Guh et al. Journal of Biomedical Science (2020) 27:44
https://doi.org/10.1186/s12929-020-00640-3

MOST 科技部 Ministry of Science and Technology
The cost of publication in Journal of Biomedical Science is borne by the Ministry of Science and Technology, Taiwan.

Journal of Biomedical Science

REVIEW Open Access

Functions and properties of nuclear lncRNAs—from systematically mapping the interactomes of lncRNAs

Chia-Yu Guh, Yu-Hung Hsieh and Hsueh-Ping Chu*

Abstract

Protein and DNA have been considered as the major components of chromatin. But beyond that, an increasing number of studies show that RNA occupies a large amount of chromatin and acts as a regulator of nuclear architecture. A significant fraction of long non-coding RNAs (lncRNAs) prefers to stay in the nucleus and cooperate with protein complexes to modulate epigenetic regulation, phase separation, compartment formation, and nuclear organization. An RNA strand also can invade into double-stranded DNA to form RNA:DNA hybrids (R-loops) in living cells, contributing to the regulation of gene expression and genomic instability. In this review, we discuss how nuclear lncRNAs orchestrate cellular processes through their interactions with proteins and DNA and summarize the recent genome-wide techniques to study the functions of lncRNAs by revealing their interactomes in vivo.

Keywords: Long non-coding RNA, RNA-interactome, Epigenetics, Chromatin, Nuclear architecture, R-loops, Phase separation

Introduction

Over the past decade, intensive studies have demonstrated that long non-coding RNAs (lncRNAs) are key regulators in embryonic development [1], DNA damage responses [2], and human diseases such as neuronal disorders [3], heart dysfunction [4], and cancers [5, 6]. LncRNAs can act as effectors to direct biological processes. Unlike DNA, RNA is more mobile, flexible, able to self-fold into a distinct structure and interact with DNA or RNA by base pairing. Revealing the molecular mechanisms by which lncRNAs regulate those biological functions relies on the study of their interaction between DNA and proteins. Several methodologies have been developed to systematically identify RNA interacting chromatin, RNA interacting proteins and RNA-RNA interactions in vivo, thus uncovering the network of protein-RNA-DNA (Table 1).

Chromatin associated RNAs

The first study showed that RNA constitutes a significant fraction of chromatin date back over 50 years ago [41, 42]. Later studies provide evidence to support the idea that a nuclear matrix is made of insoluble proteins and RNA in an interphase nucleus [43]. Treatment of RNase A leads to clumping of chromatin onto nuclear lamina and nucleolus [44, 45], indicating an indispensable role of RNA in the nuclear architecture. However, identifying a specific RNA contributes to a specific phenotype in the nucleus had facing technical challenges before the 1990s. Recently, the burst of studies utilized next-generation sequencing to characterize the interactions of lncRNA and chromatin. While using lncRNA ablation, accumulating evidence has suggested that RNA-protein complexes conduct various functions such as the formation of chromatin compartment, gene regulation, and inter- or intra-chromosomal interactions [44, 45] (Fig. 1).

It has been reported that the CCCTC-binding transcription factor (CTCF), which promotes chromatin loop formation, binds to thousands of RNAs [46]. Depletion

* Correspondence: cchu2017@ntu.edu.tw
Institute of Molecular and Cellular Biology, National Taiwan University, No. 1 Sec. 4 Roosevelt Road, Taipei, Taiwan, Republic of China

© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Table 1 In vivo mapping RNA interactomes

Method	Cross-linking	Concept	Advances	Identifier	Ref
RNA-chromatin					
ChIRP-seq	Glutaraldehyde	Use biotinylated antisense oligos to pull down a targeted RNA with its associated DNA.	Robust elution by RNase A and RNase H.	DNA that is associated with a specific RNA	[7]
CHART-seq	Formaldehyde	Use biotinylated antisense oligos to pull down a targeted RNA with its associated DNA.	Apply RNase H to specifically elute RNA mediated interacting chromatin.	DNA that is associated with a specific RNA	[8, 9]
CHIRT-seq	Glutaraldehyde	A hybrid method of ChIRP and CHART.	Combination of glutaraldehyde fixation and RNase H elution.	DNA that is associated with a specific RNA	[10]
MARGI-seq	Formaldehyde	Use a linker to ligate RNA and DNA in proximity to form of RNA-linker-DNA.	Reveal all interactions between DNA and RNA.	All RNA-DNA contacted sequences	[11]
ChAR-seq	Formaldehyde	Use a linker to ligate RNA and DNA in proximity to form of RNA-linker-DNA.	Reveal all interactions between DNA and RNA.	All RNA-DNA contacted sequences	[12]
GRID-seq	Formaldehyde and disuccinimidyl glutarate (DSG)	Use a linker to ligate RNA and DNA in proximity to form of RNA-linker-DNA.	Strong crosslinking to reveal long-range interaction between DNA and RNA.	All RNA-DNA contacted sequences	[13]
HiChIRP-seq	Glutaraldehyde	Combine ChIRP and Hi-C. Use CLICK chemistry to conjugate a biotin for subsequent contact enrichment.	Characterize a specific RNA that involves in chromosomal interaction.	Chromosome conformation at a specific RNA associated sites	[14]
RNA-proteins					
CLIP-seq	UV irradiation (254 nm)	Pull down RNA-protein complexes by immunoprecipitation and perform reverse transcription.	Identify all RNAs that interact with a targeted protein.	RNA that binds to a specific protein	[15]
iCLIP-seq	UV irradiation (254 nm)	Pull down RNA-protein complexes by immunoprecipitation and perform reverse transcription.	A random barcode to mark individual cDNA molecules to solve the problems of PCR duplicates.	RNA that binds to a specific protein	[16]
PAR-CLIP-seq	Incorporate 4-thiouridine (4-SU) and 6-thioguanosine (6-SG) into nascent RNA. UV (365 nm)	Builds on CLIP. Incorporation of 4-SU or 6-SG results in U to C and G to A mutations respectively that allows mutational analysis to identify cross-linked sites.	Use 4-SU or 6-SG incorporation to increase the cross-linking efficiency.	RNA that binds to a specific protein	[17]
RAP-MS	UV irradiation (254 nm)	Use biotinylated antisense RNA probes to capture a specific RNA.	Identify direct RNA interacting proteins.	Proteins that bind to a specific RNA	[18]
ChIRP-MS	Formaldehyde	Use biotinylated antisense DNA probes to capture a specific RNA.	Identify direct and indirect RNA interacting proteins.	Proteins that bind to a specific RNA	[19]
iDRiP-MS	UV irradiation (254 nm)	Use biotinylated antisense DNA probes to capture a specific RNA.	Identify direct RNA interacting proteins.	Proteins that bind to a specific RNA	[20] [10]
RBR-ID	UV (312 nm) + 4-thiouridine (4-SU)	Comparison of 4-SU and non-4-SU treatments, an RNA-crosslinked peptide has a different mass.	Identify all proteins bound to RNAs.	All RNA binding proteins	[21]
RNA structure/ RNA-RNA interactions					
FragSeq	N/A	RNA is digested by P1 endonuclease. Nuclease probing.	Map P1 endonuclease digestion sites.	In vitro RNA structure	[22]
PARS	N/A	RNA is digested by RNase V1 or S1 to determine double stranded or single stranded regions. Nuclease probing.	Map RNase V1 or S1 digestion sites.	In vitro RNA structure	[23]
SHAPE-seq	Covalently modify RNA molecules in vitro.	SHAPE reagents (1 M7, NAI-N_3) modify RNAs.	Single nucleotide resolution; each RNA in the experiment is bar-coded.	In vitro RNA structure	[24] [25, 26]
icSHAPE-seq	Covalently modify RNA molecules in vivo	SHAPE reagent (NAI-N_3). Copper-free click chemistry, a biotin moiety is selectively and efficiently added to NAI-N_3-modified RNA.	Identify In vivo RNA structure.	In vivo RNA structure	[27]
DMS-seq (Structure-seq)	Covalently modify RNA molecules	Dimethyl sulphate (DMS) methylates the base-pairing faces of A and C of RNA in loops.	Nucleotide resolution. Map RNA structure in vivo.	In vivo RNA structure	[28]

Table 1 In vivo mapping RNA interactomes *(Continued)*

Method	Cross-linking	Concept	Advances	Identifier	Ref
COMRADES	Psoralen + UV irradiation (365 nm)	Pull down a specific RNA using biotinylated DNA oligos and perform proximity ligation.	Reveal RNA structures and interactions of a specific RNA in vivo.	In vivo RNA structures and interactions of a targeted RNA	[29]
CLASH	UV irradiation (254 nm)	Immunoprecipitation to enrich a specific RNA binding protein and perform linker ligation.	Find mRNA target sequences for miRNA.	RNA hybrids bound by a specific RNA-binding protein	[30] [31]
hiCLIP	UV irradiation (254 nm)	Immunoprecipitate RNA-protein complexes by using antibodies against a specific RNA-binding protein and ligate RNA duplexes in proximity.	Reveal RNA duplexes bound to a specific protein.	RNA duplexes bound by a specific RNA-binding protein	[32]
MARIO	Formaldehyde and EthylGlycol bis	Apply a biotinylated linker to ligate two RNA fragments in proximity.	Reveal all RNA-RNA interactions.	All RNA fragments in proximity	[33]
PARIS	Psoralen + UV irradiation (365 nm)	Purify RNA-duplexes by 2D gel and ligate two RNA fragments in proximity	Reveal all RNA-RNA interactions	All RNA duplexes	[34]
SPLASH	Psoralen + UV irradiation (365 nm)	Use biotinylated psoralen to crosslink RNA and perform proximity ligation.	Reveal all RNA-RNA interactions.	All RNA-RNA hybrids	[35]
RNA-DNA hybrids (R-loops)					
DRIP-seq	N/A	Pull down RNA/DNA hybrids using S9.6 antibody that specifically recognizes RNA/DNA hybrids.	Reveal DNA-RNA hybrids.	DNA that forms hybrids with RNA	[36, 37]
bisDRIP-seq	Bisulfite modification	Use bisulfite to convert cytosine residues into uracil residues within genomic DNA regions that contain single-stranded DNA. Enrich DNA/RNA hybrids by S9.6 IP.	Define the boundaries of the R-loop, high resolution.	Single-stranded DNA of R loops	[38]
R-ChIP	Formaldehyde	Use catalytic-dead RNase H to capture R loops, similar to ChIP.	Reveal DNA-RNA hybrids, high resolution.	RNase H target sites, R loops	[39]
DRIPc-seq	N/A	Builds on DRIP. Sequence RNA of DNA-RNA hybrids.	Reveal DNA-RNA hybrids, high resolution.	RNA of R loops	[40]

of steroid receptor *RNA* activator (SRA) noncoding RNA that is associated with CTCF, reduced CTCF-mediated insulator activity [47] at the *IGF2/H19* imprinting control region and increased *IGF2* expression. It was also reported that a lncRNA, Jpx, activates Xist RNA expression by evicting CTCF from binding to *Xist* promoter [48]. These results imply that lncRNAs in the nucleus cooperate with RNA-binding proteins to regulate the gene expression, perhaps through modulating the chromatin conformation.

