diseases[218]. Therefore, nuclear RNAs and their interactions with specific RBPs may be viewed as linkers for different parts of the genome.

To keep different RNP droplets separate, it appears that at least two large forces are operating. One is through the attachment of heterochromatic regions to the nuclear envelope[219] and the other may be related to the poorly defined and highly debated nuclear matrix[220]. One of the main components of the operationally defined nuclear matrix is SAF-A, and several studies highlight its role in establishing nuclear architecture for regulated gene expression. Depletion of SAF-A prevents *Xist*-containing complexes from targeting the X chromosome to be inactivated[63,64], which is accompanied by the induction of additional Cajal bodies[221] and impaired long-distance DNA-DNA interactions[222]. A more recent study further revealed that the ATPase activity of SAF-A regulates its interaction with chromatin-associated RNAs and that oligomerization of this nuclear matrix protein drives chromatin decompaction[223]. Therefore, the nuclear matrix may provide a general meshwork for docking of different nuclear compartments as well as active exchanges of components from those compartments.

Conclusions and perspectives

The organization of a functional genome in the nucleus has been largely investigated from the perspective of DNA-DNA interactions. In this Review, we focus on the contribution of chromatin-associated RNAs to genome organization and the formation of various nuclear sub-domains. Given prevalent transcription from mammalian genomes, many different classes of ncRNAs are now appreciated for their regulatory functions, not just as products of gene expression. The apparent formation of transcription hubs decorated by protein-coding RNA suggests that pre-mRNAs may also function as lncRNAs for regulated gene expression before and/or during their processing into mature mRNAs[26,131]. Future research will be devoted to understanding the potential function of chromatin-associated RNAs, including both nascent pre-mRNAs and different classes of ncRNAs that are likely to interact with chromatin as RNPs. Different RNAs may function either as organizers of specific transcription hubs for gene activation or repression, or as mediators for defined phase-separation processes in the nucleus. Most published studies have concentrated on RNAs transcribed from well-annotated genomic regions, but less effort has been geared towards the chromatin interaction profiles of repeat-derived RNAs, which are difficult to differentiate between their *cis*- and *trans*-modes of action. Much progress is anticipated in this fast-moving field, which will provide critical insights not only into gene function and regulation in development and differentiation but also into vital disease processes.

Acknowledgements

The authors are grateful for the critical comments of B. Hamilton and the referees, which helped to improve the manuscript. The work in the authors' laboratories was supported by grants HG004659, GM049369 and GM052872 (to X.-D.F.) from the US National Institutes of Health.

Glossary

Protein-coding mRNAs
RNAs that have protein-coding capacity after processing.

Non-coding RNAs (ncRNAs).
RNAs that do not have protein-coding capability, although some RNAs classified as ncRNAs actually encode small peptides.

Ribonucleoprotein particles (RNPs).
Particles consisting of RNAs and RNA-binding proteins.

3D genome
Genomic DNA organized in different functional states, forming higher-order structures in the 3D space of the nucleus.

RNA clouds
Traditionally referring to large-scale regions of accumulated RNA in microscopy images (particularly mammalian *Xist* RNA coating the entire inactive X chromosome), here we use a broader definition encompassing chromatin-associated RNA foci at various genomic

scales, including RNAs associating with clusters of DNA elements or segments of chromosomes.

Promoters
DNA segments that nucleate the formation of the transcription initiation complex to drive transcription.

Enhancers
DNA segments that enhance transcription. Many active enhancers are now known to also generate RNA known as enhancer RNA.

Liquid-liquid phase separation
Formation of membrane-less assemblies or condensates of proteins or protein complexes that undergo active fusion and fission.

R-loop
An RNA:DNA hybrid with the non-template strand remaining as single-stranded DNA.

Heterochromatin
Compacted chromosomal regions that largely lack transcription activity.

Triplex
A three-strand RNA:DNA structure that involves non-canonical base-pairing between RNA and double-stranded DNA.

Shadow enhancer
A DNA element that is used as an alternative enhancer when another enhancer is inactivated.

Transcription hubs
Active gene clusters that may be far apart in terms of linear chromosomal distance but are folded into 3D spatial proximity. Genes in transcription hubs may functionally modulate one another, and their close spatial proximity allows local recycling of the transcription machinery.

Super-enhancers
DNA segments ranging in size from 10 to 40 kb that contain clusters of multiple active enhancer elements.

Typical enhancers
DNA segments, usually <0.5 kb in length, that can enhance transcription of nearby promoters.

Transcription factories
Nuclear regions that have relatively high concentrations of the transcription machinery.

Low-complexity domains (LCDs).
Protein domains containing dipeptide repeats enriched with glycine, tyrosine, serine, glutamine, phenylalanine, arginine and so forth.

References

1. Djebali S et al. Landscape of transcription in human cells. Nature 489, 101–108 (2012). [PubMed: 22955620]
2. Rinn JL & Chang HY Genome regulation by long noncoding RNAs. Annu. Rev. Biochem 81, 145–166 (2012). [PubMed: 22663078]
3. Fu XD Non-coding RNA: a new frontier in regulatory biology. Natl Sci. Rev 1, 190–204 (2014). [PubMed: 25821635]
4. Vance KW & Ponting CP Transcriptional regulatory functions of nuclear long noncoding RNAs. Trends Genet. 30, 348–355 (2014). [PubMed: 24974018]
5. Mattick JS The state of long non-coding RNA biology. Noncod. RNA 4, 17 (2018).
6. Hyman AA, Weber CA & Julicher F Liquid-liquid phase separation in biology. Annu. Rev. Cell Dev. Biol 30, 39–58 (2014). [PubMed: 25288112]
7. Banani SF, Lee HO, Hyman AA & Rosen MK Biomolecular condensates: organizers of cellular biochemistry. Nat. Rev. Mol. Cell Biol 18, 285–298 (2017). [PubMed: 28225081]
8. Sollier J & Cimprich KA Breaking bad: R-loops and genome integrity. Trends Cell Biol. 25, 514–522 (2015). [PubMed: 26045257]
9. Chen L et al. R-ChIP using inactive RNase H reveals dynamic coupling of R-loops with transcriptional pausing at gene promoters. Mol. Cell 68, 745–757 (2017). [PubMed: 29104020]
10. Ginno PA, Lott PL, Christensen HC, Korf I & Chedin F R-Loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol. Cell 45, 814–825 (2012). [PubMed: 22387027]
11. Chedin F Nascent connections: R-loops and chromatin patterning. Trends Genet. 32, 828–838 (2016). [PubMed: 27793359]
12. Stork CT et al. Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. eLife 5, e17548 (2016). [PubMed: 27552054]
13. Chen L et al. The augmented R-loop is a unifying mechanism for myelodysplastic syndromes induced by high-risk splicing factor mutations. Mol. Cell 69, 412–425 (2018). [PubMed: 29395063]
14. Skourti-Stathaki K, Proudfoot NJ & Gromak N Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol. Cell 42, 794–805 (2011). [PubMed: 21700224]
15. Kato H et al. RNA polymerase II is required for RNAi-dependent heterochromatin assembly. Science 309, 467–469 (2005). [PubMed: 15947136]
16. Castel SE & Martienssen RA RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat. Rev. Genet 14, 100–112 (2013). [PubMed: 23329111]
17. Wassenegger M The role of the RNAi machinery in heterochromatin formation. Cell 122, 13–16 (2005). [PubMed: 16009128]
18. Holoch D & Moazed D RNA-mediated epigenetic regulation of gene expression. Nat. Rev. Genet 16, 71–84 (2015). [PubMed: 25554358]
19. Pal-Bhadra M et al. Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. Science 303, 669–672 (2004). [PubMed: 14752161]
20. Iwasaki YW, Siomi MC & Siomi H PIWI-interacting RNA: its biogenesis and functions. Annu. Rev. Biochem 84, 405–433 (2015). [PubMed: 25747396]
21. Bentley DL Coupling mRNA processing with transcription in time and space. Nat. Rev. Genet 15, 163–175 (2014). [PubMed: 24514444]
22. Naftelberg S, Schor IE, Ast G & Kornblihtt AR Regulation of alternative splicing through coupling with transcription and chromatin structure. Annu. Rev. Biochem 84, 165–198 (2015). [PubMed: 26034889]
23. Zhang Y et al. Circular intronic long noncoding RNAs. Mol. Cell 51, 792–806 (2013). [PubMed: 24035497]
24. Zhang XO et al. Complementary sequence-mediated exon circularization. Cell 159, 134–147 (2014). [PubMed: 25242744]