It was proposed that lncRNA transcription guides genome organization [49]. For example, Oakes et al. demonstrated that inhibition of rRNA transcription leads to nucleolus disassembly [50]. Induction of the transcription of rRNA genes that are inserted into new chromosomal locations can generate nucleolus-like structures [51]. These results suggest that rRNA transcription could be responsible for the nucleolus organization. Lately, Nozawa et al. had suggested that chromatin-associated RNAs can form a chromatin mesh with scaffold attachment factor A (SAF-A), also known as heterogeneous ribonucleoprotein U (HNRNP-U) [52]. They showed that SAF-A oligomerization, which is dependent on ATP and RNA, drives chromatin decompaction, whereas its monomerization compacts large-scale chromatin organization. The global change of large-scale chromatin by depletion of SAF-A did not alter dramatically gene expression [52], nevertheless, it resulted in excessive DNA damages and genomic instability [52].

Xist RNA regulates gene silencing and chromatin conformation

Xist RNA is one of the well-studied lncRNAs involved in epigenetic regulation and gene silencing. It is a 17-kb lncRNA that is expressed from the inactive X-chromosome (Xi) [53], coating the entire X chromosome to repress gene expression through its ability to recruit repressive complexes such as polycomb complexes PRC1 and PRC2 [54–56]. In addition, Xist plays an important role in three-dimensional (3D) conformation and maintains the heterochromatin structure in Xi (Fig. 2). When Xist was depleted, topologically associated domains (TADs) were restored *in cis* and the Xi (inactive X chromosome) became Xa (active X chromosome) like conformation [20]. Lee et al. also showed that Xist repels positive chromatin factors such as BRG1 (also known as SMARCA4) and cohesin from the Xi [20, 57] to prevent

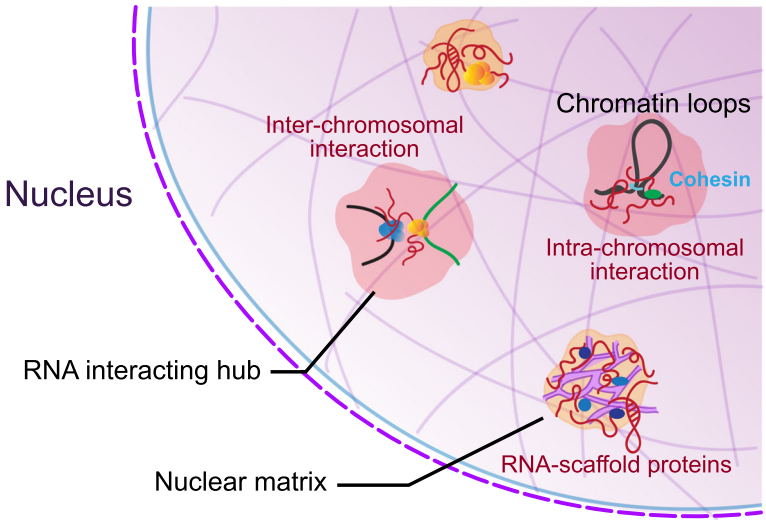

Fig. 1 lncRNA-interacting hub in the nucleus. LncRNAs serve as scaffolds for protein complexes and bring two or several distant DNA loci together. RNA involves the maintenance of nuclear architecture, facilitates chromatin looping, as well as directs the inter- or intra-chromosomal interactions

the formation of TADs and chromatin superloops. In a higher order of chromatin structure, mammalian chromosomes are organized into alternating "A/B compartments" in 3D conformation. Spatial compartments usually possess chromatin of similar states, with A compartments being actively transcribed and gene-rich, and B compartments being transcriptionally inactive and gene-poor [58, 59]. Remarkably, ablating SMCHD1 (known as an architectural protein) displayed another layer of compartments--S1/S2 [60] in Xi. Their results indicate that SMCDH1 binds to Xist and facilitates the folding of the S1/S2 compartments into a compartmentless (more compact) structure in Xi (Fig. 2). Altogether, robust evidence has shown that a lncRNA can regulate gene expression and chromatin 3D conformation through its abilities to recruit epigenetic factors, or to act as a decoy or a scaffold for protein complexes.

Enhancer RNAs promote chromatin looping

Enhancer RNAs were first discovered by two early studies through the genome-wide sequencing such as RNA-seq or ChIP-seq [61, 62]. They demonstrated that enhancers were occupied by RNA polymerase II (RNAP II) and transcribed into a class of ncRNAs termed enhancer RNAs (eRNAs). The epigenetic features of enhancers consist of histone 3 lysine 4 monomethylation (H3K4me1),

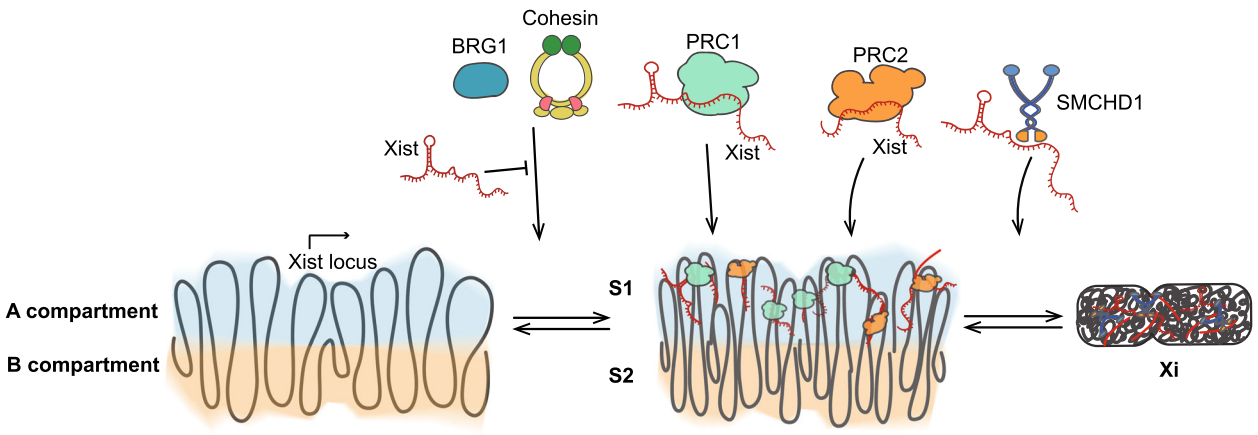

Fig. 2 Xist RNA tethers epigenetic regulators and contributes 3D conformation of the X chromosome. Xist RNA recruits repressive complexes such as polycomb complexes PRC1 and PRC2 to the inactive X chromosome (Xi) to facilitate heterochromatin formation, thus leading to gene silencing. On the other hand, Xist binds to BRG1 and cohesin and repels them from the Xi to prevent TAD (topologically associated domain) formation. Xist mediates the Xi folding in 3D space by tethering SMCHD1, which facilitates the merge of chromatin compartments. A/B compartments are first fused into "S1" and "S2" compartments, after SMCHD1 recruitment, further merging to compact Xi structure

histone 3 lysine 27 acetylation (H3K27ac), histone variants (H2AZ, H3.3), co-activators (mediator complex), and an open-chromatin architecture (DNase I hypersensitivity) [63]. The expression of enhancer RNAs could be regulated by various stimuli such as estrogen (ER) or androgen (AR) [64–66]. It is generally believed that eRNAs are unstable, transcribed quickly after induction, and degraded rapidly [62]. Several lines of evidence suggest that eRNAs can promote enhancer-promoter looping or facilitate RNA pol II loading, thus upregulating their target genes [65, 67]. For example, the gonadotropin hormone α-subunit gene is regulated by the eRNA in a cell-type-specific manner [68, 69]. The depletion of the eRNA results in a loss of interactions of the enhancers and promoters of the gonadotropin gene [68]. These results further support that lncRNAs enable to direct chromatin looping.

lncRNAs mediate interchromosomal interactions

Homologous chromosome pairing rarely occurs in somatic cells under normal growth conditions with only a few exceptions. Transvection was first observed in 1908 in *Drosophila*, where homologous chromosomes were closely paired in somatic nuclei [70]. It is an epigenetic phenomenon that can lead to either gene activation or repression. X chromosome pairing is one of the best-known examples of somatic homologous pairing in mammals [71]. Tsix, a lncRNA, is highly expressed in undifferentiated ES cells, and antagonizes Xist action. During ES differentiation in female cells, transient homologous pairing occurs, and the Xist RNA expression increases to initiate chromosome-wide silencing. It was reported that Tsix RNA mediates the homologous pairing between two X chromosomes, thus breaking epigenetic symmetry between the two X chromosomes, as well as driving the random choices for the selection of inactive or active X chromosomes [72, 73]. Moreover, another lncRNA derived from the ends of sex chromosomes, dubbed PAR-TERRA (telomeric repeat-containing RNA), facilitates the pairing by clustering the ends of the sex chromosomes, and creates a hub to constrain the DNA loci in 3D space [10]. Not limited to X chromosomes, several studies indicated that somatic allelic pairing also occurs in a number of autosomal loci, such as *Oct4* and various cytokine genes [74–77]. Although how interchromosomal pairing impacts gene expression and epigenetic regulation remains elusive, it has been proposed that the alignment of the two homologous alleles could allow biallelically bound transcription factors to redistribute onto one allele to achieve their lowest free energy state (the aggregated state) [72, 78, 79], thus inducing the transition from biallelic to monoallelic expression.