25. Han J, Xiong J, Wang D & Fu XD Pre-mRNA splicing: where and when in the nucleus. Trends Cell Biol. 21, 336–343 (2011). [PubMed: 21514162]
26. Skalska L, Beltran-Nebot M, Ule J & Jenner RG Regulatory feedback from nascent RNA to chromatin and transcription. Nat. Rev. Mol. Cell Biol 18, 331–337 (2017). [PubMed: 28270684]
27. Morlando M et al. Primary microRNA transcripts are processed co-transcriptionally. Nat. Struct. Mol. Biol 15, 902–909 (2008). [PubMed: 19172742]
28. Slobodin B et al. Transcription impacts the efficiency of mRNA translation via co-transcriptional N6-adenosine methylation. Cell 169, 326–337 (2017). [PubMed: 28388414]
29. Martianov I, Ramadass A, Serra Barros A, Chow N & Akoulitchev A Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. Nature 445, 666–670 (2007). [PubMed: 17237763]
30. Li Y, Syed J & Sugiyama H RNA-DNA triplex formation by long noncoding RNAs. Cell Chem. Biol 23, 1325–1333 (2016). [PubMed: 27773629]
31. Wang KC & Chang HY Molecular mechanisms of long noncoding RNAs. Mol. Cell 43, 904–914 (2011). [PubMed: 21925379]
32. Hacisuleyman E et al. Topological organization of multichromosomal regions by the long intergenic noncoding RNA *Firre*. Nat. Struct. Mol. Biol 21, 198–206 (2014). [PubMed: 24463464] This work suggests a role for the lncRNA *Firre* in organizing inter-chromosomal interactions in the nucleus.
33. Baltz AG et al. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol. Cell 46, 674–690 (2012). [PubMed: 22681889]
34. Castello A et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149, 1393–1406 (2012). [PubMed: 22658674] This work systematically identifies RBPs by capturing proteins ultraviolet-crosslinked to RNA followed by mass spectrometry analysis.
35. Kwon SC et al. The RNA-binding protein repertoire of embryonic stem cells. Nat. Struct. Mol. Biol 20, 1122–1130 (2013). [PubMed: 23912277]
36. Huang R, Han M, Meng L & Chen X Transcriptome-wide discovery of coding and noncoding RNA-binding proteins. Proc. Natl Acad. Sci. USA 115, E3879–E3887 (2018). [PubMed: 29636419]
37. Bao X et al. Capturing the interactome of newly transcribed RNA. Nat. Methods 15, 213–220 (2018). [PubMed: 29431736]
38. Trendel J et al. The human RNA-binding proteome and its dynamics during translational arrest. Cell 176, 391–403 (2019). [PubMed: 30528433]
39. He C et al. High-resolution mapping of RNA-binding regions in the nuclear proteome of embryonic stem cells. Mol. Cell 64, 416–430 (2016). [PubMed: 27768875]
40. Queiroz RML et al. Comprehensive identification of RNA-protein interactions in any organism using orthogonal organic phase separation (OOPS). Nat. Biotechnol 37, 169–178 (2019). [PubMed: 30607034]
41. Brannan KW et al. SONAR discovers RNA-binding proteins from analysis of large-scale protein-protein interactomes. Mol. Cell 64, 282–293 (2016). [PubMed: 27720645]
42. Baber JL, Libutti D, Levens D & Tjandra N High precision solution structure of the C-terminal KH domain of heterogeneous nuclear ribonucleoprotein K, a c-myc transcription factor. J. Mol. Biol 289, 949–962 (1999). [PubMed: 10369774]
43. Michelotti EF, Michelotti GA, Aronsohn AI & Levens D Heterogeneous nuclear ribonucleoprotein K is a transcription factor. Mol. Cell. Biol 16, 2350–2360 (1996). [PubMed: 8628302]
44. Takimoto M et al. Specific binding of heterogeneous ribonucleoprotein particle protein K to the human c-myc promoter, in vitro. J. Biol. Chem 268, 18249–18258 (1993). [PubMed: 8349701]
45. Monsalve M et al. Direct coupling of transcription and mRNA processing through the thermogenic coactivator PGC-1. Mol. Cell 6, 307–316 (2000). [PubMed: 10983978]
46. Cifuentes-Rojas C, Hernandez AJ, Sarma K & Lee JT Regulatory interactions between RNA and polycomb repressive complex 2. Mol. Cell 55, 171–185 (2014). [PubMed: 24882207]
47. Zhao J, Sun BK, Erwin JA, Song JJ & Lee JT Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 322, 750–756 (2008). [PubMed: 18974356]

48. Kanhere A et al. Short RNAs are transcribed from repressed Polycomb target genes and interact with polycomb repressive complex-2. Mol. Cell 38, 675–688 (2010). [PubMed: 20542000]
49. Beltran M et al. The interaction of PRC2 with RNA or chromatin is mutually antagonistic. Genome Res. 26, 896–907 (2016). [PubMed: 27197219] This study reveals that the PRC2 component SUZ12 is able to bind to chromatin independently of all other PRC2 subunits, raising the possibility that PRC2 may not always function as part of a multi-subunit core complex in the cell.
50. Wang X et al. Targeting of Polycomb repressive complex 2 to RNA by short repeats of consecutive guanines. Mol. Cell 65, 1056–1067 (2017). [PubMed: 28306504]
51. Belak ZR & Ovsenek N Assembly of the Yin Yang 1 transcription factor into messenger ribonucleoprotein particles requires direct RNA binding activity. J. Biol. Chem 282, 37913–37920 (2007). [PubMed: 17974562]
52. Belak ZR, Ficzycz A & Ovsenek N Biochemical characterization of Yin Yang 1-RNA complexes. Biochem. Cell Biol 86, 31–36 (2008). [PubMed: 18364743]
53. Missra A & Gilmour DS Interactions between DSIF (DRB sensitivity inducing factor), NELF (negative elongation factor), and the Drosophila RNA polymerase II transcription elongation complex. Proc. Natl Acad. Sci. USA 107, 11301–11306 (2010). [PubMed: 20534440]
54. Pagano JM et al. Defining NELF-E RNA binding in HIV-1 and promoter-proximal pause regions. PLOS Genet. 10, e1004090 (2014). [PubMed: 24453987]
55. Sun S et al. Jpx RNA activates Xist by evicting CTCF. Cell 153, 1537–1551 (2013). [PubMed: 23791181]
56. Di Ruscio A et al. DNMT1-interacting RNAs block gene-specific DNA methylation. Nature 503, 371–376 (2013). [PubMed: 24107992]
57. Savell KE et al. Extra-coding RNAs regulate neuronal DNA methylation dynamics. Nat. Commun 7, 12091 (2016). [PubMed: 27384705]
58. Johnson WL et al. RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. eLife 6, e25299 (2017). [PubMed: 28760200]
59. Velazquez Camacho O et al. Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. eLife 6, e25293 (2017). [PubMed: 28760199]
60. Lee JT Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. Genes Dev. 23, 1831–1842 (2009). [PubMed: 19684108]
61. Wutz A Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nat. Rev. Genet 12, 542–553 (2011). [PubMed: 21765457]
62. Jegu T, Aeby E & Lee JT The X chromosome in space. Nat. Rev. Genet 18, 377–389 (2017). [PubMed: 28479596]
63. Hasegawa Y et al. The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. Dev. Cell 19, 469–476 (2010). [PubMed: 20833368]
64. Pullirsch D et al. The Trithorax group protein Ash2l and Saf-A are recruited to the inactive X chromosome at the onset of stable X inactivation. Development 137, 935–943 (2010). [PubMed: 20150277]
65. Brockdorff N Polycomb complexes in X chromosome inactivation. Phil. Trans. R. Soc. B 372, 20170021 (2017). [PubMed: 28947664]
66. da Rocha ST et al. Jarid2 is implicated in the initial Xist-induced targeting of PRC2 to the inactive X chromosome. Mol. Cell 53, 301–316 (2014). [PubMed: 24462204]
67. Almeida M et al. PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. Science 356, 1081–1084 (2017). [PubMed: 28596365] This recent work suggests that PRC1, rather than PRC2, may initiate X-chromosome inactivation through docking on the XN region of *Xist*.
68. Werner MS & Ruthenburg AJ Nuclear fractionation reveals thousands of chromatin-tethered noncoding RNAs adjacent to active genes. Cell Rep. 12, 1089–1098 (2015). [PubMed: 26257179]
69. Chu C, Qu K, Zhong FL, Artandi SE & Chang HY Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol. Cell 44, 667–678 (2011). [PubMed: 21963238]