Methods for RNA-chromatin interactions

To map the RNA-associated chromatin in vivo, Chu et al. and Simon et al. utilized biotinylated antisense DNA oligo probes to capture a specific RNA that is associated with chromatin [7, 8] (Table 1 and Fig. 3), named ChIRP-seq (Chromatin isolation by RNA purification) and CHART-seq (Capture Hybridization Analysis of RNA Targets) respectively. They first fixed cells with either glutaraldehyde (ChIRP) or formaldehyde (CHART) and sheared chromatin into small pieces by physical sonication. To minimize the noises caused by the non-specific interactions of DNA probes and chromatin DNA or proteins, CHART includes an RNase H step to elute the RNA-chromatin complexes. This is to ensure that only lncRNAs-complexes that are targeted by DNA probes will be eluted by RNase H, which specifically degrades RNA of RNA-DNA hybrids. Later, a hybrid method was developed, called CHIRT-seq, which combines glutaraldehyde fixation and RNase H elution [10] to identify genomic binding sites for TERRA RNA. Accumulating studies have used ChIRP-seq to determine the genomic bindings of many lncRNAs, including NEAT1, MALAT1, HOTAIR and MEG3, TERC, LTR ERV-9 [80, 81]. CHART-seq had also successfully identified Xist, NEAT1, and MALAT1 genomic binding sites [8, 9].

Recently, several methods were developed to reveal all interactions between RNA and DNA in the nucleus, including MARGI-seq (Mapping RNA-genome interactions), ChAR-seq (Chromatin-Associated RNA sequencing), and GRID-seq (Global RNA Interactions with DNA) [11–13]. The idea of these methods is to ligate chromatin-associated RNAs with their target genomic sequences by proximity ligation using a linker, thus forming RNA-DNA chimeric fragments. These techniques revealed hundreds of chromatin-associated RNAs including previously known lncRNAs *and* a large set of non-coding RNAs that are bound to active promoters, enhancers, and super-enhancers in a tissue-specific manner [11–13]. All-to-all mapping ideally can discover all interactions of RNA-chromatin, however, the coverage depends on the ligation efficacy, the distances between RNA and genomic DNA, as well as the abundance of RNAs.

As discussed previously, long non-coding RNAs may serve as scaffolds to bring two or several distant DNA loci into spatial proximity. To understand how a specific RNA interacts with chromatin in a 3D space, Mumbach et al. developed a method, named HiChIRP, which combines Hi-C and ChIRP [14]. They incorporated azido-CTP into chromatin contacts, captured RNA-chromatin using biotinylated probes, conducted the copper-free dibenzocyclooctyne (DIBO) 'CLICK' chemistry to covalently conjugate a biotin for subsequent contacts, and sequenced the DNA contacts. They performed HiChIRP on 7SK small nuclear RNA (snRNA), telomerase RNA component (TERC) and lincRNA-EPS [14]. They found that thousands of loops were enriched on 7SK HiChIRP, and some of them in promoters and active regulatory elements. They also showed that TERC is not only associated with loops that formed between telomeres but also

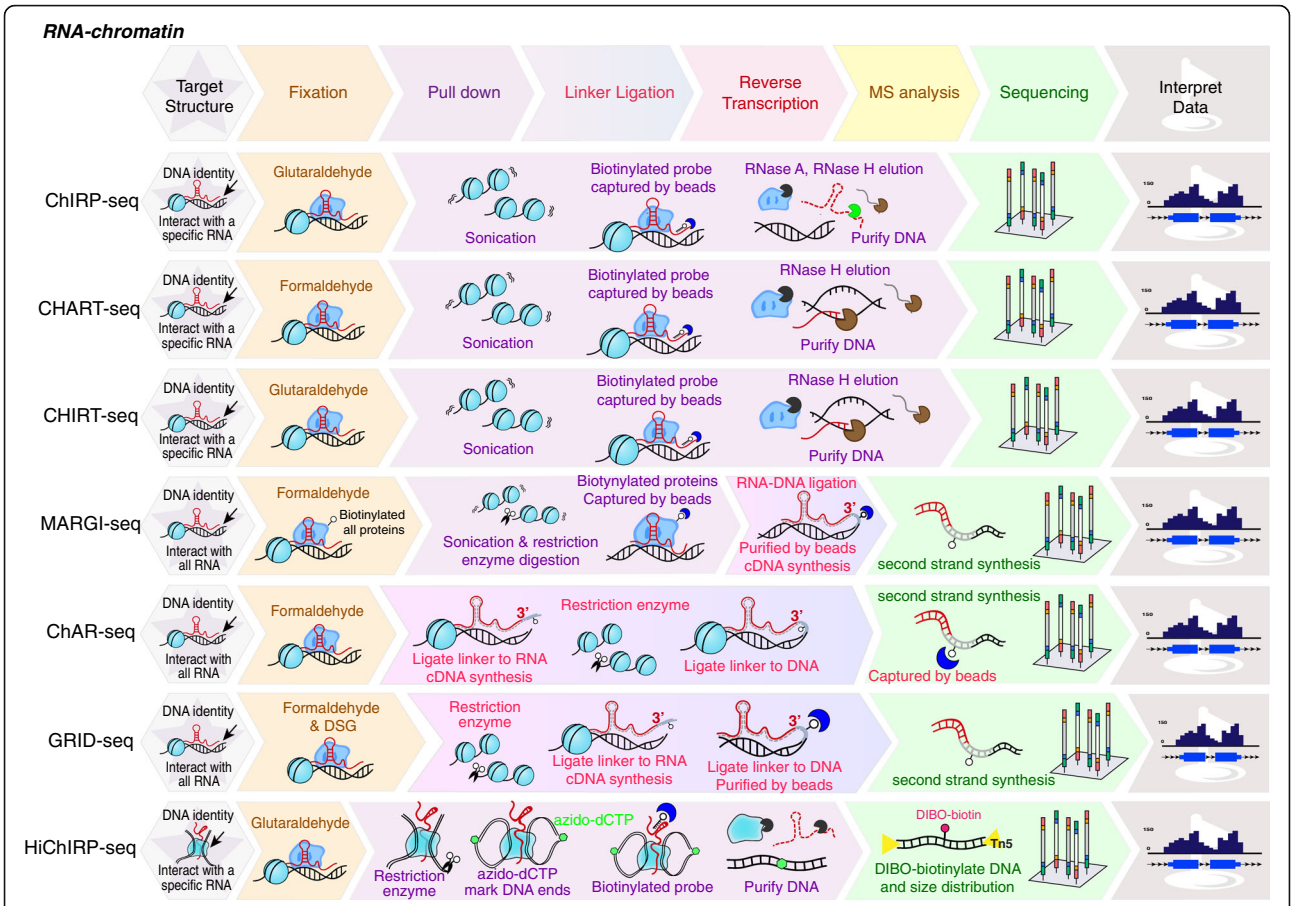

Fig. 3 Mapping RNA-chromatin interactions. In ChIRP-seq, CHART-seq, and CHIRT-seq, RNA associated chromatin complexes are crosslinked by formaldehyde or glutaraldehyde, captured by antisense oligos that target a specific lncRNA, and DNA fragments that are associated with RNA-protein complex are sequenced. For all-to-all interactions (MARGI-seq, ChAR-seq, and GRID-seq), a linker is ligated to connect DNA and RNA. DNA-RNA chimeras are sequenced. HiChIRP-seq combines Hi-C and ChIRP to identify the interacting chromatin (inter- and intra-chromatin interactions) of a specific lncRNA

with enhancer-promoter loops at some oncogenes, implying a role of TERC outside of telomeres. Therefore, their results provide insights into how lncRNAs mediate inter- or intra-chromatin looping.

Methods for RNA-protein interactions

RNA binding proteins (RBPs) are associated with a large number of human disorders, such as autoimmune and neurologic diseases [82–84]. Remarkable examples include FMRP in the Fragile-X mental retardation protein [83], the neuron-specific Nova and Hu proteins in the paraneoplastic neurologic degenerations [84] and the small nuclear ribonucleoproteins (snRNPs) in systemic lupus erythematosus [82]. To identify RNAs that interact with these proteins in vivo, Ule et al. combined UV cross-linking with immunoprecipitation (CLIP) to pull down RNA-protein complexes [15] (Table 1 and Fig. 4), and the captured RNAs are sequenced. Because CLIP relies on reverse transcription to pass over residual amino acids that covalently attach to the RNA at the cross-link site, cDNAs tend to prematurely truncate immediately before the cross-link nucleotide [15]. Later on, they resolved this problem by PCR amplification of truncated cDNAs via self-circularization of cDNAs, and developed individual-nucleotide resolution UV cross-linking and immunoprecipitation (iCLIP) to precisely map protein–RNA interactions [16]. Due to the fact of low efficiency of RNA-protein crosslinking by UV 254 nm in CLIP, Tuchi et al. developed a method, named PAR-CLIP (photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation) [17], to improve the crosslinking efficiency by incorporating 4-thiouridine (4SU) into nascent RNA transcripts in living cells. The cells were then irradiated by UV 365 nm to induce efficient crosslinking of photoreactive nucleoside (4SU)-labeled RNAs to interacting proteins (Table 1). PAR-CLIP has been applied to identify the transcriptome-wide binding sites of several RBPs and microRNA-containing ribonucleoprotein complexes.

Recently, researchers have elaborated on the methods for pulling down RNA-proteins complexes via antisense oligos that are complementary to targeted RNA sequences.