70. Simon MD et al. The genomic binding sites of a noncoding RNA. Proc. Natl Acad. Sci. USA 108, 20497–20502 (2011). [PubMed: 22143764]
71. Engreitz JM et al. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science 341, 1237973 (2013). [PubMed: 23828888] This report shows that *Xist*-mediated spreading of the silenced chromatin state occurs through exploiting folded X-chromosome organization rather than proceeding in a linear fashion.
72. Lee N, Moss WN, Yario TA & Steitz JA EBV noncoding RNA binds nascent RNA to drive host PAX5 to viral DNA. Cell 160, 607–618 (2015). [PubMed: 25662012]
73. Quinn JJ et al. Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. Nat. Biotechnol 32, 933–940 (2014). [PubMed: 24997788] This work develops domain-specific CHIRP for identifying RNA domain-specific interactions with RNA, protein and DNA, revealing an architecture of the lncRNA *roX2* involved in dosage compensation in *D. melanogaster*.
74. McHugh CA et al. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature 521, 232–236 (2015). [PubMed: 25915022]
75. West JA et al. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. Mol. Cell 55, 791–802 (2014). [PubMed: 25155612]
76. Engreitz JM et al. RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites. Cell 159, 188–199 (2014). [PubMed: 25259926]
77. Sugimoto Y et al. hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. Nature 519, 491–494 (2015). [PubMed: 25799984]
78. Lieberman-Aiden E et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293 (2009). [PubMed: 19815776]
79. Sridhar B et al. Systematic mapping of RNA-chromatin interactions in vivo. Curr. Biol 27, 602–609 (2017). [PubMed: 28132817] The work reports a strategy to map RNA-chromatin interactions in fixed nuclei.
80. Li X et al. GRID-seq reveals the global RNA-chromatin interactome. Nat. Biotechnol 35, 940–950 (2017). [PubMed: 28922346] This study reports a technology called GRID-seq to detect the RNA-chromatin interactome in fly and mammalian cells and revealed nascent-RNA-associated transcription hubs that link long-distance promoters and enhancers into spatial proximity.
81. Bell JC et al. Chromatin-associated RNA sequencing (ChAR-seq) maps genome-wide RNA-to-DNA contacts. eLife 7, e27024 (2018). [PubMed: 29648534] This work reports a technology called ChAR-seq, which is similar to GRID-seq, to determine the RNA-chromatin interactome in a ,*D. melanogaster* cell line.
82. Quinodoz SA et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. Cell 174, 744–757 (2018). [PubMed: 29887377] This study introduces a ligation-free approach to study 3D genome organization and suggests the existence of inter-chromosomal hubs.
83. Dejardin J & Kingston RE Purification of proteins associated with specific genomic loci. Cell 136, 175–186 (2009). [PubMed: 19135898]
84. Liu X et al. In situ capture of chromatin interactions by biotinylated dCas9. Cell 170, 1028–1043 (2017). [PubMed: 28841410]
85. Preker P et al. RNA exosome depletion reveals transcription upstream of active human promoters. Science 322, 1851–1854 (2008). [PubMed: 19056938]
86. Lam MT, Li W, Rosenfeld MG & Glass CK Enhancer RNAs and regulated transcriptional programs. Trends Biochem. Sci 39, 170–182 (2014). [PubMed: 24674738]
87. Chen PB, Chen HV, Acharya D, Rando OJ & Fazzio TG R loops regulate promoter-proximal chromatin architecture and cellular differentiation. Nat. Struct. Mol. Biol 22, 999–1007 (2015). [PubMed: 26551076]
88. Kim TK et al. Widespread transcription at neuronal activity-regulated enhancers. Nature 465, 182–187 (2010). [PubMed: 20393465]
89. Wang D et al. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature 474, 390–394 (2011). [PubMed: 21572438]
90. Hah N, Murakami S, Nagari A, Danko CG & Kraus WL Enhancer transcripts mark active estrogen receptor binding sites. Genome Res. 23, 1210–1223 (2013). [PubMed: 23636943]

91. Bose DA et al. RNA binding to CBP stimulates histone acetylation and transcription. Cell 168, 135–149 (2017). [PubMed: 28086087]
92. Hsieh CL et al. Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. Proc. Natl Acad. Sci. USA 111, 7319–7324 (2014). [PubMed: 24778216]
93. Lai F et al. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. Nature 494, 497–501 (2013). [PubMed: 23417068]
94. Li W et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature 498, 516–520 (2013). [PubMed: 23728302]
95. Schaukowitch K et al. Enhancer RNA facilitates NELF release from immediate early genes. Mol. Cell 56, 29–42 (2014). [PubMed: 25263592]
96. Li W, Notani D & Rosenfeld MG Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nat. Rev. Genet 17, 207–223 (2016). [PubMed: 26948815]
97. Moorthy SD et al. Enhancers and super-enhancers have an equivalent regulatory role in embryonic stem cells through regulation of single or multiple genes. Genome Res 27, 246–258 (2017). [PubMed: 27895109]
98. Hong JW, Hendrix DA & Levine MS Shadow enhancers as a source of evolutionary novelty. Science 321, 1314 (2008). [PubMed: 18772429]
99. Nozawa RS & Gilbert N RNA: nuclear glue for folding the genome. Trends Cell Biol. 29, 201–211 (2019). [PubMed: 30630665]
100. Li G et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell 148, 84–98 (2012). [PubMed: 22265404]
101. Sabari BR et al. Coactivator condensation at super-enhancers links phase separation and gene control. Science 361, 379 (2018).This study reveals that transcription factors and coactivators are able to form phase-separated condensates at super-enhancers to concentrate the transcription machinery at cell-identity genes.
102. Cho WK et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science 361, 412–415 (2018). [PubMed: 29930094]
103. Chong S et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science 361, 378 (2018).
104. Tan Y et al. Dismissal of RNA polymerase II underlies a large ligand-induced enhancer decommissioning program. Mol. Cell 71, 526–539 (2018). [PubMed: 30118678]
105. Wang KC et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 472, 120–124 (2011). [PubMed: 21423168]
106. Isoda T et al. Non-coding transcription instructs chromatin folding and compartmentalization to dictate enhancer-promoter communication and T cell fate. Cell 171, 103–119 (2017). [PubMed: 28938112]
107. Sutherland H & Bickmore WA Transcription factories: gene expression in unions? Nat. Rev. Genet 10, 457–466 (2009). [PubMed: 19506577]
108. Yao J, Ardehali MB, Fecko CJ, Webb WW & Lis JT Intranuclear distribution and local dynamics of RNA polymerase II during transcription activation. Mol. Cell 28, 978–990 (2007). [PubMed: 18158896]
109. Arab K et al. GADD45A binds R-loops and recruits TET1 to CpG island promoters. Nat. Genet 51, 217–223 (2019). [PubMed: 30617255]
110. Ji X et al. SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. Cell 153, 855–868 (2013). [PubMed: 23663783] This study demonstrates that the classic SR protein SRSF2 has a splicing-independent role in transcriptional activation through mediating pTEFb translocation from the *7SK* complex to nascent RNA, a mechanism first elucidated with HIV Tat.
111. Ott M, Geyer M & Zhou Q The control of HIV transcription: keeping RNA polymerase II on track. Cell Host Microbe 10, 426–435 (2011). [PubMed: 22100159]
112. Peterlin BM & Price DH Controlling the elongation phase of transcription with P-TEFb. Mol. Cell 23, 297–305 (2006). [PubMed: 16885020]

113. Komeno Y et al. SRSF2 is essential for hematopoiesis, and its myelodysplastic syndrome-related mutations dysregulate alternative pre-mRNA splicing. Mol. Cell. Biol 35, 3071–3082 (2015). [PubMed: 26124281]
114. Kim E et al. SRSF2 mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. Cancer Cell 27, 617–630 (2015). [PubMed: 25965569]
115. Cao R et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298, 1039–1043 (2002). [PubMed: 12351676]
116. Davidovich C, Zheng L, Goodrich KJ & Cech TR Promiscuous RNA binding by Polycomb repressive complex 2. Nat. Struct. Mol. Biol 20, 1250–1257 (2013). [PubMed: 24077223]
117. Davidovich C et al. Toward a consensus on the binding specificity and promiscuity of PRC2 for RNA. Mol. Cell 57, 552–558 (2015). [PubMed: 25601759]
118. Zhao J et al. Genome-wide identification of Polycomb-associated RNAs by RIP-seq. Mol. Cell 40, 939–953 (2010). [PubMed: 21172659]
119. Kaneko S et al. Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. Mol. Cell 53, 290–300 (2014). [PubMed: 24374312]
120. Wang X et al. Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. Nat. Struct. Mol. Biol 24, 1028–1038 (2017). [PubMed: 29058709] This work shows that RNA prevents PRC2 from binding to DNA but does not inhibit its methyltransferase activity.
121. Blackledge NP, Rose NR & Klose RJ Targeting Polycomb systems to regulate gene expression: modifications to a complex story. Nat. Rev. Mol. Cell Biol 16, 643–649 (2015). [PubMed: 26420232]
122. Margueron R & Reinberg D The Polycomb complex PRC2 and its mark in life. Nature 469, 343–349 (2011). [PubMed: 21248841]
123. Voigt P, Tee WW & Reinberg D A double take on bivalent promoters. Genes Dev. 27, 1318–1338 (2013). [PubMed: 23788621]
124. Hosogane M, Funayama R, Shirota M & Nakayama K Lack of transcription triggers H3K27me3 accumulation in the gene body. Cell Rep. 16, 696–706 (2016). [PubMed: 27396330]
125. Kaneko S, Son J, Bonasio R, Shen SS & Reinberg D Nascent RNA interaction keeps PRC2 activity poised and in check. Genes Dev. 28, 1983–1988 (2014). [PubMed: 25170018] This study reveals the prevalent interaction of PRC2 with nascent RNA.
126. Riising EM et al. Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. Mol. Cell 55, 347–360 (2014). [PubMed: 24999238]
127. Bernstein BE et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125, 315–326 (2006). [PubMed: 16630819]
128. Marks H et al. The transcriptional and epigenomic foundations of ground state pluripotency. Cell 149, 590–604 (2012). [PubMed: 22541430]
129. Theunissen TW et al. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell 15, 471–487 (2014). [PubMed: 25090446]
130. Chu C et al. Systematic discovery of Xist RNA binding proteins. Cell 161, 404–416 (2015). [PubMed: 25843628]
131. Wei C et al. RBFox2 binds nascent RNA to globally regulate Polycomb complex 2 targeting in mammalian genomes. Mol. Cell 62, 875–889 (2016). [PubMed: 27211866] This study reveals that the RBP RBFox2 modulates global interaction between PRC2 and nascent RNA, which is particularly important for maintaining bivalent promoters.
132. Margueron R et al. Role of the polycomb protein EED in the propagation of repressive histone marks. Nature 461, 762–767 (2009). [PubMed: 19767730]
133. Xu K et al. EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. Science 338, 1465–1469 (2012). [PubMed: 23239736]
134. Youmans DT, Schmidt JC & Cech TR Live-cell imaging reveals the dynamics of PRC2 and recruitment to chromatin by SUZ12-associated subunits. Genes Dev. 32, 794–805 (2018). [PubMed: 29891558]