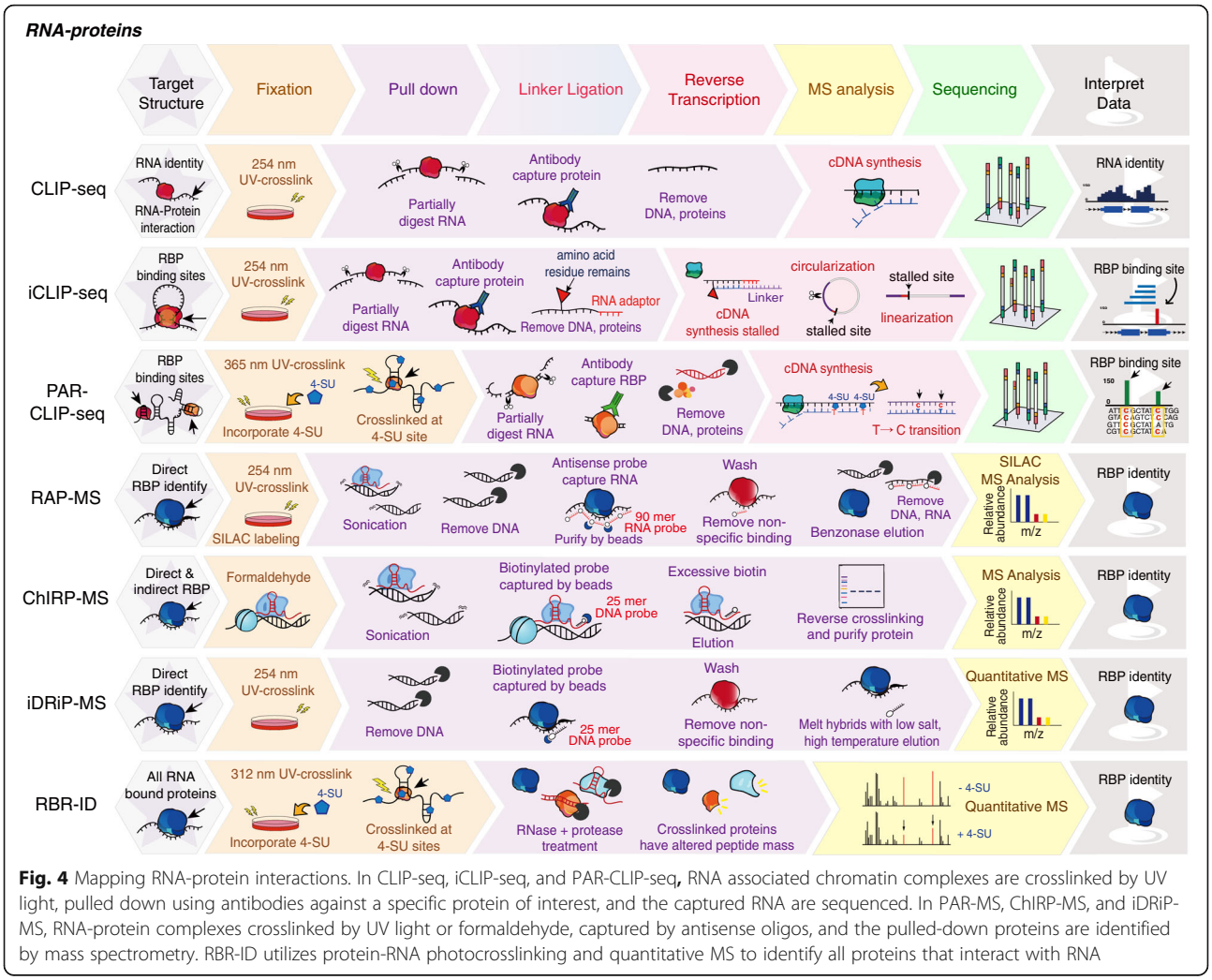

Fig. 4 Mapping RNA-protein interactions. In CLIP-seq, iCLIP-seq, and PAR-CLIP-seq, RNA associated chromatin complexes are crosslinked by UV light, pulled down using antibodies against a specific protein of interest, and the captured RNA are sequenced. In PAR-MS, ChIRP-MS, and iDRiP-MS, RNA-protein complexes crosslinked by UV light or formaldehyde, captured by antisense oligos, and the pulled-down proteins are identified by mass spectrometry. RBR-ID utilizes protein-RNA photocrosslinking and quantitative MS to identify all proteins that interact with RNA

Therefore, this type of capture is not limited by the requirement of antibodies against the proteins of interest for immunoprecipitation. In 2015, three groups established RNA-centric capture methods, including iDRiP-MS (*i*dentification of *d*irect *R*NA *i*nteracting *p*roteins), RAP-MS (RNA antisense purification coupled with mass spectrometry) and ChIRP-MS (comprehensive identification of RNA-binding proteins by mass spectrometry) to determine Xist RNA interacting proteins [18–20] (Table 1 and Fig. 4). In these three studies, cells were initially crosslinked to preserve RNA-protein interactions, and RNA-protein complexes were further purified by biotinylated antisense oligos along with highly denaturing purification conditions. ChIRP-MS utilizes formaldehyde to crosslink RNA-protein complexes, whereas RAP-MS and iDRiP-MS apply UV light for crosslinking (Table 1 and Fig. 4). Given that UV light is a short-range crosslinker, the RNA-protein interactions revealed by RAP-MS and iDRiP-MS tend to be direct. In contrast, the formaldehyde crosslinking fixes much larger macromolecular networks and generally leads to identify both direct and indirect factors.

To identify all RNA-binding proteins (RBPs), RBR-ID (proteomic identification of RNA-binding regions) introduces 4SU (4-thiouridine) into RNAs, crosslinks RNA-proteins by UV light, and compares the mass spectrometry of 4SU treated and non-treated samples [21]. An RNA-crosslinked peptide has a different mass, so the intensity of the signal is generally lower in the crosslinked sample compared to the non-crosslinked sample (Table 1 and Fig. 4). Using RBR-ID, about 800 previously unknown and known RNA-binding proteins, such as chromatin factors (CTCF, ATRX, HDAC1, DNMT3, EZH2, TET1, TET2 and HP1), were identified as RBPs in mouse embryonic stem cells [21].

Methods for RNA-RNA interactions

Unlike DNA, RNA is a single strand of nucleic acids and is capable of folding into 3D structures that range from simple helical elements to complex tertiary structures and quaternary ribonucleoprotein assemblies [85]. The

changes in RNA structures directly affect their functions in response to cellular conditions. Recently, high-throughput techniques that combine nuclease digestion [22, 23, 86] or chemical probing [24–26, 28] with next-generation sequencing have revealed the single-stranded or double-stranded regions of RNA molecules. FragSeq (fragmentation sequencing) [22] utilizes nuclease P1, which specifically cleaves single-stranded nucleic acids (Table 1 and Fig. 5). PARS (parallel analysis of RNA structure) [23] maps RNA structure using RNase V1 or S1 nuclease to digest RNA to determine the double or single-stranded regions, respectively. In SHAPE-Seq (selective 2′-hydroxyl acylation analyzed by primer extension sequencing) [24–26], RNAs are treated with chemical probes (such as 1 M7) that covalently modify the RNA in loops and bulges, thus blocking reverse transcription at the modified sites. However, these methods may not represent the RNA structures in vivo due to applying in vitro transcription and in vitro folding of RNAs.

To profile genome-wide RNA structures in vivo, Ding et al. developed a method, called DMS-seq [28] (Table 1 and Fig. 5). In the DMS-seq, cells are treated with dimethyl sulfate (DMS) that methylates unprotected adenines (As) and cytosines (Cs) of RNA, resulting in the reverse transcriptase stalling at one nucleotide before DMS-modified As and Cs during cDNA synthesis. A year later, Chang's group developed a method, termed icSHAPE [27] (in vivo click selective 2′-hydroxyl acylation and profiling experiment and NAI-N_3) to probe RNA secondary structures in living cells for all four bases. They used a cell-permeable SHAPE reagent, NAI-N_3, which adds an azide group to NAI (2-methylnicotinic acid imidazolide) to label flexible regions of RNA. After RNA isolation, a biotin moiety is selectively added to NAI-N_3-modified RNA by copper-free click chemistry, allowing the biotin-streptavidin purification of modified RNA after RNA fragmentation. By comparison of in vivo and in vitro icSHAPE, they observed that all RNAs are less folded in vivo, suggesting that RNA structures largely depend on intracellular environments. Moreover, they found that regulatory RNAs, such as lncRNAs and primary microRNA (miRNA) precursors, preserve their structures better than mRNAs in vivo. To selectively enrich some specific RNA molecules, COMRADES (cross-linking of matched RNAs and deep sequencing) combines RNA capture and CLICK chemistry to probe RNA structures and RNA-RNA interactions [29]. In the COMRADES method, cells are crosslinked by azide-modified psoralen, following RNA capture using biotinylated probes, and a copper-free click-chemistry reaction is carried out to link a biotin moiety to the cross-linked regions, allowing the second selection of the cross-linked regions for sequencing (Table 1 and Fig. 5).

In addition, an RNA molecule can base-pair with other RNA molecules to form RNA duplexes (such as miRNA and its target RNA) bound by RBPs. To identify miRNA targets, Tollervey et al. developed a method, named CLASH (crosslinking, ligation, and sequencing of hybrids) [30, 31] to identify human AGO1 interacting RNA duplexes (Table 1 and Fig. 5). In the CLASH method, RNA-protein complexes are UV-crosslinked and purified by antibodies against RNA binding proteins, following the ligation of RNA-RNA hybrids, and the chimeric RNAs of RNA-RNA duplexes are deep sequenced. Similarly, hiCLIP [32] also can map RNA-RNA duplexes bound by RBPs. In the hiCLIP method, a linker-adapter is introduced to the ligation step of RNA-RNA duplexes to improve the ligation efficiency (Table 1 and Fig. 5). Moreover, the ligation of two chimeric RNAs has been applied to other methods including MARIO (Mapping RNA interactome in vivo) [33], PARIS [34] and SPLASH [35], which map all RNA-RNA interactions (RNA structures and interactome) in living cells.

R-loops in gene regulation and genomic instability

R-loops are three-stranded nucleic acid structures in which a strand of RNA hybridizes with a strand of DNA, while the other strand of DNA loops out. The R-loop structure was first described in 1976 in the study [87] in which Thomas et al. showed that RNA could hybridize to double-stranded DNA in the presence of 70% formamide in vitro. A year later, Roberts and Sharp used R-loop hybridization technique to map an adenovirus 2 (Ad2) mRNA to its genome and found that some DNA sequences of the Ad2 coding gene were not hybridized with the matured RNA, suggesting that Ad2 genome consists of non-coding sequences, later known as introns [88, 89]. Recently, genome-wide studies have shown that R-loops are found in vivo, especially enriched in promoter regions [36]. Ginno et al. demonstrated that R-loop formation is involved in gene regulation via its potential to protect DNA from methylation [36]. Moreover, recent reports showed that antisense long noncoding TARID (TCF21 antisense RNA inducing promoter demethylation) forms an R-loop at the *TCF21* promoter, thus facilitating GADD45A binding, local DNA demethylation and *TCF21* expression [90, 91]. Another example is GATA3-AS1 lncRNA, which forms an R-loop structure with the central intron of *GATA3-AS1* and tethers the MLL H3K4 methyltransferase to *GATA3* gene locus, thereby regulating *GATA3* expression [92]. Furthermore, it has been proposed recently that R-loops act as intrinsic Pol II promoters to induce lncRNA transcription [93]. The depletion of R-loops by overexpressing RNase H1 causes the selective reduction of antisense lncRNA transcription [93].

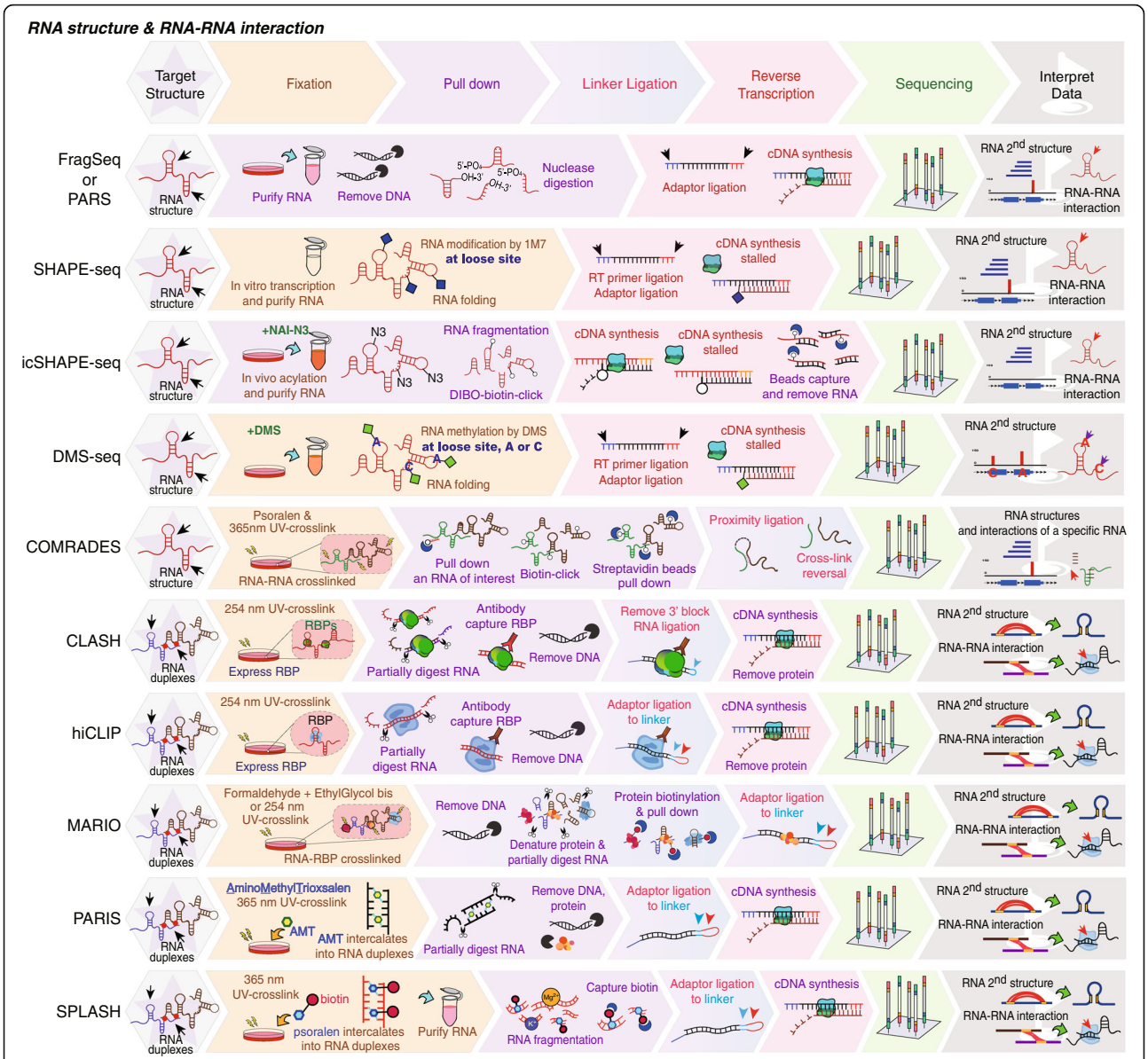

Fig. 5 Profiling RNA structure and RNA-RNA interactions. FragSeq, PARS and SHAPE-seq reveal in vitro RNA structures by nuclease or chemical probing. DMS-seq, icSHAPE-seq, and COMRADES map in vivo RNA structures by modifying RNA with chemicals. CLASH and hiCLIP can identify RNA-RNA duplexes bound by a specific protein. MARIO, PARIS, and SPLASH map all RNA-RNA interactions by introducing a linker to ligate RNA and RNA

It is generally believed that R-loops form *in cis* during transcription when a nascent RNA hybridizes to the DNA template behind the moving RNA polymerase [94]. However, research in yeast suggests that RNA:DNA hybrids can form *in trans*, which means that an RNA transcript at one locus hybridizes with homologous DNA at another locus [95], and the hybrids are likely involved in homologous recombination. Moreover, excellent studies have shown that genome instability arises from lesions generated from the formation of R-loops [95, 96]. Because transcription and replication share a common DNA template, when replication forks encounter transcription machinery, it results in transcription-replication collisions that lead to DNA damage. Therefore, the persistent RNA:DNA hybrids could be threats for genomic rearrangements [96].

Methods for mapping R-loops

To map R-loops in a genome-wide scale, Ginno et al. developed a method, named DRIP-seq (DNA-RNA immunoprecipitation coupled to high-throughput sequencing), which utilizes an antibody (S9.6) [36] to specifically

purify RNA:DNA hybrids (Table 1 and Fig. 6), and the captured DNA fragments are further sequenced. Conventional DRIP-seq generally produces a robust signal but has a limited (approximately kilobase) resolution, a higher background, and a lack of strand specificity. Another method, named DRIPc-seq (DNA:RNA immunoprecipitation followed by cDNA conversion coupled to high-throughput sequencing), which builds on DRIP, except that a strand-specific RNA sequencing is performed to profile R-loops [40]. DRIPc-seq increases the resolution of R-loop profiling and shows the strand-specificity. However, the sensitivity of DRIP for revealing authentic R-loops in vivo has been doubted due to the fact that the immunoprecipitation for R-loops is usually performed using isolated genomic DNA without any crosslinking. Thus R-loops could be disrupted or formed in vitro after cell lysis. Dumelie et al. developed a method, named bisDRIP-seq (bisulfite-based DRIP-seq) [38], which selectively labels the single-stranded DNA that loops out from an R-loop. They used bisulfite to convert cytosine residues into uracil residues within the genomic DNA regions that contain single-stranded DNA simultaneously when cells were lysed. Their results showed that bisDRIP-seq could map R-loops at near nucleotide resolution and detect the boundaries of R-loops. In contrast to DRIP, R-ChIP (RNase H chromatin immunoprecipitation) utilizes an RNase H, which specifically recognizes DNA:RNA hybrids in vivo to map R-loops [39]. By expressing a catalytically dead enzyme, R-ChIP captures R-loops using a standard ChIP-seq protocol, which involves both fixation to stabilize R-loop/RNase H complex and sonication to increase the resolution (Table 1 and Fig. 6).

lncRNAs mediate phase separation of membrane-less organelles

Cellular organelles such as mitochondria and Golgi apparatuses composed of lipid bilayer membrane structures, which help the formation of compartments and separate biological processes within a cell. In contrast, membraneless structures are formed through a process known as liquid-liquid phase separation (LLPS) and are made of RNA-protein droplets [97]. One of the most prominent membraneless structures in the nucleus is the nucleolus [98], which produces the ribosomal RNA and consists of a variety of proteins and RNA. Studies have shown that rRNA transcription is important for nucleolar assembly [50, 51, 99, 100]. Other membraneless structures, such as paraspeckles, Cajal bodies (CBs), histone locus bodies (HLBs) and promyelocytic leukemia (PML) bodies are also found in the nucleoplasm as phase separated-like droplets (Fig. 7), while stress bodies and process bodies are RNA granules in the cytoplasm. DNA is typically absent from the interior of these liquid-like droplets, whereas lncRNAs serve as scaffolds for their formation and maintenance [101–103].