135. Grewal SI & Jia S Heterochromatin revisited. Nat. Rev. Genet 8, 35–46 (2007). [PubMed: 17173056]
136. Holoch D & Moazed D Small-RNA loading licenses Argonaute for assembly into a transcriptional silencing complex. Nat. Struct. Mol. Biol 22, 328–335 (2015). [PubMed: 25730778]
137. Tschiersch B et al. The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3–9 combines domains of antagonistic regulators of homeotic gene complexes. EMBO J. 13, 3822–3831 (1994). [PubMed: 7915232]
138. Verdel A et al. RNAi-mediated targeting of heterochromatin by the RITS complex. Science 303, 672–676 (2004). [PubMed: 14704433]
139. Mutazono M et al. The intron in centromeric noncoding RNA facilitates RNAi-mediated formation of heterochromatin. PLOS Genet. 13, e1006606 (2017). [PubMed: 28231281]
140. Ragunathan K, Jih G & Moazed D Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment. Science 348, 1258699 (2015). [PubMed: 25831549]
141. Akhtar A, Zink D & Becker PB Chromodomains are protein-RNA interaction modules. Nature 407, 405–409 (2000). [PubMed: 11014199]
142. Shirai A et al. Impact of nucleic acid and methylated H3K9 binding activities of Suv39h1 on its heterochromatin assembly. eLife 6, e25317 (2017). [PubMed: 28760201]
143. Muerdter F et al. A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in *Drosophila*. Mol. Cell 50, 736–748 (2013). [PubMed: 23665228]
144. Yu Y et al. Panoramix enforces piRNA-dependent cotranscriptional silencing. Science 350, 339–342 (2015). [PubMed: 26472911]
145. Brennecke J et al. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell 128, 1089–1103 (2007). [PubMed: 17346786]
146. Johnson WL & Straight AF RNA-mediated regulation of heterochromatin. Curr. Opin. Cell Biol 46, 102–109 (2017). [PubMed: 28614747]
147. Nishibuchi G & Dejardin J The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. Chromosome Res. 25, 77–87 (2017). [PubMed: 28078514]
148. Lu J & Gilbert DM Proliferation-dependent and cell cycle regulated transcription of mouse pericentric heterochromatin. J. Cell Biol 179, 411–421 (2007). [PubMed: 17984319]
149. Allo M et al. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. Nat. Struct. Mol. Biol 16, 717–724 (2009). [PubMed: 19543290]
150. Cernilogar FM et al. Chromatin-associated RNA interference components contribute to transcriptional regulation in *Drosophila*. Nature 480, 391–395 (2011). [PubMed: 22056986]
151. Huang V et al. Ago1 Interacts with RNA polymerase II and binds to the promoters of actively transcribed genes in human cancer cells. PLOS Genet. 9, e1003821 (2013). [PubMed: 24086155]
152. Benhamed M, Herbig U, Ye T, Dejean A & Bischof O Senescence is an endogenous trigger for microRNA-directed transcriptional gene silencing in human cells. Nat. Cell Biol 14, 266–275 (2012). [PubMed: 22366686]
153. Wan G et al. Spatiotemporal regulation of liquid-like condensates in epigenetic inheritance. Nature 557, 679–683 (2018). [PubMed: 29769721]
154. Greenstein RA et al. Noncoding RNA-nucleated heterochromatin spreading is intrinsically labile and requires accessory elements for epigenetic stability. eLife 7, e32948 (2018). [PubMed: 30020075]
155. Yu R, Wang X & Moazed D Epigenetic inheritance mediated by coupling of RNAi and histone H3K9 methylation. Nature 558, 615–619 (2018). [PubMed: 29925950] This work reveals that endogenous siRNAs can act in trans to target a homologous euchromatic region to silence gene expression in fission yeast.
156. Lee NN et al. Mtr4-like protein coordinates nuclear RNA processing for heterochromatin assembly and for telomere maintenance. Cell 155, 1061–1074 (2013). [PubMed: 24210919]

157. Yamanaka S et al. RNAi triggered by specialized machinery silences developmental genes and retrotransposons. Nature 493, 557–560 (2013). [PubMed: 23151475]

158. Floyd JA et al. A natural allele of Nxf1 suppresses retrovirus insertional mutations. Nat. Genet 35, 221–228 (2003). [PubMed: 14517553]

159. Concepcion D, Ross KD, Hutt KR, Yeo GW & Hamilton BA Nxf1 natural variant E610G is a semi-dominant suppressor of IAP-induced RNA processing defects. PLOS Genet. 11, e1005123 (2015). [PubMed: 25835743]

160. Kato M et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell 149, 753–767 (2012). [PubMed: 22579281] This study first demonstrates that the LCDs in various RBPs are able to drive phase separation, providing a conceptual framework for the formation of membrane-less cellular condensates.

161. Wang M et al. Stress-induced low complexity RNA activates physiological amyloidogenesis. Cell Rep. 24, 1713–1721 (2018). [PubMed: 30110628]

162. Boisvert FM, van Koningsbruggen S, Navascues J & Lamond AI The multifunctional nucleolus. Nat. Rev. Mol. Cell Biol 8, 574–585 (2007). [PubMed: 17519961]

163. Spector DL & Lamond AI Nuclear speckles. Cold Spring Harb. Perspect. Biol 3, a000646 (2011). [PubMed: 20926517]

164. Lamond AI & Spector DL Nuclear speckles: a model for nuclear organelles. Nat. Rev. Mol. Cell Biol 4, 605–612 (2003). [PubMed: 12923522]

165. Shevtsov SP & Dundr M Nucleation of nuclear bodies by RNA. Nat. Cell Biol 13, 167–173 (2011). [PubMed: 21240286] This study demonstrates the ability of specific RNAs to induce de novo formation of nuclear bodies, including Cajal bodies, nuclear speckles and paraspeckles.

166. Feric M et al. Coexisting liquid phases underlie nucleolar subcompartments. Cell 165, 1686–1697 (2016). [PubMed: 27212236] This work demonstrates that multiple subcompartments of nucleoli are immiscible liquid phases due to different droplet surface tension and that such immiscibility is conferred by key RNA-binding domains of specific nucleolar proteins to facilitate sequential rRNA processing.