How do RNAs promote phase separation synergistically with protein-protein interactions? The in vitro studies reported that RNA repeats, such as trinucleotide repeat and G-quadruplex, can undergo a phase transition to form either a condensed liquid or a gel-like state [104] through the multivalent base-pairing between RNA molecules. Droplet-like assemblies of RNA are associated with certain neurodegenerative diseases, including Huntington disease, muscular dystrophy, and

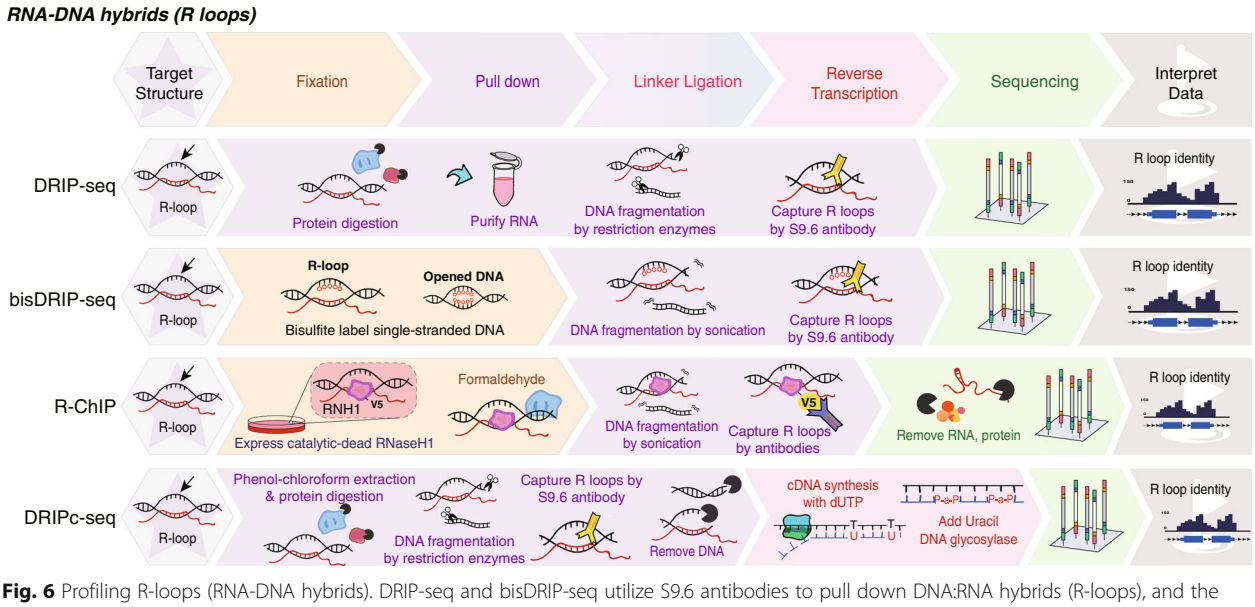

Fig. 6 Profiling R-loops (RNA-DNA hybrids). DRIP-seq and bisDRIP-seq utilize S9.6 antibodies to pull down DNA:RNA hybrids (R-loops), and the captured DNA fragments are sequenced. DRIPc-seq builds on DRIP-seq, and the captured RNA fragments are reversed transcribed and sequenced. R-ChIP captures R-loops using a standard ChIP-seq protocol to pull down R-loop/catalytic-dead RNase H complex

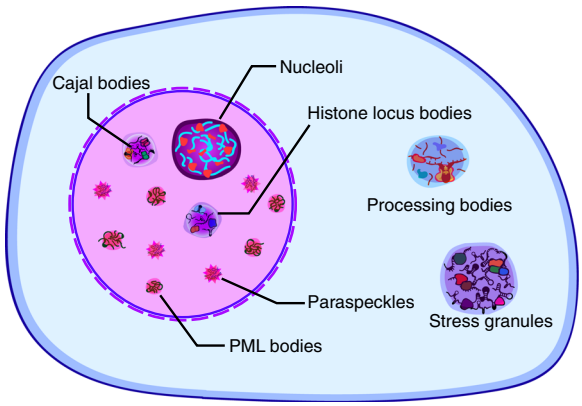

Fig. 7 RNA-protein droplets in cells. RNA granules are made out of protein and RNA complexes, and their formations are driven by liquid-liquid phase separation. Nucleoli, paraspeckles, Cajal bodies (CBs), histone locus bodies (HLBs) and promyelocytic leukemia (PML) bodies are RNA assemblies in the nucleus. P-bodies and stress granules are found in the cytoplasm. RNAs act as a regulatory element to control their sizes and properties. The imbalance of RNA/protein ratio in such RNA assemblies could lead to human diseases

amyotrophic lateral sclerosis [105]. It has been proposed that such gel-like RNA foci may contribute to neuronal dysfunction in vivo. A recent study showed that RNA plays an important role in the phase behavior of prion-like RBPs, such as TDP43 and FUS [106], which are largely soluble in the nucleus but form solid pathological aggregates when mislocalized to the cytoplasm. The in vitro studies indicated that the ratio of RNA/protein is important for droplet formation and phase separation [106]. Remarkably, reduction of nuclear RNAs or disruption of RNA binding leads to excessive phase separation and the formation of cytosolic assemblies in cells [106], suggesting that the higher RNA concentrations in the nucleus act as a buffer to prevent aggregation of RBPs in the cytoplasm. Then the question is how the RNA/protein ratio is regulated in such droplets. Hondele et al. recently reported that RNA flux into and out of phase-separated organelles is controlled by RNA-dependent DEAD-box ATPases (DDXs), which contain low-complexity domains (LCDs) that are crucial for the formation of multivalent meshworks within membraneless droplets [107]. In addition, Ries et al. showed that methylation on mRNAs triggers phase separation of endogenous compartments, such as P-bodies, stress granules or neuronal RNA granules [108]. These studies indicate that the abundance of nuclear RNAs and the modifications of RNA contribute to the dynamics of such membraneless organelles. However, there are intriguing questions that remain elusive. How do RNA structures impact on phase separation? How do RNA-mediated droplets involve in chromatin organization and gene regulation? What signals trigger the reorganization of such structures?

Conclusions

The Human Genome and ENCODE Projects have shown that above 90% of the genome is transcribed [1, 72] into non-coding RNAs. However, the functions of most non-coding RNAs remain largely unknown. In the last decade, emerging new techniques combined with high throughput sequencing have profiled the interactions between DNA, RNA, and proteins, thus facilitating the studies of lncRNAs functions in cells. LncRNAs can act as a guide, or decoy, or a scaffold for protein complexes to mediate the epigenetic regulation [109]. Chromatin-associated lncRNAs involves nuclear architecture and chromatin conformation. Given that lncRNAs are long and mobile, they could serve as bridges to mediate chromatin looping and to drive the inter- or intra-chromosomal interactions [44, 45]. Moreover, RNA is able to hybridize with DNA to form R-loops, which have been found to contribute to gene regulation and genomic instability. RNA also mediates the liquid-liquid phase separation through its ability being as a multivalent scaffold for the binding of RBPs, thus regulating the sizes and the dynamics of membraneless organelles that carry biological processes. It is still a beginning to uncover the surface of the lncRNA world. There are still a lot we don't know and plenty of work that needs to be done. Because RNA molecules consist of specific sequences, it is realistic and easy to design drug targets by blocking their actions using antisense oligos, RNA interferences, or aptamers. We hope that by understanding the mechanisms of lncRNAs action, RNA-centric therapies could be potential options to treat human diseases.

Abbreviations

lncRNAs: Long non-coding RNAs; R-loops: RNA:DNA hybrids; CTCF: CCCTC-binding transcription factor; TADs: Topologically associated domains; ChIP: Chromatin Immunoprecipitation; eRNAs: Enhancer RNAs; RBPs: RNA binding proteins; 4SU: 4-thiouridine; IGF2: Insulin-like growth factor 2; ATRX: Alpha Thalassemia/Mental Retardation Syndrome X-Linked; HDAC1: Histone Deacetylase 1; DNMT3: DNA methyltransferase 3; EZH2: Enhancer Of Zeste Homolog 2; TET1: Ten-eleven translocation *methylcytosine dioxygenase* 1; TET2: Ten-eleven translocation *methylcytosine dioxygenase* 2; HP1: Heterochromatin Protein 1; PRC1: Polycomb Repressive Complex 1; PRC2: Polycomb Repressive Complex 2; *AGO1*: Argonaute RISC Component 1; 1 M7: 1-methyl-7-nitroisatoic anhydride; GADD45A: Growth arrest and DNA damage inducible alpha; GATA3: GATA binding protein 3

Acknowledgements

We would like to thank Tse-Chun Kuo and Hsing-Jien Kung for the invitation of this review article. We apologize to any of our colleagues whose work could not be cited here for space constraints.

Authors' contributions

C-Y. G. and **Y-H. H. prepared the figures and Tables.** H-P. C. wrote the manuscript. The authors read and approved the final manuscript.

Funding

This work was supported by Ministry of Science and Technology (MOST) grants 106–2311-B-002-039-MY2 (H-P.C.), 108–2628-B-002-004 (H-P. C.), and National Taiwan University grants NTU-CDP-108 L7732, NTU-AS-108 L104305, NTU-CC-107 L892003.

Availability of data and materials
Not applicable.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 7 November 2019 Accepted: 6 March 2020
Published online: 17 March 2020