167. Falahati H, Pelham-Webb B, Blythe S & Wieschaus E Nucleation by rRNA dictates the precision of nucleolus assembly. Curr. Biol 26, 277–285 (2016). [PubMed: 26776729]

168. Zhang L, Wu C, Cai G, Chen S & Ye K Stepwise and dynamic assembly of the earliest precursors of small ribosomal subunits in yeast. Genes Dev. 30, 718–732 (2016). [PubMed: 26980190]

169. Xing YH et al. SLERT regulates DDX21 rings associated with Pol I transcription. Cell 169, 664–678 (2017). [PubMed: 28475895]

170. Li D et al. Activity dependent LoNA regulates translation by coordinating rRNA transcription and methylation. Nat. Commun 9, 1726 (2018). [PubMed: 29712923]

171. Dundr M & Misteli T Biogenesis of nuclear bodies. Cold Spring Harb. Perspect. Biol 2, a000711 (2010). [PubMed: 21068152]

172. Mao YS, Zhang B & Spector DL Biogenesis and function of nuclear bodies. Trends Genet. 27, 295–306 (2011). [PubMed: 21680045]

173. Nizami Z, Deryusheva S & Gall JG The Cajal body and histone locus body. Cold Spring Harb. Perspect. Biol 2, a000653 (2010). [PubMed: 20504965]

174. Shopland LS et al. Replication-dependent histone gene expression is related to Cajal body (CB) association but does not require sustained CB contact. Mol. Biol. Cell 12, 565–576 (2001). [PubMed: 11251071]

175. Frey MR & Matera AG RNA-mediated interaction of Cajal bodies and U2 snRNA genes. J. Cell Biol 154, 499–509 (2001). [PubMed: 11489914]

176. Wang Q et al. Cajal bodies are linked to genome conformation. Nat. Commun 7, 10966 (2016). [PubMed: 26997247]

177. Mourelatos Z, Abel L, Yong J, Kataoka N & Dreyfuss G SMN interacts with a novel family of hnRNP and spliceosomal proteins. EMBO J 20, 5443–5452 (2001). [PubMed: 11574476]

178. Yong J, Kasim M, Bachorik JL, Wan L & Dreyfuss G Gemin5 delivers snRNA precursors to the SMN complex for snRNP biogenesis. Mol. Cell 38, 551–562 (2010). [PubMed: 20513430]

179. Machyna M, Neugebauer KM & Stanek D Coilin: the first 25 years. RNA Biol. 12, 590–596 (2015). [PubMed: 25970135]
180. Raska I et al. Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies. Exp. Cell Res 195, 27–37 (1991). [PubMed: 2055273]
181. Dundr M et al. Actin-dependent intranuclear repositioning of an active gene locus in vivo. J. Cell Biol 179, 1095–1103 (2007). [PubMed: 18070915]
182. Platani M, Goldberg I, Lamond AI & Swedlow JR Cajal body dynamics and association with chromatin are ATP-dependent. Nat. Cell Biol 4, 502–508 (2002). [PubMed: 12068306]
183. So BR et al. A U1 snRNP-specific assembly pathway reveals the SMN complex as a versatile hub for RNP exchange. Nat. Struct. Mol. Biol 23, 225–230 (2016). [PubMed: 26828962]
184. Lin Y, Protter DS, Rosen MK & Parker R Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. Mol. Cell 60, 208–219 (2015). [PubMed: 26412307]
185. Protter DSW et al. Intrinsically disordered regions can contribute promiscuous interactions to RNP granule assembly. Cell Rep. 22, 1401–1412 (2018). [PubMed: 29425497]
186. Nakagawa S et al. Malat1 is not an essential component of nuclear speckles in mice. RNA 18, 1487–1499 (2012). [PubMed: 22718948]
187. Tripathi V et al. SRSF1 regulates the assembly of pre-mRNA processing factors in nuclear speckles. Mol. Biol. Cell 23, 3694–3706 (2012). [PubMed: 22855529]
188. Shin Y et al. Liquid nuclear condensates mechanically sense and restructure the genome. Cell 175, 1481–1491 (2018). [PubMed: 30500535]
189. Misteli T, Caceres JF & Spector DL The dynamics of a pre-mRNA splicing factor in living cells. Nature 387, 523–527 (1997). [PubMed: 9168118]
190. Hnisz D, Shrinivas K, Young RA, Chakraborty AK & Sharp PAA Phase separation model for transcriptional control. Cell 169, 13–23 (2017). [PubMed: 28340338] This review/hypothesis article proposes a phase separation model to establish multivalent interactions for transcriptional control.
191. Boija A et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. Cell 175, 1842–1855 (2018). [PubMed: 30449618]
192. Lu H et al. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. Nature 558, 318–323 (2018). [PubMed: 29849146]
193. Boehning M et al. RNA polymerase II clustering through carboxy-terminal domain phase separation. Nat. Struct. Mol. Biol 25, 833–840 (2018). [PubMed: 30127355]
194. Ying Y et al. Splicing activation by Rbfox requires self-aggregation through its tyrosine-rich domain. Cell 170, 312–323 (2017). [PubMed: 28708999] This work shows that the LCD in RBFox proteins mediates higher-order assembly of a large complex of splicing regulators that are important for differential splicing control and localization in nuclear speckles.
195. Gueroussov S et al. Regulatory expansion in mammals of multivalent hnRNP assemblies that globally control alternative splicing. Cell 170, 324–339 (2017). [PubMed: 28709000]
196. Clemson CM et al. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Mol. Cell 33, 717–726 (2009). [PubMed: 19217333]
197. Stanek D & Fox AH Nuclear bodies: news insights into structure and function. Curr. Opin. Cell Biol 46, 94–101 (2017). [PubMed: 28577509]
198. Hennig S et al. Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles. J. Cell Biol 210, 529–539 (2015). [PubMed: 26283796]
199. Maharana S et al. RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science 360, 918–921 (2018). [PubMed: 29650702]
200. Jiang L et al. NEAT1 scaffolds RNA-binding proteins and the Microprocessor to globally enhance pri-miRNA processing. Nat. Struct. Mol. Biol 24, 816–824 (2017). [PubMed: 28846091] This work elucidates a key biological function of nuclear paraspeckles where the lncRNA *NEAT1* helps to organize a large assembly of RBPs that collectively bind to almost all expressed pri-miRNAs, as well as the microprocessor complex through pseudo miR-612 to enhance global miRNA biogenesis.

201. Kim YK & Kim VN Processing of intronic microRNAs. EMBO J. 26, 775–783 (2007). [PubMed: 17255951]
202. Strom AR et al. Phase separation drives heterochromatin domain formation. Nature 547, 241–245 (2017). [PubMed: 28636597]
203. Larson AG & Narlikar GJ The role of phase separation in heterochromatin formation, function, and regulation. Biochemistry 57, 2540–2548 (2018). [PubMed: 29644850]
204. Bandaria JN, Qin P, Berk V, Chu S & Yildiz A Shelterin protects chromosome ends by compacting telomeric chromatin. Cell 164, 735–746 (2016). [PubMed: 26871633]
205. Lallemand-Breitenbach V & de The H PML nuclear bodies. Cold Spring Harb. Perspect. Biol 2, a000661 (2010). [PubMed: 20452955]
206. Mannen T, Yamashita S, Tomita K, Goshima N & Hirose T The Sam68 nuclear body is composed of two RNase-sensitive substructures joined by the adaptor HNRNPL. J. Cell Biol 214, 45–59 (2016). [PubMed: 27377249]
207. Saito K et al. Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in *Drosophila*. Genes Dev. 24, 2493–2498 (2010). [PubMed: 20966047]
208. Olivieri D, Sykora MM, Sachidanandam R, Mechtler K & Brennecke J An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in *Drosophila*. EMBO J. 29, 3301–3317 (2010). [PubMed: 20818334]
209. Cremer T & Cremer C Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat. Rev. Genet 2, 292–301 (2001). [PubMed: 11283701]
210. Cremer T & Cremer M Chromosome territories. Cold Spring Harb. Perspect. Biol 2, a003889 (2010). [PubMed: 20300217]
211. Bickmore WA The spatial organization of the human genome. Annu. Rev. Genom. Hum. Genet 14, 67–84 (2013).
212. Jain A & Vale RD RNA phase transitions in repeat expansion disorders. Nature 546, 243–247 (2017). [PubMed: 28562589] This study reveals that particular RNAs are able to undergo phase transition, linking sequence-specific gelation of RNAs to the aetiology of neurodegenerative diseases.
213. Langdon EM et al. mRNA structure determines specificity of a polyQ-driven phase separation. Science 360, 922–927 (2018). [PubMed: 29650703]
214. Guo L et al. Nuclear-import receptors reverse aberrant phase transitions of RNA-binding proteins with prion-like domains. Cell 173, 677–692 (2018). [PubMed: 29677512]
215. Hofweber M et al. Phase separation of FUS is suppressed by its nuclear import receptor and arginine methylation. Cell 173, 706–719 (2018). [PubMed: 29677514]
216. Qamar S et al. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-pi interactions. Cell 173, 720–734 (2018). [PubMed: 29677515]
217. Yoshizawa T et al. Nuclear import receptor inhibits phase separation of FUS through binding to multiple sites. Cell 173, 693–705 (2018). [PubMed: 29677513]
218. Mikhaleva S & Lemke EA Beyond the transport function of import receptors: what's all the FUS about? Cell 173, 549–553 (2018). [PubMed: 29677508]
219. van Steensel B & Belmont AS Lamina-associated domains: links with chromosome architecture, heterochromatin, and gene repression. Cell 169, 780–791 (2017). [PubMed: 28525751]
220. Pederson T Thinking about a nuclear matrix. J. Mol. Biol 277, 147–159 (1998). [PubMed: 9514762]
221. Xiao R et al. Nuclear matrix factor hnRNP U/SAF-A exerts a global control of alternative splicing by regulating U2 snRNP maturation. Mol. Cell 45, 656–668 (2012). [PubMed: 22325991]
222. Fan H et al. The nuclear matrix protein HNRNPU maintains 3D genome architecture globally in mouse hepatocytes. Genome Res. 28, 192–202 (2018). [PubMed: 29273625]
223. Nozawa RS et al. SAF-A regulates interphase chromosome structure through oligomerization with chromatin-associated RNAs. Cell 169, 1214–1227 (2017). [PubMed: 28622508]
224. Hendrickson DG, Kelley DR, Tenen D, Bernstein B & Rinn JL Widespread RNA binding by chromatin-associated proteins. Genome Biol. 17, 28 (2016). [PubMed: 26883116]

225. Shi Y Mechanistic insights into precursor messenger RNA splicing by the spliceosome. Nat. Rev. Mol. Cell Biol 18, 655–670 (2017). [PubMed: 28951565]

226. Martin RM, Rino J, Carvalho C, Kirchhausen T & Carmo-Fonseca M Live-cell visualization of pre-mRNA splicing with single-molecule sensitivity. Cell Rep. 4, 1144–1155 (2013). [PubMed: 24035393]

227. Maxwell ES & Fournier MJ The small nucleolar RNAs. Annu. Rev. Biochem 64, 897–934 (1995). [PubMed: 7574504]

228. Cassidy SB, Schwartz S, Miller JL & Driscoll DJ Prader-Willi syndrome. Genet. Med 14, 10–26 (2012). [PubMed: 22237428]

229. Kishore S & Stamm S The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. Science 311, 230–232 (2006). [PubMed: 16357227]

230. Yin QF et al. Long noncoding RNAs with snoRNA ends. Mol. Cell 48, 219–230 (2012). [PubMed: 22959273]

231. Flynn RA et al. 7SK-BAF axis controls pervasive transcription at enhancers. Nat. Struct. Mol. Biol 23, 231–238 (2016). [PubMed: 26878240]

232. Wang J, Jia ST & Jia S New insights into the regulation of heterochromatin. Trends Genet. 32, 284–294 (2016). [PubMed: 27005444]

Different classes of chromatin-associated RNAs

The RNA-chromatin interactome of mammalian cells generated by global RNA interaction with DNA sequencing (GRID-seq) and that of *Drosophila melanogaster* produced by both GRID-seq and chromatin-associated RNA sequencing (ChAR-seq) provide a general landscape of chromatin-associated RNAs that belong to several important classes.

Pre-mRNAs

Nascent protein-coding RNAs interact with chromatin predominately at their sites of transcription, and the interaction levels are proportional to their levels of transcription[80]. These findings are in line with the widespread association of both coding and non-coding transcripts with multiple chromatin remodellers[224].

snRNAs

Among small nuclear RNAs (snRNAs) functioning in pre-mRNA splicing, u2 snRNA and its variants are most prevalent on chromatin (similarly, u6 snRNA is also quite abundant (X.l. and X.-D.F., unpublished observations)). Both u2 and u6 snRNAs constitute the catalytic core of the spliceosome[225], indicating that splicing takes place in the proximity of chromatin, consistent with real-time analysis of intron turnover[226].

snoRNAs

A few small nucleolar RNAs (snoRNAs) have been well characterized for their functions in guiding ribosomal RNA (rRNA) modifications[227], but the function of most other expressed snoRNAs from multicellular organisms remains unknown. The detection of various snoRNAs on chromatin raises an intriguing possibility that some of these snoRNAs may participate in epigenetic control of gene expression by modifying other chromatin-associated RNAs, which represents a new direction in which to explore potential regulatory roles of snoRNAs in the genome. A cluster of snoRNAs has been linked to diseases in humans, such as Prader-Willi syndrome[228]. A specific snoRNA has been suggested to modulate alternative splicing via base-pairing interactions with an intronic element[229]. snoRNAs have also been detected at both ends of certain long non-coding RNAs (lncRNAs) that appear to function in sequestering splicing regulators as a way to regulate alternative splicing[230]. How these mechanisms might be related to specific diseases remains unclear.

lncRNAs

Multiple abundant lncRNAs have been detected on chromatin by one-to-many approaches, revealing potential roles of *MALAT1* and *NEAT1* in organizing gene clusters for transcription[75], *Xist* in spreading gene silencing on the inactive X chromosome[71] and distinct mechanisms for *7SK* RNA to modulate promoter and enhancer activities[231]. These lncRNAs have also been consistently detected by all-to-all strategies[79–81].

Repeat-derived RNAs

These RNAs are traditionally analysed by fluorescence in situ hybridization (FISH), but they are also readily detectable with all-to-all methods[79–81]. However, it has been challenging to precisely assign these repeat RNAs to specific genomic loci. Functional studies suggest that these repeat-derived RNAs play central roles in heterochromatin formation and/or maintenance[18,232].

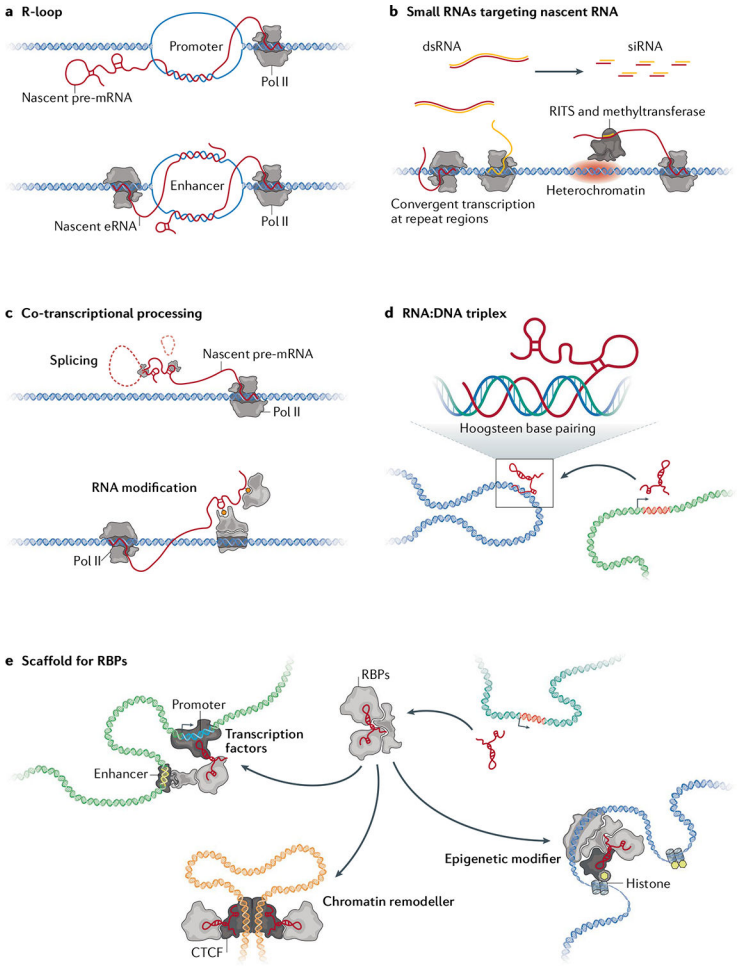

Fig. 1 |. Different modes of RNA-chromatin interactions.

a | Nascent RNAs can form R-loops, which predominately occur at promoters and enhancers (top). Given divergent transcription from many promoters and most enhancers, R-loops may consist of divergent transcripts (bottom) that could jointly contribute to the formation and dynamics of the resulting R-loops[8–10]. **b** | Nascent RNA is targeted by amplified small RNAs on its site of transcription, a process frequently associated with heterochromatin formation[15–20]. **c** | Nascent pre-mRNA remains on chromatin for co-transcriptional processing (top) and RNA modification (bottom), which may play important roles in feedback control of transcription[21,22,25,26,28]. **d** | *Trans*-acting RNA may form a triplex structure on double-stranded DNA using the Hoogsteen base-pairing rule, a strategy exploited by various long non-coding RNAs[29–31]. **e** | RNA serves as a scaffold to nucleate the formation of specific protein complexes on chromatin to control gene expression by promoting enhancer-promoter interactions, facilitating chromatin remodelling or directly contributing to specific chromatin modification activities[46–50,55,87,91]. CTCF, CCCTC-binding factor; dsRNA, double-stranded RNA; eRNA, enhancer RNA; Pol II, RNA polymerase II; RBP, RNA-binding protein; RITS, RNA-induced transcriptional silencing complex; siRNA, small interfering RNA.