References
1. Kung JT, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. Genetics. 2013;193(3):651–69. https://doi.org/10.1534/genetics.112.146704.
2. Thapar R. Regulation of DNA Double-Strand Break Repair by Non-Coding RNAs. Molecules. 2018;23(11). https://doi.org/10.3390/molecules23112789.
3. Sparber P, Filatova A, Khantemirova M, et al. The role of long non-coding RNAs in the pathogenesis of hereditary diseases. BMC Med Genet. 2019; 12(Suppl 2):42. https://doi.org/10.1186/s12920-019-0487-6.
4. Turton N, Swan R, Mahenthiralingam T, et al. The Functions of Long Non-Coding RNA during Embryonic Cardiovascular Development and Its Potential for Diagnosis and Treatment of Congenital Heart Disease. J Cardiovasc Dev Dis. 2019;6(2). https://doi.org/10.3390/jcdd6020021.
5. Schmitt AM, Chang HY. Long Noncoding RNAs: At the Intersection of Cancer and Chromatin Biology. Cold Spring Harb Perspect Med. 2017;7(7). https://doi.org/10.1101/cshperspect.a026492.
6. Schmitz SU, Grote P, Herrmann BG. Mechanisms of long noncoding RNA function in development and disease. Cell Mol Life Sci. 2016;73(13):2491–509. https://doi.org/10.1007/s00018-016-2174-5.
7. Huang RC, Bonner J. Histone-bound RNA, a component of native nucleohistone. Proc Natl Acad Sci U S A. 1965;54(3):960–7. https://doi.org/10.1073/pnas.54.3.960.
8. Holmes DS, Mayfield JE, Sander G, et al. Chromosomal RNA: its properties. Science. 1972;177(4043):72–4. https://doi.org/10.1126/science.177.4043.72.
9. Pederson T. Thinking about a nuclear matrix. J Mol Biol. 1998;277(2):147–59. https://doi.org/10.1006/jmbi.1997.1618.
10. Nickerson JA, Krochmalnic G, Wan KM, et al. Chromatin architecture and nuclear RNA. Proc Natl Acad Sci U S A. 1989;86(1):177–81. https://doi.org/10.1073/pnas.86.1.177.
11. Caudron-Herger M, Rippe K. Nuclear architecture by RNA. Curr Opin Genet Dev. 2012;22(2):179–87. https://doi.org/10.1016/j.gde.2011.12.005.
12. Kung JT, Kesner B, An JY, et al. Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. Mol Cell. 2015; 57(2):361–75. https://doi.org/10.1016/j.molcel.2014.12.006.
13. Yao H, Brick K, Evrard Y, et al. Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. Genes Dev. 2010;24(22):2543–55. https://doi.org/10.1101/gad.1967810.
14. Sun S, Del Rosario BC, Szanto A, et al. Jpx RNA activates Xist by evicting CTCF. Cell. 2013;153(7):1537–51. https://doi.org/10.1016/j.cell.2013.05.028.
15. Mele M, Rinn JL. "Cat's Cradling" the 3D Genome by the Act of LncRNA Transcription. Mol Cell. 2016;62(5):657–64. https://doi.org/10.1016/j.molcel.2016.05.011.
16. Oakes M, Nogi Y, Clark MW, et al. Structural alterations of the nucleolus in mutants of Saccharomyces cerevisiae defective in RNA polymerase I. Mol Cell Biol. 1993;13(4):2441–55. https://doi.org/10.1128/mcb.13.4.2441.
17. Karpen GH, Schaefer JE, Laird CD. A Drosophila rRNA gene located in euchromatin is active in transcription and nucleolus formation. Genes Dev. 1988;2(12B):1745–63. https://doi.org/10.1101/gad.2.12b.1745.
18. Nozawa RS, Boteva L, Soares DC, et al. SAF-A Regulates Interphase Chromosome Structure through Oligomerization with Chromatin-Associated RNAs. Cell. 2017;169(7):1214–27.e18. https://doi.org/10.1016/j.cell.2017.05.029.
19. Brown CJ, Hendrich BD, Rupert JL, et al. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell. 1992;71(3):527–42. https://doi.org/10.1016/0092-8674(92)90520-m.
20. Wang J, Mager J, Chen Y, et al. Imprinted X inactivation maintained by a mouse Polycomb group gene. Nat Genet. 2001;28(4):371–5. https://doi.org/10.1038/ng574.
21. Plath K, Talbot D, Hamer KM, et al. Developmentally regulated alterations in Polycomb repressive complex 1 proteins on the inactive X chromosome. J Cell Biol. 2004;167(6):1025–35. https://doi.org/10.1083/jcb.200409026.
22. Zhao J, Sun BK, Erwin JA, et al. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science. 2008;322(5902):750–6. https://doi.org/10.1126/science.1163045.
23. Minajigi A, Froberg J, Wei C, et al. Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. Science. 2015;349(6245). https://doi.org/10.1126/science.aab2276.
24. Jegu T, Blum R, Cochrane JC, et al. Xist RNA antagonizes the SWI/SNF chromatin remodeler BRG1 on the inactive X chromosome. Nat Struct Mol Biol. 2019;26(2):96–109. https://doi.org/10.1038/s41594-018-0176-8.
25. Bickmore WA, van Steensel B. Genome architecture: domain organization of interphase chromosomes. Cell. 2013;152(6):1270–84. https://doi.org/10.1016/j.cell.2013.02.001.
26. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326(5950):289–93. https://doi.org/10.1126/science.1181369.
27. Wang CY, Jegu T, Chu HP, et al. SMCHD1 Merges Chromosome Compartments and Assists Formation of Super-Structures on the Inactive X. Cell. 2018;174(2):406–21.e25. https://doi.org/10.1016/j.cell.2018.05.007.
28. Kim TK, Hemberg M, Gray JM, et al. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010;465(7295):182–7. https://doi.org/10.1038/nature09033.
29. De Santa F, Barozzi I, Mietton F, et al. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol. 2010;8(5):e1000384. https://doi.org/10.1371/journal.pbio.1000384.
30. Ding M, Liu Y, Liao X, et al. Enhancer RNAs (eRNAs): New Insights into Gene Transcription and Disease Treatment. J Cancer. 2018;9(13):2334–40. https://doi.org/10.7150/jca.25829.
31. Hah N, Murakami S, Nagari A, et al. Enhancer transcripts mark active estrogen receptor binding sites. Genome Res. 2013;23(8):1210–23. https://doi.org/10.1101/gr.152306.112.
32. Hsieh CL, Fei T, Chen Y, et al. Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. Proc Natl Acad Sci U S A. 2014;111(20):7319–24. https://doi.org/10.1073/pnas.1324151111.
33. Wang D, Garcia-Bassets I, Benner C, et al. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature. 2011; 474(7351):390–4. https://doi.org/10.1038/nature10006.
34. Yang Y, Su Z, Song X, et al. Enhancer RNA-driven looping enhances the transcription of the long noncoding RNA DHRS4-AS1, a controller of the DHRS4 gene cluster. Sci Rep. 2016;6:20961. https://doi.org/10.1038/srep20961.
35. Pnueli L, Rudnizky S, Yosefzon Y, et al. RNA transcribed from a distal enhancer is required for activating the chromatin at the promoter of the gonadotropin alpha-subunit gene. Proc Natl Acad Sci U S A. 2015;112(14): 4369–74. https://doi.org/10.1073/pnas.1414841112.
36. Li W, Notani D, Ma Q, et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature. 2013;498(7455):516–20. https://doi.org/10.1038/nature12210.
37. Stevens NM. A study of the germ cells of certain Diptera with reference to the heterochromosomes and the phenomena of synapsis. J Exp Zool. 1908; 5(3):359–74. https://doi.org/10.1002/jez.1400050304.
38. Lee JT. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. Genes Dev. 2009;23(16):1831–42. https://doi.org/10.1101/gad.1811209.
39. Consortium EP, Birney E, Stamatoyannopoulos JA, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature. 2007;447(7146):799–816. https://doi.org/10.1038/nature05874.
40. Xu N, Tsai CL, Lee JT. Transient homologous chromosome pairing marks the onset of X inactivation. Science. 2006;311(5764):1149–52. https://doi.org/10.1126/science.1122984.
41. Chu HP, Cifuentes-Rojas C, Kesner B, et al. TERRA RNA Antagonizes ATRX and Protects Telomeres. Cell. 2017;170(1):86–101.e16. https://doi.org/10.1016/j.cell.2017.06.017.