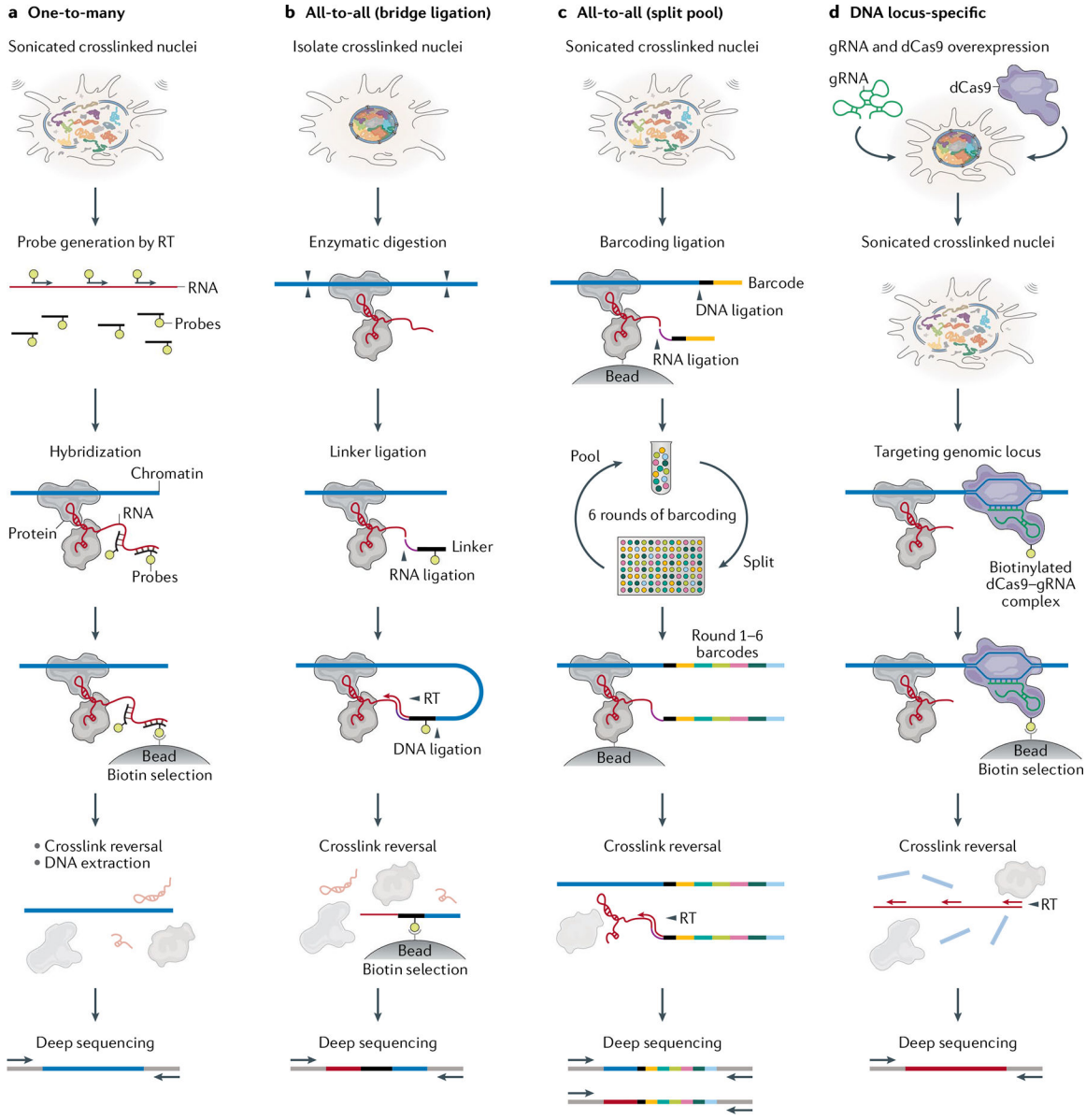

Fig. 2 |. Strategies for global analysis of RNA-chromatin interactions.

a | One-to-many approach based on RNA capture followed by deep sequencing of associated DNA fragments[69–71]. **b** | All-to-all strategy through in situ linker ligation to RNA and DNA on fixed nuclei[79–81]. **c** | A split-pool strategy to barcode isolated nuclear particles, thus enabling the profiling of RNA-RNA, RNA-DNA and DNA-DNA interactions[82]. **d** | Locus-specific capture either through a hybridized probe or using a nuclease-dead Cas9 (dCas9)-based method, which can be used for de novo identification of RNAs and proteins on specific genomic loci[84]. gRNA, guide RNA; RT, reverse transcription.

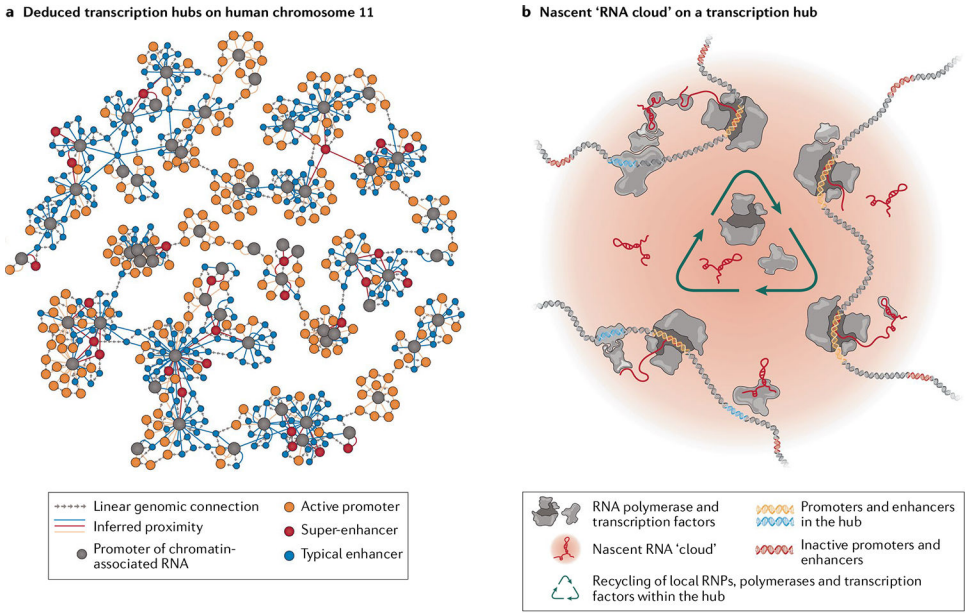

Fig. 3 |. Nascent-RNA-decorated transcription hubs and a model for RNA-nested hub formation. **a** | Deduced transcription hubs on a human chromosome (chromosome 11) in MDA-MB-231 breast cancer cells based on the RNA-chromatin interactome detected by global RNA interaction with DNA sequencing (GRID-seq)[80]. **b** | 'RNA cloud' on a transcription hub, suggesting a potential role of nascent RNA in networking regulatory DNA elements in the hub[99,100]. Also illustrated is the potential local recycling of RNA polymerase II and other transcription factors within the hub[110]. RNP, ribonucleoprotein particle.

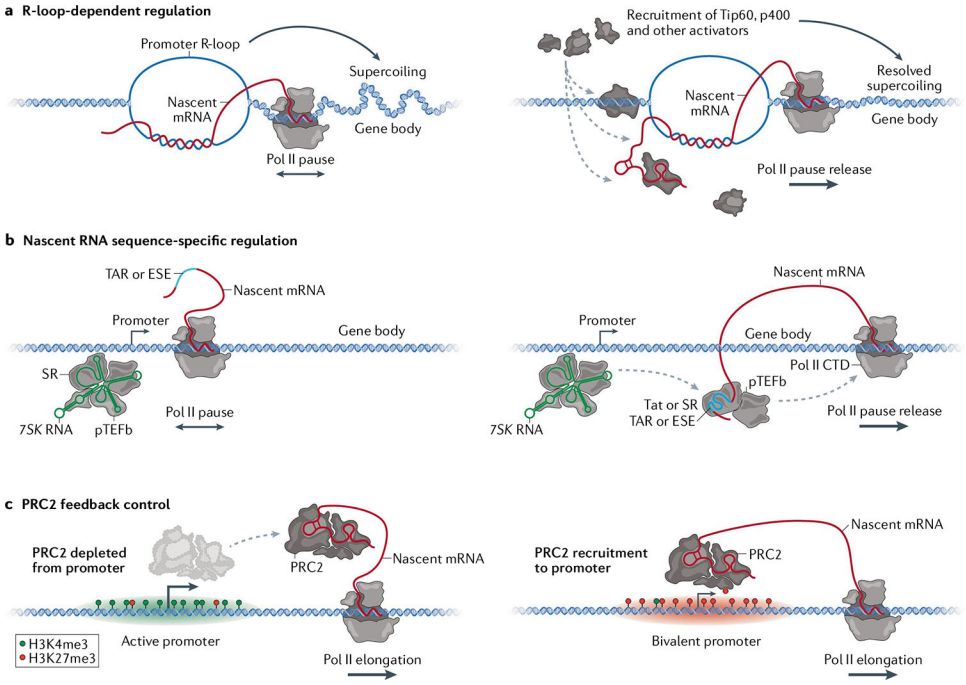

Fig. 4 |. Roles of chromatin-associated nascent RNA in transcriptional activation or repression. **a** | Formation of an R-loop in the promoter proximal region leads to transcriptional repression through pausing of RNA polymerase II (Pol II) (left)[9]. R-Loop-dependent recruitment of transcriptional co-activators may contribute to transcriptional pause release (right)[87]. **b** | Specific sequences from nascent RNA induce translocation of HIV Tat or cellular SRSF2 from the *7SK* complex to the nascent RNA (left), thereby relocating the Pol II C-terminal domain (CTD) kinase pTEFb from the inhibitory *7SK* complex to the paused Pol II complex to drive transcription pause release (right)[110–112]. **c** | Two competing models for nascent-RNA-dependent recruitment of Polycomb repressive complex 2 (PRC2) to provide a feedback control mechanism for gene expression. One model proposes that nascent RNA competes with nearby chromatin for PRC2 binding (left)[49,120,124–126], whereas the other suggests that PRC2 may use nascent RNA as a stepping stone to catalyse trimethylation of K27 on histone H3 (H3K27me3) (right)[74,130,131]. These two models appear to respectively apply to highly active promoters predominately associated with trimethylation of K4 on histone H3 (H3K4me3) and bivalent promoters marked with both H3K4me3 and H3K27me3. ESE, exonic splicing enhancer; SR, splicing factor characterized by repeats of serine (S) and arginine (R) amino acids; TAR, transactivation-responsive region.

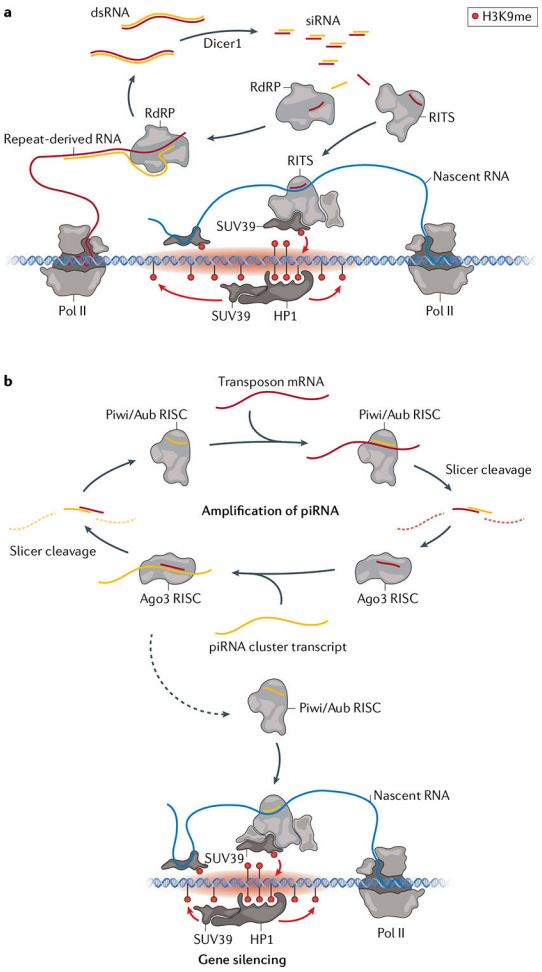

Fig. 5 |. RNA-dependent and RNA-independent feedback loops for establishing and locally spreading heterochromatin.

a | In fission yeast, RNA-dependent RNA polymerase (RdRP)-dependent amplification of local repeat-derived RNA occurs and then Dicer-processed small interfering RNAs (siRNAs) target nascent RNA to induce a network of interactions, resulting in the recruitment of SUV39 to deposit H3K9me on chromatin. This inhibitory histone mark next recruits heterochromatin protein 1 (HP1) and, together with SUV39, facilitates H3K9me spreading into adjacent genomic regions[15–18,139,140]. **b** | In *Drosophila melanogaster*, the primary transposon transcript is processed into Piwi-interacting RNAs (piRNAs), which then target nascent RNA to recruit SUV39 and HP1 to establish and spread methylation of K9 on histone H3 (H3K9me)[19,20,145–147]. Ago3, Argonaute 3; Aub, Aubergine; dsRNA, double-stranded RNA; Pol II, RNA polymerase II; RISC, RNA-induced silencing complex; RITS, RNA-induced transcriptional silencing complex.

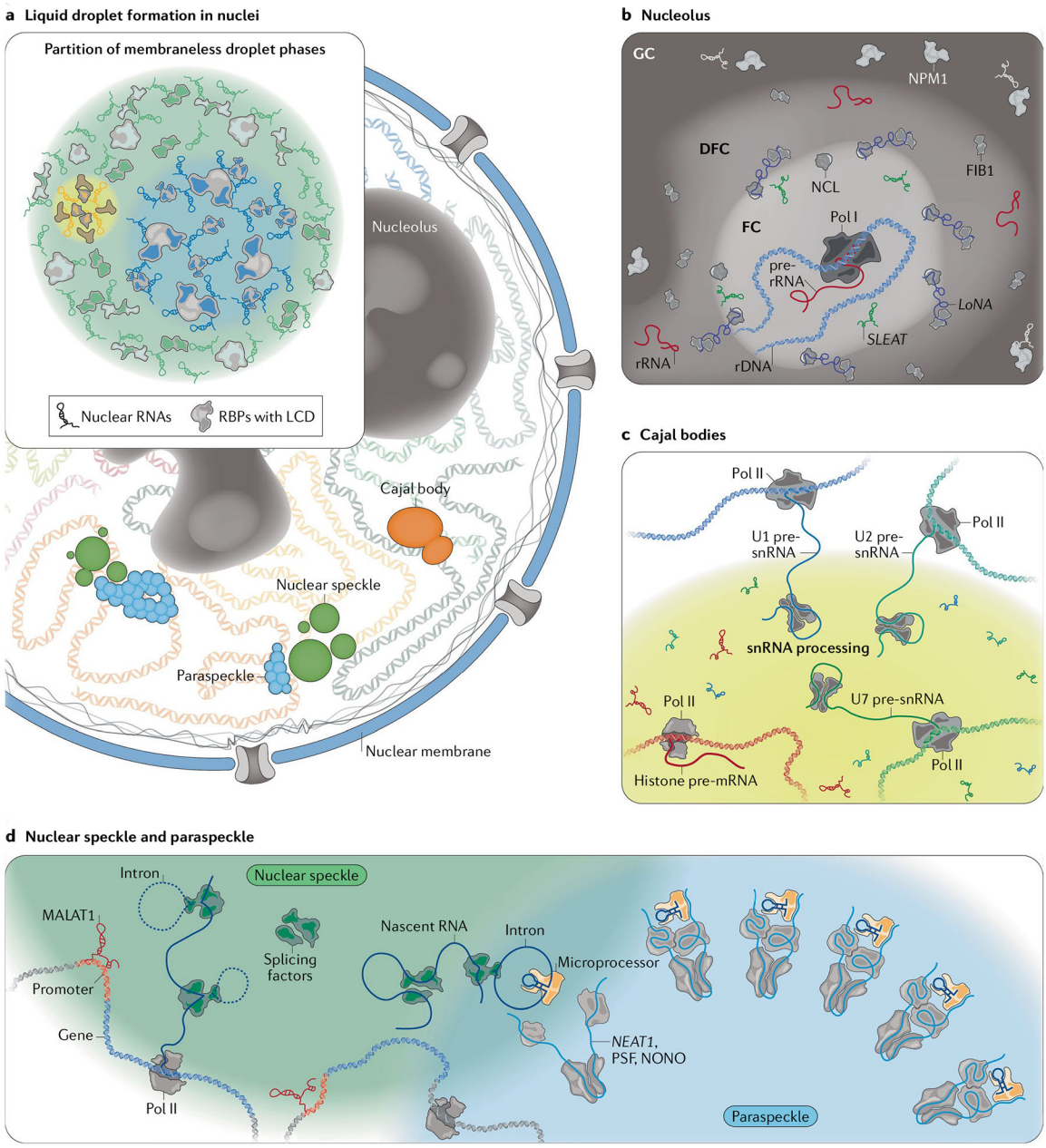

Fig. 6 |. RNAs in liquid-liquid phase separation and nuclear body formation.

a | Multivalent interactions in low-complexity domain (LCD)-containing proteins promote the formation of liquid droplets. Their interactions with RNAs may then partition liquid droplets into spatially distinct domains[6,7,160]. **b** | Scheme of dynamic liquid droplets in three different phases in the nucleolus for ribosomal DNA (rDNA) transcription (in the fibrillar centre (FC) domain), ribosomal RNA (rRNA) processing (in the dense fibrillar compartment (DFC) domain) and ribosome assembly (in the granule compartment (GC) domain)[162,166–170]. **c** | Both locally produced and *trans*-supplied small nuclear RNAs (snRNAs) and histone mRNAs drive the formation of liquid droplets in Cajal bodies[171–174]. **d** | Formation of nuclear speckles induced by co-transcriptional and post-transcriptional pre-

mRNA processing. Paraspeckles are known to form adjacently to nuclear speckles, and a large long non-coding RNA *NEAT1* is essential for the formation and maintenance of this nuclear subdomain. The connection between nuclear speckles and paraspeckles might result from a phase separation induced by primary microRNA (pri-miRNA) processing from various introns[163,196,197,201]. FIB1, fibrillarin; NPM1, nucleoplasmin; NCL, nucleolin; Pol I, RNA polymerase I; Pol II, RNA polymerase II; Pol III, RNA polymerase III; RBP, RNA-binding protein.