42. Hogan MS, Parfitt DE, Zepeda-Mendoza CJ, et al. Transient pairing of homologous Oct4 alleles accompanies the onset of embryonic stem cell differentiation. Cell Stem Cell. 2015;16(3):275–88. https://doi.org/10.1016/j.stem.2015.02.001.
43. Zorca CE, Kim LK, Kim YJ, et al. Myosin VI regulates gene pairing and transcriptional pause release in T cells. Proc Natl Acad Sci U S A. 2015;112(13):E1587–93. https://doi.org/10.1073/pnas.1502461112.
44. Brandt VL, Hewitt SL, Skok JA. It takes two: communication between homologous alleles preserves genomic stability during V(D)J recombination. Nucleus. 2010;1(1):23–9. https://doi.org/10.4161/nucl.1.1.10595.
45. Spilianakis CG, Lalioti MD, Town T, et al. Interchromosomal associations between alternatively expressed loci. Nature. 2005;435(7042):637–45. https://doi.org/10.1038/nature03574.
46. Donohoe ME, Silva SS, Pinter SF, et al. The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting. Nature. 2009;460(7251):128–32. https://doi.org/10.1038/nature08098.
47. Scialdone A, Nicodemi M. Mechanics and dynamics of X-chromosome pairing at X inactivation. PLoS Comput Biol. 2008;4(12):e1000244. https://doi.org/10.1371/journal.pcbi.1000244.
48. Chu C, Qu K, Zhong FL, et al. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell. 2011;44(4):667–78. https://doi.org/10.1016/j.molcel.2011.08.027.
49. Simon MD, Wang CI, Kharchenko PV, et al. The genomic binding sites of a noncoding RNA. Proc Natl Acad Sci U S A. 2011;108(51):20497–502. https://doi.org/10.1073/pnas.1113536108.
50. Hu T, Pi W, Zhu X, et al. Long non-coding RNAs transcribed by ERV-9 LTR retrotransposon act in cis to modulate long-range LTR enhancer function. Nucleic Acids Res. 2017;45(8):4479–92. https://doi.org/10.1093/nar/gkx055.
51. Wang F, Chainani P, White T, et al. Deep learning identifies genome-wide DNA binding sites of long noncoding RNAs. RNA Biol. 2018;15(12):1468–76. https://doi.org/10.1080/15476286.2018.1551704.
52. Simon MD, Pinter SF, Fang R, et al. High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. Nature. 2013;504(7480):465–9. https://doi.org/10.1038/nature12719.
53. Sridhar B, Rivas-Astroza M, Nguyen TC, et al. Systematic Mapping of RNA-Chromatin Interactions In Vivo. Curr Biol. 2017;27(4):602–9. https://doi.org/10.1016/j.cub.2017.01.011.
54. Bell JC, Jukam D, Teran NA, et al. Chromatin-associated RNA sequencing (ChAR-seq) maps genome-wide RNA-to-DNA contacts. Elife. 2018;7. https://doi.org/10.7554/eLife.27024.
55. Li X, Zhou B, Chen L, et al. GRID-seq reveals the global RNA-chromatin interactome. Nat Biotechnol. 2017;35(10):940–50. https://doi.org/10.1038/nbt.3968.
56. Mumbach MR, Granja JM, Flynn RA, et al. HiChIRP reveals RNA-associated chromosome conformation. Nat Methods. 2019;16(6):489–92. https://doi.org/10.1038/s41592-019-0407-x.
57. Lerner MR, Steitz JA. Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. Proc Natl Acad Sci U S A. 1979;76(11):5495–9. https://doi.org/10.1073/pnas.76.11.5495.
58. Miyashiro KY, Beckel-Mitchener A, Purk TP, et al. RNA cargoes associating with FMRP reveal deficits in cellular functioning in Fmr1 null mice. Neuron. 2003;37(3):417–31. https://doi.org/10.1016/s0896-6273(03)00034-5.
59. Musunuru K, Darnell RB. Paraneoplastic neurologic disease antigens: RNA-binding proteins and signaling proteins in neuronal degeneration. Annu Rev Neurosci. 2001;24:239–62. https://doi.org/10.1146/annurev.neuro.24.1.239.
60. Ule J, Jensen KB, Ruggiu M, et al. CLIP identifies Nova-regulated RNA networks in the brain. Science. 2003;302(5648):1212–5. https://doi.org/10.1126/science.1090095.
61. Konig J, Zarnack K, Rot G, et al. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol. 2010;17(7):909–15. https://doi.org/10.1038/nsmb.1838.
62. Hafner M, Landthaler M, Burger L, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell. 2010;141(1):129–41. https://doi.org/10.1016/j.cell.2010.03.009.
63. Chu C, Zhang QC, da Rocha ST, et al. Systematic discovery of Xist RNA binding proteins. Cell. 2015;161(2):404–16. https://doi.org/10.1016/j.cell.2015.03.025.
64. McHugh CA, Chen CK, Chow A, et al. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature. 2015;521(7551):232–6. https://doi.org/10.1038/nature14443.
65. He C, Sidoli S, Warneford-Thomson R, et al. High-Resolution Mapping of RNA-Binding Regions in the Nuclear Proteome of Embryonic Stem Cells. Mol Cell. 2016;64(2):416–30. https://doi.org/10.1016/j.molcel.2016.09.034.
66. Ganser LR, Kelly ML, Herschlag D, et al. The roles of structural dynamics in the cellular functions of RNAs. Nat Rev Mol Cell Biol. 2019;20(8):474–89. https://doi.org/10.1038/s41580-019-0136-0.
67. Underwood JG, Uzilov AV, Katzman S, et al. FragSeq: transcriptome-wide RNA structure probing using high-throughput sequencing. Nat Methods. 2010;7(12):995–1001. https://doi.org/10.1038/nmeth.1529.
68. Kertesz M, Wan Y, Mazor E, et al. Genome-wide measurement of RNA secondary structure in yeast. Nature. 2010;467(7311):103–7. https://doi.org/10.1038/nature09322.
69. Wan Y, Qu K, Zhang QC, et al. Landscape and variation of RNA secondary structure across the human transcriptome. Nature. 2014;505(7485):706–9. https://doi.org/10.1038/nature12946.
70. Lucks JB, Mortimer SA, Trapnell C, et al. Multiplexed RNA structure characterization with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). Proc Natl Acad Sci U S A. 2011;108(27):11063–8. https://doi.org/10.1073/pnas.1106501108.
71. Mortimer SA, Trapnell C, Aviran S, et al. SHAPE-Seq: High-Throughput RNA Structure Analysis. Curr Protoc Chem Biol. 2012;4(4):275–97. https://doi.org/10.1002/9780470559277.ch120019.
72. Loughrey D, Watters KE, Settle AH, et al. SHAPE-Seq 2.0: systematic optimization and extension of high-throughput chemical probing of RNA secondary structure with next generation sequencing. Nucleic Acids Res. 2014;42(21). https://doi.org/10.1093/nar/gku909.
73. Ding Y, Tang Y, Kwok CK, et al. In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. Nature. 2014;505(7485):696–700. https://doi.org/10.1038/nature12756.
74. Spitale RC, Flynn RA, Zhang QC, et al. Structural imprints in vivo decode RNA regulatory mechanisms. Nature. 2015;519(7544):486–90. https://doi.org/10.1038/nature14263.
75. Ziv O, Gabryelska MM, Lun ATL, et al. COMRADES determines in vivo RNA structures and interactions. Nat Methods. 2018;15(10):785–8. https://doi.org/10.1038/s41592-018-0121-0.
76. Helwak A, Kudla G, Dudnakova T, et al. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell. 2013;153(3):654–65. https://doi.org/10.1016/j.cell.2013.03.043.
77. Kudla G, Granneman S, Hahn D, et al. Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. Proc Natl Acad Sci U S A. 2011;108(24):10010–5. https://doi.org/10.1073/pnas.1017386108.
78. Sugimoto Y, Vigilante A, Darbo E, et al. hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. Nature. 2015;519(7544):491–4. https://doi.org/10.1038/nature14280.
79. Nguyen TC, Cao X, Yu P, et al. Mapping RNA-RNA interactome and RNA structure in vivo by MARIO. Nat Commun. 2016;7:12023. https://doi.org/10.1038/ncomms12023.
80. Lu Z, Zhang QC, Lee B, et al. RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. Cell. 2016;165(5):1267–79. https://doi.org/10.1016/j.cell.2016.04.028.
81. Aw JG, Shen Y, Wilm A, et al. In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation. Mol Cell. 2016;62(4):603–17. https://doi.org/10.1016/j.molcel.2016.04.028.
82. Thomas M, White RL, Davis RW. Hybridization of RNA to double-stranded DNA: formation of R-loops. Proc Natl Acad Sci U S A. 1976;73(7):2294–8. https://doi.org/10.1073/pnas.73.7.2294.
83. Berget SM, Moore C, Sharp PA. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. Proc Natl Acad Sci U S A. 1977;74(8):3171–5. https://doi.org/10.1073/pnas.74.8.3171.
84. Chow LT, Gelinas RE, Broker TR, et al. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. Cell. 1977;12(1):1–8. https://doi.org/10.1016/0092-8674(77)90180-5.
85. Ginno PA, Lott PL, Christensen HC, et al. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol Cell. 2012;45(6):814–25. https://doi.org/10.1016/j.molcel.2012.01.017.
86. Arab K, Park YJ, Lindroth AM, et al. Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A. Mol Cell. 2014;55(4):604–14. https://doi.org/10.1016/j.molcel.2014.06.031.
87. Arab K, Karaulanov E, Musheev M, et al. GADD45A binds R-loops and recruits TET1 to CpG island promoters. Nat Genet. 2019;51(2):217–23. https://doi.org/10.1038/s41588-018-0306-6.
88. Gibbons HR, Shaginurova G, Kim LC, et al. Divergent lncRNA GATA3-AS1 Regulates GATA3 Transcription in T-Helper 2 Cells. Front Immunol. 2018;9:2512. https://doi.org/10.3389/fimmu.2018.02512.

89. Tan-Wong SM, Dhir S, Proudfoot NJ. R-Loops Promote Antisense Transcription across the Mammalian Genome. Mol Cell. 2019;76(4):600–16. e6. https://doi.org/10.1016/j.molcel.2019.10.002.
90. Groh M, Gromak N. Out of balance: R-loops in human disease. PLoS Genet. 2014;10(9):e1004630. https://doi.org/10.1371/journal.pgen.1004630.
91. Wahba L, Gore SK, Koshland D. The homologous recombination machinery modulates the formation of RNA-DNA hybrids and associated chromosome instability. Elife. 2013;2:e00505. https://doi.org/10.7554/eLife.00505.
92. Crossley MP, Bocek M, Cimprich KA. R-Loops as Cellular Regulators and Genomic Threats. Mol Cell. 2019;73(3):398–411. https://doi.org/10.1016/j.molcel.2019.01.024.
93. Sanz LA, Hartono SR, Lim YW, et al. Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. Mol Cell. 2016;63(1):167–78. https://doi.org/10.1016/j.molcel.2016.05.032.
94. Dumelie JG, Jaffrey SR. Defining the location of promoter-associated R-loops at near-nucleotide resolution using bisDRIP-seq. Elife. 2017;6. https://doi.org/10.7554/eLife.28306.
95. Chen L, Chen JY, Zhang X, et al. R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. Mol Cell. 2017;68(4):745–57.e5. https://doi.org/10.1016/j.molcel.2017.10.008.
96. Palikyras S, Papantonis A. Modes of phase separation affecting chromatin regulation. Open Biol. 2019;9(10):190167. https://doi.org/10.1098/rsob.190167.
97. Brangwynne CP, Mitchison TJ, Hyman AA. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. Proc Natl Acad Sci U S A. 2011;108(11):4334–9. https://doi.org/10.1073/pnas.1017150108.
98. Yao RW, Xu G, Wang Y, et al. Nascent Pre-rRNA Sorting via Phase Separation Drives the Assembly of Dense Fibrillar Components in the Human Nucleolus. Mol Cell. 2019. https://doi.org/10.1016/j.molcel.2019.08.014.
99. Falahati H, Pelham-Webb B, Blythe S, et al. Nucleation by rRNA Dictates the Precision of Nucleolus Assembly. Curr Biol. 2016;26(3):277–85. https://doi.org/10.1016/j.cub.2015.11.065.
100. Bond CS, Fox AH. Paraspeckles: nuclear bodies built on long noncoding RNA. J Cell Biol. 2009;186(5):637–44. https://doi.org/10.1083/jcb.200906113.
101. Naganuma T, Hirose T. Paraspeckle formation during the biogenesis of long non-coding RNAs. RNA Biol. 2013;10(3):456–61. https://doi.org/10.4161/rna.23547.
102. Fox AH, Nakagawa S, Hirose T, et al. Paraspeckles: Where Long Noncoding RNA Meets Phase Separation. Trends Biochem Sci. 2018;43(2):124–35. https://doi.org/10.1016/j.tibs.2017.12.001.
103. Jain A, Vale RD. RNA phase transitions in repeat expansion disorders. Nature. 2017;546(7657):243–7. https://doi.org/10.1038/nature22386.
104. Alberti S, Dormann D. Liquid-Liquid Phase Separation in Disease. Annu Rev Genet. 2019. https://doi.org/10.1146/annurev-genet-112618-043527.
105. Maharana S, Wang J, Papadopoulos DK, et al. RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science. 2018;360(6391):918–21. https://doi.org/10.1126/science.aar7366.
106. Hondele M, Sachdev R, Heinrich S, et al. DEAD-box ATPases are global regulators of phase-separated organelles. Nature. 2019;573(7772):144–8. https://doi.org/10.1038/s41586-019-1502-y.
107. Ries RJ, Zaccara S, Klein P, et al. m(6)A enhances the phase separation potential of mRNA. Nature. 2019;571(7765):424–8. https://doi.org/10.1038/s41586-019-1374-1.
108. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 2012;81:145–66. https://doi.org/10.1146/annurev-biochem-051410-092902.
109. Fox M, Roberts JJ. Drug resistance and DNA repair. Cancer Metastasis Rev. 1987;6(3):261–81.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions