



RNA: interactions drive functionalities

Xiaofeng Dai¹ · Shuo Zhang² · Kathia Zaleta-Rivera³

Received: 7 August 2019 / Accepted: 7 December 2019 / Published online: 14 December 2019
© Springer Nature B.V. 2019

Abstract

RNA is produced from the majority of human genomic sequences, although only a relatively small portion of these transcripts has known functions. Diverse RNA species interact with RNA, DNA, proteins, lipids, and metabolites to form intricate molecular networks. In this review, we attempt to delineate diverse RNA functions by interaction types between RNA and other macromolecules. Through such interactions RNAs participate in essentially every major molecular function and process, including information flow and storage, environment sensing, signal transduction, and gene regulation at transcriptional and posttranscriptional levels. Through such interactions, RNAs promote or inhibit diverse biological processes, and act as catalyzer or quencher to modulate the pace of these progresses. Alterations and personal variations of these interactions are mechanistically coupled with disease etiology and phenotypical variations for clinical use.

Keywords RNA interaction · Non-coding RNA · Functionalities

Introduction

At least 75% human genome is transcribed into RNAs [1], fulfilling complex functions in diverse cell types. Though novel classes of non-coding RNAs [2] (ncRNAs) are kept being discovered with technical advances, the coding and noncoding RNAs with functional annotations remain a relatively small fraction [3]. Identifying the shared features, if any, of the previous discoveries facilitates the discovery of new functional RNAs, which helps fulfilling this knowledge gap towards revealing the whole picture of genome functionalities.

RNAs are diverse and have refined roles. Besides canonical protein-encoding RNAs, various long and short ncRNAs have been continuously discovered with advances in the experimental and computational approaches of

contemporary biology. We propose to classify RNAs with (partially) characterized functionalities into three categories (Table 1). The first class encompasses protein-coding associated RNAs including RNAs designated to protein synthesis, i.e., messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA), and RNAs responsible for their maturation. While mRNAs, rRNAs and tRNAs are responsible for protein translation; small nuclear RNAs (snRNA) primarily function in pre-mRNA processing in the nucleus, ribonuclease P (RNase P) and ribonuclease MRP (RNase MRP) are responsible for tRNA and rRNA maturation, respectively, and small nucleolar RNAs (snoRNA) possess impressively diverse functions including rRNA and snRNA modification. The second class of RNAs comprises of regulatory RNAs. For example, micro RNAs (miRNAs), short hairpin RNAs (shRNAs), small interfering RNAs (siRNA) are short ncRNAs interfering with gene expression via degrading mRNA molecules post-transcriptionally; piwi-interacting RNAs (piRNAs) silence retrotransposons and other genetic elements in germline cells; antisense RNAs (aRNAs) are endogenous RNAs with partial or full sequence complementarity to other transcripts, which use diverse transcriptional and post-transcriptional gene regulatory mechanisms to carry out a wide variety of biological functions; competing endogenous RNAs (ceRNAs) regulate mRNA transcripts by competing for shared miRNAs; long noncoding RNAs (lncRNAs) are over 200

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11033-019-05230-7>) contains supplementary material, which is available to authorized users.

✉ Xiaofeng Dai
xiaofeng.dai@jiangnan.edu.cn

¹ Wuxi School of Medicine, Jiangnan University, Wuxi, China

² School of Biotechnology, Jiangnan University, Wuxi, China

³ Department of Bioengineering, University of California San Diego, San Diego, USA

Table 1 Classification and functionalities of currently known RNAs

Category	RNA name	Abbreviation	Sub-category	Functionalities
Protein-coding associated RNA	Messenger RNA	mRNA	Protein synthesis	Convey genetic information from DNA to tRNA
	Transfer RNA	tRNA	Protein synthesis	Convey genetic information from mRNA to protein
	Ribosomal RNA	rRNA	Protein synthesis	Place where protein synthesis occurs
	Small nuclear RNA	snRNA	RNA maturation	pre-mRNA processing in the nucleus
	Ribonuclease P	RNase P	RNA maturation	tRNA maturation
	Ribonuclease MRP	RNase MRP	RNA maturation	rRNA maturation
	Small nucleolar RNA	snoRNA	RNA maturation	rRNA modification, snRNA modification
Regulatory RNA	microRNA	miRNA	Short noncoding RNA	Interfering with gene expression via degrading mRNA post-transcriptionally
	Short hairpin RNA	shRNA	Short noncoding RNA	Interfering with gene expression via degrading mRNA post-transcriptionally
	Small interfering RNA	siRNA	Short noncoding RNA	Interfering with gene expression via degrading mRNA post-transcriptionally
	Piwi-interacting RNA	piRNA	Short noncoding RNA	Silence retrotransposons and other genetic elements in germline cells
	Antisense RNA	aRNA	Short noncoding RNA	Endogenous RNAs with partial/full sequence complementarity to other transcripts that use diverse transcriptional and post-transcriptional gene regulatory mechanisms to carry out multiple biological functions
	Enhancer RNA	eRNA	Relatively short noncoding RNA	Short ncRNA molecules transcribed from the enhancer regions and actively play a role in transcriptional regulation <i>in cis</i> and <i>in trans</i>
	Competing endogenous RNA	ceRNA	Long noncoding RNA	Regulate mRNA transcripts by competing for shared miRNAs
	Long noncoding RNA	lncRNA	Long noncoding RNA	Mediate epigenetic alterations via recruiting chromatin remodeling complexes to specific genomic loci
Parasitic RNA	Retrotransposon	Retrotransposon	Retrotransposon	Genetic elements that can amplify themselves in a genome and are ubiquitous components of the DNA of many eukaryotic organisms
	RNA virus	RNA virus	Virus	Virus that has RNA as its genetic material
	Viroid	Viroid	Viroid	Circular noncoding RNAs that infect plants
	CRISPR RNA	crRNA	Phage	Constitute prokaryotic immune system by helping Cas proteins recognize and cut exogenous DNA

Table 1 (continued)

Category	RNA name	Abbreviation	Sub-category	Functionalities
Unclassified RNA	Circular RNA	circRNA		Some are regulatory RNA, some are protein-coding RNA
	Sno-derived RNA	sdRNA		MiRNA-like RNAs originated from H/ACA box snoRNAs or C/D box snoRNAs with hypothetical roles during the interplay between RNA silencing and snoRNA-mediated RNA processing systems
	miRNA-offset RNA	moRNA		Produced from human miRNA precursors but have considerably lower expression levels than the corresponding miRNAs
	Transcription initiation RNA	tiRNA		Mapped within –60 to +120 nucleotides of the transcription start site (TSS) and suggested as a general feature of transcription in possibly all eukaryotes with exact functions uncharacterized

nucleotides in length and transcribed in a tissue specific and developmentally regulated pattern that can mediate epigenetic alterations via recruiting chromatin remodeling complexes to specific genomic loci; enhancer RNAs (eRNAs) are short ncRNA molecules transcribed from the enhancer regions and actively play a role in transcriptional regulation *in cis* and *in trans* [4]. The last group is comprised of parasitic RNAs including, e.g., retrotransposons capable of self-propagation, RNA viruses, and CRISPR RNAs (crRNAs) that constitute prokaryotic immune system by helping Cas proteins recognize and cut exogenous DNA. Such a classification is not mutually exclusive. Some RNAs were identified initially from one class and later found with roles played in another. For example, circular RNAs (circRNAs) [5] were firstly characterized as ncRNAs with gene regulatory potential [6], some of which were shown later with protein-coding roles [7]. The precise roles of many ncRNAs are still under investigation. For instance, sno-derived RNAs (sdRNAs) are miRNA-like RNAs originated from H/ACA box snoRNAs or C/D box snoRNAs with hypothetical roles during the interplay between RNA silencing and snoRNA-mediated RNA processing systems; miRNA-offset RNAs (moRNAs) are produced from human miRNA precursors but have considerably lower expression levels than the corresponding miRNAs; transcription initiation RNAs (tiRNAs) are mapped within –60 to +120 nucleotides of the transcription start site (TSS) and suggested as a general feature of transcription in possibly all eukaryotes with exact functions uncharacterized.

The diverse types of RNA functionalities largely depend on their interactions with macromolecules, which are often determined by the sequence and structural features of both

interplay partners. This review classifies currently identified RNA interactions into six categories depending on their interplay partners, outlines the capturing technologies in each category as well as the derived RNA functionalities, and identifies understudied niches as future directions to warrant more investigations.

RNA interactions with macromolecules

RNA and DNA are nucleic acids which, along with proteins, lipids and carbohydrates, constitute the four major macromolecules essential for known life forms. RNAs are known to interact with them as well as metabolites to achieve diverse functionalities (Fig. 1).

RNA interactions with DNA

RNA–DNA interactions possess prominent functionalities in diverse biological processes through genetic and epigenetic regulations. NcRNAs are key regulators of chromatin states and gene expression for important biological processes such as dosage compensation [8], imprinting [9], development process [10], lineage differentiation, and disease progression including carcinogenesis [11], where DNA–RNA hybrid formation plays an important role. RNAs provide the templates to orchestrate genome rearrangement in some species such as *Oxytricha* and act as the templates to facilitate DNA translocation in the ‘trans-splicing mediation’ model of non-canonical gene fusions resulted from intergenic splicing [12]. Further, the activity of the telomerase holoenzyme is determined by the binding affinity between the telomerase RNA template region and the DNA primer [13]; the CRISPR

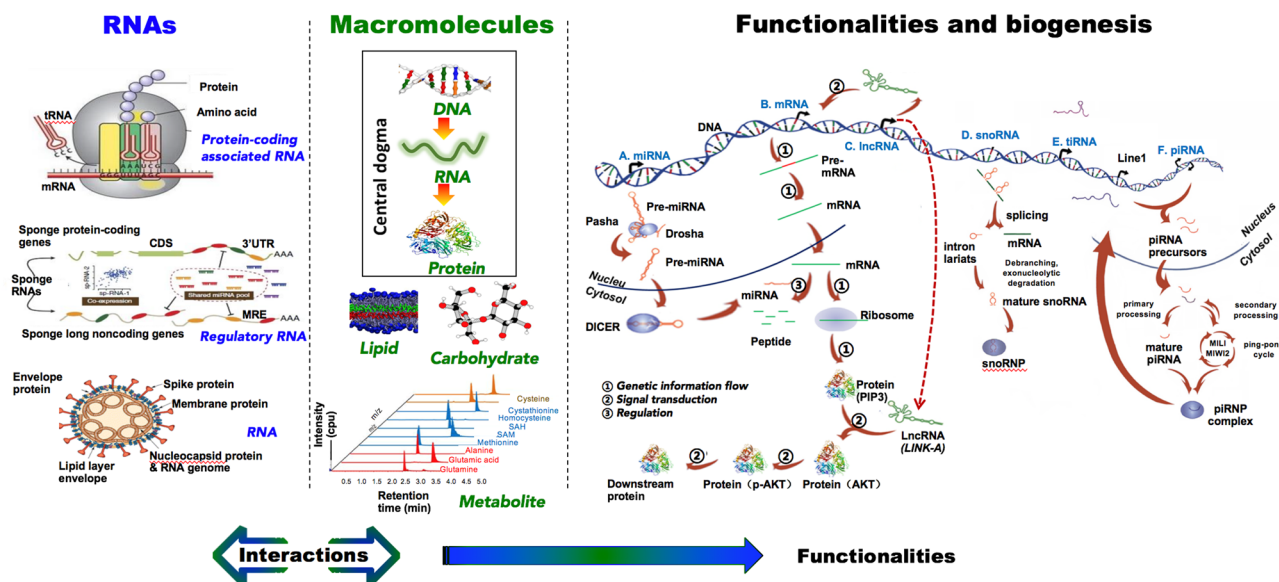


Fig. 1 Conceptual scheme representing the functional RNA realm driven by multi-player interactions. The primary classes, interacting macromolecules and functionalities of RNAs are represented in blue,

green and black, respectively, which collectively convey the message that interactions between RNAs and macromolecules manifest RNA associated functionalities

system against foreign nucleic acid invasion in bacteria and archaea involves DNA/RNA hybrids; steady-state circRNAs have been mapped to thousands of genomic loci in mammals and contribute to gene regulation [6]; and transcriptional pausing can be modeled using sequence-dependent free energy of nucleic acid interactions including DNA–RNA base pairing. While some lncRNAs work in cis on neighboring genes, others function in trans to regulate distantly located genes. For example, while both playing fundamental roles in X chromosome inactivation and dose compensation, human lncRNA Xist functions in cis [8] and *Drosophila* ncRNAs roX1 and roX2 bind numerous regions in trans on the X chromosome of male cells.

Most evidence on the interactions between RNA and DNA was observed at the chromatin level. We extract RNA–DNA interaction modes from this context and summarize them into 4 types. We define ‘type 1’ as ‘RNA interaction with single DNA strand forming R loop’, ‘type 2’ as ‘RNA interaction with double DNA strand forming triple helix’, ‘type 3’ as ‘RNA interaction with single DNA strand without R loop’, and ‘type 4’ as ‘RNA interaction with double DNA strand forming tertiary structure’. Type 1 is the mode used for mRNA elongation in the transcription machinery, where the R loop comprises the nascent RNA hybridized with the DNA template strand and the single-stranded non-template DNA [14]. Recent advances have confirmed the role of type 1 mode RNA–DNA interaction in regulating ncRNA expression. For instance, R loops are formed over the promoter region of the lncRNA COOLAIR and stabilized by AtNDX (a ssDNA-binding homeodomain

protein) to suppress its transcription [15]. The sgRNAs in the CRISPR system function via the type 1 mode for genome-editing, representing an exogenous example (Fig. 2a). The type 2 mode is adopted by many lncRNAs to regulate gene expression in cis or in trans [16]. For instance, promoter-associated RNAs (pRNAs) regulate rRNA gene expression by forming stable RNA–DNA triple helix with the promoter sequence, and the formed triplex is specifically recognized by DNA methyltransferase DNMT3b that results in de novo CpG methylation of rRNA genes (Fig. 2b). It is also reported that specific endogenous miRNAs take the type 2 mode to fight against viral invasion in eukaryotic cells by entangling with double strand DNA viruses such as HIV-1 and forming stable triplexes with viral DNA motifs [17]. The RNA template of telomerase takes the type 3 mode to interact with single DNA strand for chromosome elongation (Fig. 2c). The type 4 mode is a potential model recently proposed where RNA serves as a structural component to maintain the 3D genome conformation (Fig. 2d).

A series of experimental approaches have been developed to characterize RNA–DNA interactions. RNA-centric biochemical purification techniques such as RNA antisense purification (RAP) have enabled comprehensive mapping of RNA–DNA Interactions in vivo [8] (Table 2). Kingston et al. established a hybridization-based technique that specifically enriches endogenous RNAs along with their targets from reversibly cross-linked chromatin extracts, namely CHART (capture hybridization analysis of RNA targets), to map the genomic binding sites for endogenous RNAs [18] (Table 2). Chang et al. developed a method termed ChIRP

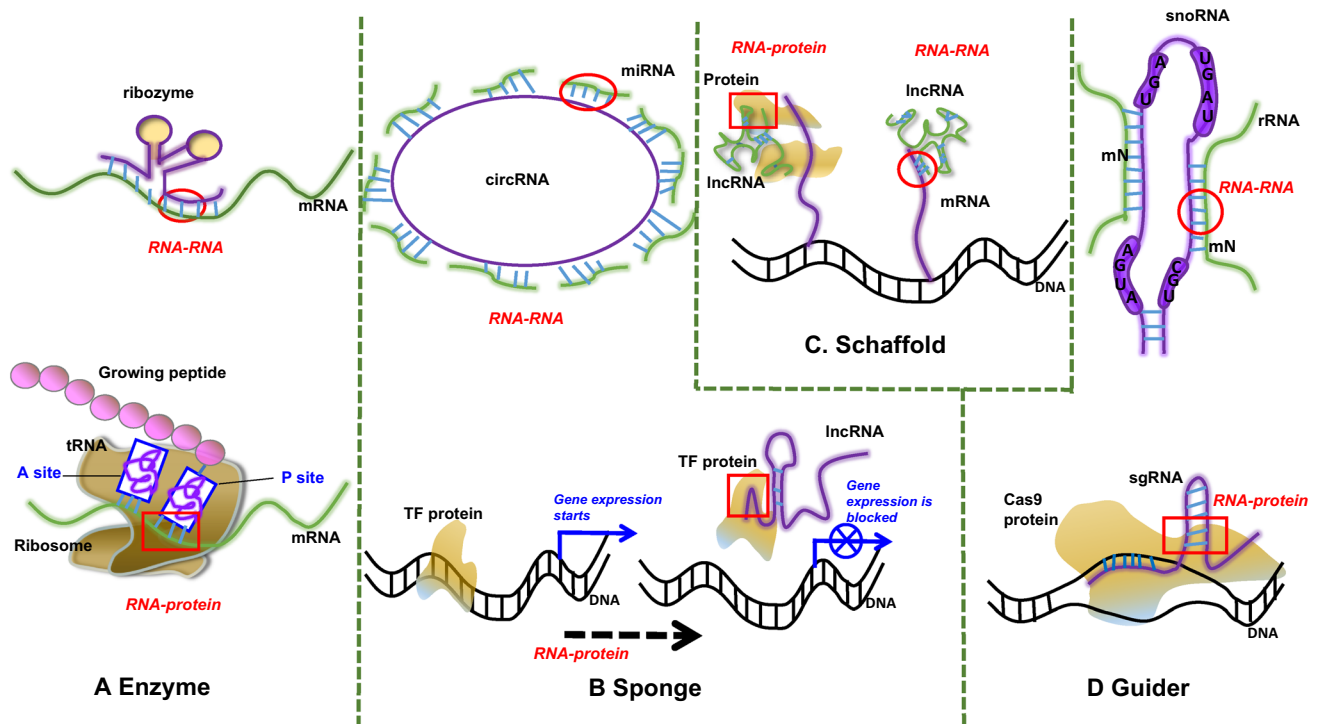


Fig. 2 Schematic representation of RNA-DNA interaction modes and classical examples. **a** Type 1 mode: RNA interaction with single DNA strand forming R loop; **b** Type 2 mode: RNA interaction with

double DNA strand forming triple helix; **c** Type 3 mode: RNA interaction with single DNA strand without R loop; **d** Type 4 mode: RNA interaction with double DNA strand forming tertiary structure

(chromatin isolation by RNA purification) to allow unbiased high-throughput discovery of RNA-bound DNA and proteins *in vivo*, where cultured cells are cross-linked and RNAs of interest are hybridized to target RNAs through biotinylated complementary oligonucleotides followed by magnetic bead isolation (Table 2). While most efforts have been devoted to approaches utilizing cell extracts, Zhong et al. developed MARGI (mapping RNA-genome interactions) to massively reveal native RNA-chromatin interactions from unperturbed cells, which is achieved through RNA–DNA proximity ligation followed by paired-end sequencing of these chimeric sequences (Table 2).

RNA interactions with RNA

RNA–RNA interactions have led to the discovery of many important functions. The milestone in this regard is perhaps the discovery of Alanine transfer RNA and decoding of the triplet system by Robert Holley for which he was awarded with the Nobel Prize. The interactions between tRNA and mRNA transfer genetic information to protein functionalities and fill up the gap in genetic information flow. Another fundamental landmark awarded with Nobel Prize is the regulatory roles of RNAs on gene expression derived from siRNA–mRNA interactions, which has led to technical advances in gene expression modulation, namely

RNA interference. RNA–RNA interactions represent a general strategy used by many ncRNAs to achieve complicated and diverse biological functionalities. For example, genome-wide RNA interactome analysis revealed that TINCR (terminal differentiation-induced ncRNA) interacts with a range of differentiation mRNAs through a 25-nucleotide ‘TINCR box’ motif, and is required for their high mRNA abundance.

Many ncRNAs interact with other RNAs either directly through base-pairing such as miRNA–mRNA, lncRNA–miRNA or snRNA–mRNA hybridization, or indirectly via protein intermediates, e.g., the ribosome is composed of multiple ncRNA components, and numerous lncRNAs associate with proteins to regulate RNA processing (Supplementary Fig. 1a) [19]. The formation of structured RNA such as duplexes represents a critical feature of RNA mediated biological processes. The lead RNA can take diverse roles when interacting with its RNA partner such as enzyme (e.g., ribozyme in mRNA degradation), sponge (e.g., circRNA in miRNA regulation), schaffold (e.g., lncRNA during nascent RNA production), and guider (e.g., snoRNA in rRNA maturation) (Fig. 3). Besides locally positioned RNA structures, long-distance intragenomic RNA–RNA interactions also exist.

Classical RNA–RNA interactions such as U1 snRNA interactions with pre-mRNA were identified through observations of sequence complementarity followed by targeted

Table 2 RNA capture interaction technologies

Technology	Interactions	Crosslinking method	Purification and fragmentation or digestion method	RNA capture method	Reverse crosslink	Ligation, pull down or purification method
AGO-CLIP [23]	RNA–RNA	4-SU and irradiation with ultraviolet light	No purification. Digestion of the lysate with RNase T1 followed by immunoprecipitation of the AGO-RNA complex	AGO antibody capture	No crosslink reversal.	PNK phosphorylation (leaves 3' ends blocked) followed by proximity ligation with T4 RNA Ligase (Ligates 3' hydroxyl end of complete miRNAs to bound RNA fragments. cDNA library prep and sequencing.
hiCLIP [24]	RNA–RNA	Irradiation with ultraviolet light (254 nm)	No purification. Partial digestion of the lysate with RNase I followed by immunoprecipitation of RBP-RNA complex	Antibody capture and RNA based oligonucleotide adaptors (A and B)	No crosslink reversal.	Ligation of two adaptors (A and B) to both strands of the RNA duplex followed by removal of the 3' block from adaptor B and proximity ligation. The RNA hybrid product is then converted into a cDNA library and sequenced.
CLASH [25, 66]	RNA–RNA	Irradiation with ultraviolet light (254 nm)	AGO1-RNA complex purification with IgG-Dynabeads followed by partial digestion with RNases A/T1. Second purification of AGO1-RNA complex with Ni-NTA resin.	Expression of ZZ-TEV-His6-AGO1 (AGO1 acts as bait protein to pull miRNAs) for antibody capture.	Crosslink reversal with Proteinase K.	On Ni-NTA beads: 5' phosphorylation, miRNA target proximity ligation, 3' and 5' dephosphorylation, 3' adaptor ligation and 5' radiolabeling. AGO1-RNA complexes are eluted from nickel beads for protein and RNA analysis.
RAP-RNA [AMT] [26]	RNA–RNA	AMT and irradiation with ultraviolet light (365 nm)	TRIZOL purification. Fragmentation with RNA fragmentation buffer and biotin pull down with Streptavidin beads.	Pool of barcoded antisense 5'-biotinylated ssDNA oligos.	Crosslink reversal by heat (65 °C) and Proteinase K.	No ligation. Purification of nucleic acids with Silane beads. cDNA library prep and sequencing.
RAP-RNA [FA] [26]	RNA–RNA	Formaldehyde	No purification. Fragmentation of the lysate with sonication (DNase I treatment) and purification with Silane beads.	Pool of barcoded antisense 5'-biotinylated ssDNA oligos.	Crosslink reversal by heat (65 °C) and Proteinase K.	No ligation. Biotin pull down with Streptavidin beads and purification with Silane beads. cDNA library prep and sequencing.

Table 2 (continued)

Technology	Interactions	Crosslinking method	Purification and fragmentation or digestion method	RNA capture method	Reverse crosslink	Ligation, pull down or purification method
RAP-RNA/ <i>FA-DSG</i> [26]	RNA–RNA	Formaldehyde and DSG	No purification. Fragmentation of the lysate with sonication (DNase I treatment) and purification with Silane beads.	Pool of barcoded antisense 5'-biotinylated ssDNA oligos.	Crosslink reversal by heat (65 °C) and Proteinase K.	No ligation. Biotin pull down with Streptavidin beads and purification with Silane beads. cDNA library prep and sequencing.
SPLASH [67]	RNA–RNA	EZ-Link-Psoralen-PEG ₃ -biotin and irradiation with ultraviolet light (365 nm)	Magnesium-based fragmentation. Size selection using TBE-Urea gel. Biotin pull down and end repair.	Biotinylated psoralen: EZ-Link-Psoralen-PEG3-biotin.	Crosslink reversal by ultraviolet light radiation (254 nm).	Proximity ligation of chimeras with T4 ligase followed by 3' adaptor ligation and RNA purification with TRIzol. cDNA library prep and sequencing.
LIGR-seq [27]	RNA–RNA	AMT and irradiation with ultraviolet light (365 nm)	TRIzol purification (DNase I treatment and rRNA depletion). Digestion with S1 nuclease.	No capture method.	Crosslink reversal by ultraviolet light radiation (254 nm).	Ligation with circRNA ligase (ligate free overhangs adjacent to duplexes) and purification with phenol–chloroform. cDNA library prep and sequencing.
PARIS [28]	RNA–RNA	AMT and irradiation with ultraviolet light (365 nm)	RNA purification with TRIzol. Digestion with ShortCut RNase III.	No capture method.	Crosslink reversal by ultraviolet light radiation (254 nm).	Proximity ligation of barcoded adapters with T4 RNA ligase, Reverse transcription, cDNA library prep and sequencing.
MARIO [29]	RNA–RNA	Formaldehyde	No purification. Digestion with RNase I or fragmentation with sonication (Turbo DNase treatment). Protein denaturation and biotinylation. Biotinylated RBPs are immobilized on Streptavidin beads.	Protein biotinylation for immobilization of RBPs and biotinylation of RNA probe.	Crosslink reversal by heat (55 °C) and Proteinase K.	On Streptavidin beads: Ligation of biotinylation linker and Proximity ligation. Biotin pull down with Streptavidin beads. cDNA library prep and sequencing.
RAP-DNA [8]	RNA–DNA	Formaldehyde and DSG	No purification. Fragmentation of the lysate (chromatin) with sonication (DNase I treatment).	Multiple biotinylated conjugated RNA probes.	Crosslink reversal by heat and Proteinase K.	No ligation. Biotin pull down with Streptavidin beads. cDNA library prep and sequencing.

Table 2 (continued)

Technology	Interactions	Crosslinking method	Purification and fragmentation or digestion method	RNA capture method	Reverse crosslink	Ligation, pull down or purification method
CHART [18]	RNA–DNA RNA–Protein	Formaldehyde	Fragmentation with sonication. Biotin pull down with Streptavidin beads	Cocktail of biotinylated 3' DNA oligonucleotides.	Crosslink reversal by heat.	No ligation. Protein purification using SDS-PAGE and nucleic acid purification with phenol–chloroform-DNase digestion, reverse transcription, cDNA library prep and sequencing.
ChIRP [68]	RNA–DNA RNA–Protein	Glutaraldehyde	Fragmentation with sonication.	Biotinylated antisense tilting DNA oligonucleotides will hybridize and capture target RNAs.	No crosslink reversal	No ligation. Biotin pull down with Streptavidin beads. Elution of RNAs bound to DNA or proteins with a cocktail of RNase A and H. cDNA library prep and sequencing.
MARGI [69]	RNA–DNA	Formaldehyde and DSG	Digestion with RNase I or sonication & Hae III restriction digestion.	Biotinylated linker adaptor.	Crosslink reversal by heat and Proteinase K.	Proximity ligation. Biotin pull down with Streptavidin beads. cDNA library prep and sequencing.
RIP [70]	RNA–Protein	Formaldehyde	Fragmentation with sonication. Immunoprecipitation with antibody against target protein.	Antibody to capture RNA–protein complex.	Crosslink reversal by heat.	No ligation. RNA purification with TRIzol. RNA analysis by RT-PCR using specific primers. cDNA library prep and sequencing.
RIP-Chip [71]	RNA–Protein	No crosslinking	No purification. No Fragmentation or digestion. The lysate is incubated with pre-bound antibody to protein A/G beads to immunoprecipitate protein-RNA complex. RNA extraction with Trizol	Antibody to capture RNA–protein complex.	Crosslink reversal with Proteinase K.	No ligation. cDNA library prep and sequencing. Optional: Before Proteinase K treatment, beads are suspended and boiled to recover associated proteins for WB analysis. After Proteinase K, proteins associated with the complex can be analyzed by MS or proteomics.

Table 2 (continued)

Technology	Interactions	Crosslinking method	Purification and fragmentation or digestion method	RNA capture method	Reverse crosslink	Ligation, pull down or purification method
RIP-seq [33]	RNA–Protein	No crosslinking	No purification. No fragmentation or digestion. Lysate is incubated with antibody for immunoprecipitation of the protein–RNA complex followed by RNA extraction.	Antibody to capture RNA–protein complex.	No crosslink reversal.	cDNA synthesis. Ligation of the 3' and 5' adaptors, PCR amplification, gel size selection and sequencing.
CLIP [72]	RNA–Protein	Irradiation with ultraviolet light	Fragmentation with sonication and partial digestion with RNase I. Pull down with antibody immunoprecipitation. SDS-PAGE transfer.	Antibody to capture target protein.	Crosslink reversal with Proteinase K.	5' phosphorylation-linker ligation, RT-PCR, cDNA library prep and sequencing.
HITS-CLIP (CLIP-Seq) [34]	RNA–Protein	Irradiation with ultraviolet light (365 nm)	Partial RNA digestion with RNase. The target protein is immunoprecipitated with antibody conjugated to magnetic beads.	³² P-labeled 3' linker.	Crosslink reversal with Proteinase K.	On beads: 3' dephosphorylation, 3' linker ligation (radiolabeled), 5' phosphorylation and 5' linker ligation. Complexes are eluted from beads and separated by SDS-PAGE. Transfer to nitrocellulose membranes RNAs are extracted from membranes by protease treatment for cDNA library prep and sequencing.
PAR-CLIP [35]	RNA–Protein	4-SU or 6-TG and irradiation with ultraviolet light (365 nm)	Digestion with RNase T1 followed by 5' phosphorylation radiolabeling with T4 PNK [gamma- ³² P]. Protein G magnetic beads purification and SDS-PAGE. Purification of crosslinked RNAs from membrane.	No capture method.	Crosslink reversal with Proteinase K.	3' and 5' adaptor ligation, cDNA library prep and sequencing.

Table 2 (continued)

Technology	Interactions	Crosslinking method	Purification and fragmentation or digestion method	RNA capture method	Reverse crosslink	Ligation, pull down or purification method
iCLIP [73]	RNA–Protein	Irradiation with ultraviolet light (254 nm)	Fragmentation with sonication and partial digestion with RNase I. Pull down with antibody immunoprecipitation. SDS-PAGE transfer. RT-PCR with primer containing two cleavable adapter regions and barcodes. Size selection with PAGE-Urea gel.	3' L3 linker adaptor and 5' linker radioactively labeled.	Crosslink reversal with Proteinase K.	Self-ligation of cDNA products (circularization), followed by restriction digestion to linearize cDNA products. High-throughput sequencing.
TRAP/RAT [74]	RNA–Protein RNA–RNA	No crosslinking	No fragmentation. Purification with IgG Dynabeads followed by TEV cleavage. Second purification with Tobramycin resin.	Co-expression constructs: RAT-7SK-PT (contain 5' end of human 7SK RNA tagged with hairpins that bind PP7 coat protein and Tobramycin) and ProteinA-TEV-PP7CP.	No crosslink reversal.	Proteins are digested with trypsin to generate peptides for MS analysis. RNA is purified with phenol–chloroform method, size selection using gel and northern blot analysis.
RiboTrap [75]	RNA–Protein RNA–RNA	No crosslinking	Partial digestion with RNase A. Purification with IgG Sepharose beads followed by protease TEV cleavage.	Co-expression constructs: MS2 CP-TEV-Protein A and MS2 tagged RNA construct MS2 hairpins-pIII/3'UTR.	Crosslink reversal with Proteinase K.	No ligation. Protein pull down using SDS-PAGE and nitrocellulose transfer. RNA pull down with RT-PCR. cDNA library prep and sequencing.
MS2-BioTRAP [36]	RNA–Protein	Irradiation with ultraviolet light (365 nm)	Fragmentation with sonication, no digestion. Purification with MAP-SILAC or PAM-SILAC methods to isolate RNA–Protein complexes from lysates followed by affinity purification with Streptavidin beads. Optional: TEV cleavage. RNA purification with TriZol reagent.	Co-expression RNA constructs: IRES-Luc-MS2-PolyA, CAP-Luc-MS2-PolyA, And MS2 CP-TEV-HB. Stable isotope labeling with amino acids in cell culture (SILAC) to label proteins.	No crosslink reversal.	No ligation. On Streptavidin beads: Proteins are digested with trypsin to generate peptides for LC–MS/MS analysis.

Table 2 (continued)

Technology	Interactions	Crosslinking method	Purification and fragmentation or digestion method	RNA capture method	Reverse crosslink	Ligation, pull down or purification method
CRAC [76]	RNA–Protein	Irradiation with ultraviolet light (254 nm)	Partial digestion with RNase A and T1, followed by Nickel bead purification. On nickel beads: 3' dephosphorylation and 5' radioactive labelling by phosphorylation.	Expression of RBP protein bait fused to His6-TEV-ProteinA (HTP) or CBP-TEV-ProteinA (TAP). Release of His6 tagged proteins by TEV cleavage.	Crosslink reversal with Proteinase K.	On nickel beads: 3' linker ligation 5' 32P-ATP labeling and 5' linker ligation. Pull down using SDS-PAGE and nitrocellulose transfer. Band cut from membrane and RNA extraction for cDNA library prep and sequencing.
RAP-MS [37]	RNA–Protein RNA–RNA	Irradiation with ultraviolet light (254 nm) Before X-link, cells are subjected to SILAC to label proteins and induced to express the lncRNA bait.	Fragmentation with sonication (Turbo DNase treatment) and Streptavidin beads for pull down.	Pool of long antisense 5'-biotinylated ssDNA oligos to capture target RBPs.	Crosslink reversal with Proteinase K and Benzonase.	No ligation. Purification of proteins and nucleic acids with DynaMag magnetic beads. Proteins are precipitated with TCA for MS and RNAs are purified with SILANE beads. cDNA library prep and sequencing.
RaPID [77]	RNA–Protein RNA–RNA	Formaldehyde	Fragmentation with sonication. Blocking of intracellular biotin and pull down of the mRNA::MS2-CP-GFP-SBP, complexes using streptavidin beads.	Co-expression constructs: MS2-aptamer-tagged mRNA and MS2-CP-GFP-SBP	Crosslink reversal by heat (70 °C).	No ligation. For protein fraction: SDS-PAGE, western blot and Silver staining. For RNA fraction: RT-PCR. cDNA library prep and sequencing.
RNA-Compete [78]	RNA–Protein binding specificity	No crosslinking	No fragmentation. GST purification beads.	RNA pool (30–40 nt) generation using microarray and RBP-protein bait construct.	No crosslink reversal.	Microarray (RNA pool generation): 5'Cy3-T7 primer annealing for primer extension and Cy5 linker ligation. Purification of ssDNA, PCR amplification, linker cleavage, RNA seq.

Table 2 (continued)

Technology	Interactions	Crosslinking method	Purification and fragmentation or digestion method	RNA capture method	Reverse crosslink	Ligation, pull down or purification method
SEQRS [39]	RNA–Protein binding specificity	No crosslinking	No purification. No fragmentation or digestion.	DNA library: contain a random region of 20-mer flanked by 20 bp constant regions (primer sites). RNA pool is generated by transcription using designed library.	No crosslink reversal.	No ligation. Protein immobilization on resin to capture RNA–protein complex. RT-PCR. Attached barcodes and sequencing adapters by PCR for cDNA library prep and sequencing.
RBNS [40]	RNA–Protein binding specificity	No crosslinking	No purification. No fragmentation or digestion.	Pool of diverse RNA oligonucleotides containing adaptors. Expression construct: PreScission-GST-RBP-SBP (Express RBP protein bait)	No crosslink reversal.	No ligation. Pull down of RBP with Streptavidin magnetic beads. RNA elution, cDNA library prep and sequencing.
RNA-MITOMI [41]	RNA–Protein binding specificity	No crosslinking	No purification. No fragmentation or digestion.	ssDNA capture probe with 5' biotin- and 3' fluorescein - labeled poly (T). ssDNA capture probe is immobilized in each chamber of the MITOMI chip. RNA library generation by transcription using designed ssDNA oligonucleotides, which are spotted by microarray.	No crosslink reversal.	MITOMI microfluidic chip is overlaid onto the microarray such that each spot was compartmentalized in a unique microchamber. On MITOMI chip: -in vitro transcription. The transcribed poly(A)-tailed RNA molecules hybridize to the immobilized capture probe. A quencher probe is used to quantify RNA capture; -protein detection with fluorescent dye conjugated RBP. Protein pull down is done with Texas red dye conjugated to GST antibody.

Table 2 (continued)

Technology	Interactions	Crosslinking method	Purification and fragmentation or digestion method	RNA capture method	Reverse crosslink	Ligation, pull down or purification method
RNA-MaP [43]	RNA–Protein binding specificity	No crosslinking	No purification. No fragmentation or digestion.	DNA library. DNA molecules in library contain barcode adapter, RNAP promoter, stall site, variant region and RNAP footprint/adaptor. Capture of DNA strands on illumina flow cell and convert individual molecules within the library to clusters on the flow cell through sequencing.	No crosslink reversal.	No ligation. On illumina flow cell: Annealing of 5' biotinylated primer. Reverse transcription (RT). Transcription Initiation by RNA polymerase. Stalled RNAP at the biotin-Streptavidin roadblock. RNA transcription; generates RNA tethered to DNA.
HiTS-RAP [42]	RNA–Protein binding specificity	No crosslinking	No purification. No fragmentation or digestion.	DNA library. DNA molecules contain Tus site to halt RNAP and RNAP T7 promoter. Capture of DNA strands on illumina flow cell to convert Individual molecules within the library to clusters on the flow cell through sequencing.	No crosslink reversal.	No ligation. On illumina flow cell: Annealing of primer. Reverse transcription (RT). Transcription initiation by RNA polymerase. Stalled RNAP at the Tus site. RNA transcription generates RNA tethered to DNA. Bind fluorescently labeled protein. The method determines dissociation constants (Kd values).
MS2-TRAP [45]	RNA-Lipid	No crosslinking	No fragmentation or digestion. Lysate is incubated with GST antibodies for pull down of the Protein-RNA-Lipid complex.	Co-expression constructs: MS2-tagged FL IncRNA (LINK-A) and MS2-CP GST	No crosslink reversal.	No ligation. Assays to detect RNA-Lipid interactions: ELISA PIP3 kinase assay, P13K immunoblotting and RIP assay using PIP3 antibody.

Table 2 (continued)

Technology	Interactions	Crosslinking method	Purification and fragmentation or digestion method	RNA capture method	Reverse crosslink	Ligation, pull down or purification method
NAIM-NAIS [79]	RNA-Metabolite (Riboswitch)	No crosslinking	No purification. No fragmentation or digestion.	Pool of RNAs with phos-phorothioate analogs randomly incorporated. RNAs are selected by self-cleavage reaction.	No crosslink reversal.	Uncleaved RNAs are 5' dephosphorylated. Uncleaved and cleaved RNAs are 5'-end radiolabeled. Radiolabeled RNAs are separated in a polyacrylamide gel and scanned for NAIM and NAIS analysis.

AGO-CLIP crosslinking and immunoprecipitation, *hiCLIP* RNA hybrid and individual-nucleotide resolution UV cross-linking and immunoprecipitation, *CLASH* cross-linking, ligation, and sequencing of hybrids, *RAP-RNA* RNA antisense purification-RNA, *SPLASH* sequencing of psoralen crosslinked, ligated, and selected hybrids, *LIGR-Seq* ligation of interacting RNA followed by high-throughput sequencing, *PARIS* psoralen analysis of RNA interactions and structures, *MARIO* mapping RNA interactome in vivo, *RAP-DNA* RNA antisense purification-DNA, *CHART* capture hybridization analysis of RNA targets, *ChIRP* chromatin isolation by RNA purification, *MARGI* mapping RNA-genome interactions, *RIP* ribonucleoprotein immunoprecipitation, *RIP-Chip* RIP microarray, *RIP-seq* RIP sequencing, *CLIP* UV-crosslinking and immunoprecipitation, *HITS-CLIP* high-throughput sequencing-UV cross-linking and immunoprecipitation, *PAR-CLIP* photoactivatable ribonucleoside-enhanced crosslinking an immunoprecipitation, *iCLIP* individual-nucleotide resolution UV-crosslinking and immunoprecipitation, *TRAP/RAT* tandem RNA-affinity purification/RNA affinity in tandem, *MS2-BioTRAP* MS2 in vivo biotin tagged RNA affinity purification, *CRAC* cross-linking and analysis of cDNAs, *RAP-MS* RNA antisense purification by mass spectrometry, *RaPID* RNA purification and identification, *RNA-Competition* competition between individual RNA sequences binding to proteins, *SEQRS* in vitro selection, high-throughput sequencing of RNA and sequence specificity landscapes, *RBNS* RNA Bind-n-Seq, *RNA-MITOMI* RNA-mechanically induced trapping of molecular interactions, *RNA-MaP* RNA on a massively parallel array, *HITS-RAP* high-throughput sequencing-RNA affinity profiling, *MS2-TRAP* MS2-tagged RNA affinity purification, *NAIM* nucleotide analog interference mapping, *NAIS* nucleotide analog interference suppression, *UV* ultraviolet light, *4-SU* 4-thiouridine, *6-TG* 6-thioguanosine, *X-link* crosslink, *AGO* argonaute protein, *DSG* disuccinimidyl glutarate, *AMT* 4'-aminomethyl trioxsalen (psoralen derivative), *RT-PCR* reverse transcription-polymerase chain reaction, *SDS-PAGE* sodium dodecyl sulfate-polyacrylamide gel electrophoresis, *GST* glutathione-S-transferase, *RBP* RNA binding protein, *RNAP* RNA polymerase, *FAM* fluorescein, *ssDNA* single strand DNA, *RAT* RNA affinity in tandem, *CP* coated protein, *GFP* green fluorescent protein, *SBP* streptavidin binding peptide, *Luc* Luciferase, *TEV* tobacco etch virus protease, *TCA* trichloroacetic acid, *SILAC* stable isotope labeling with amino acids in cell culture, *MAP-SILAC* mix after purification-SILAC, *PAM-SILAC* purification after mixing-SILAC

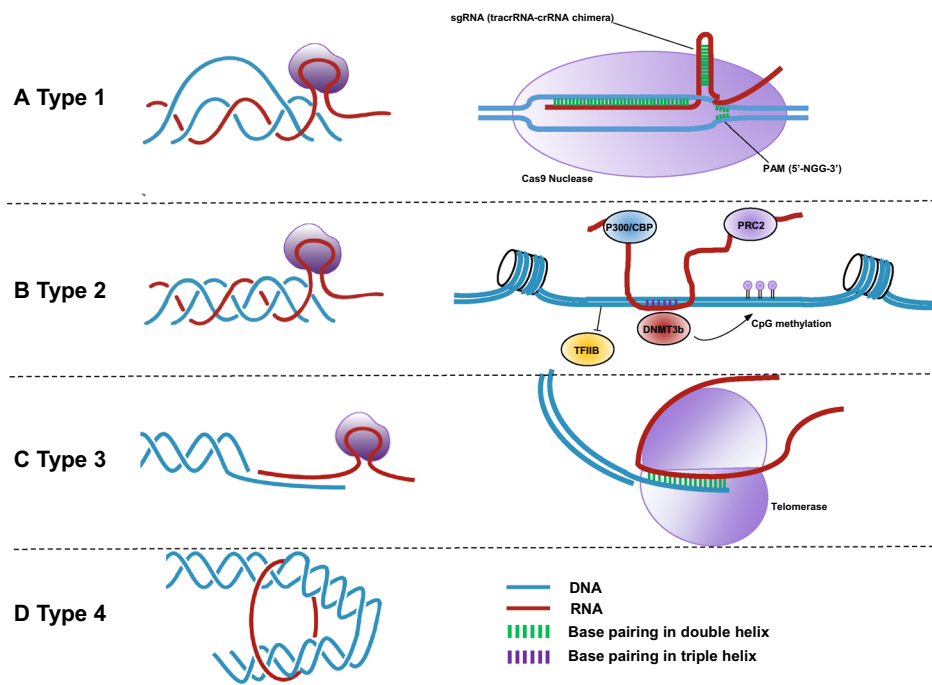


Fig. 3 Typical examples conveying diverse RNA functionalities through RNA-RNA or RNA-protein interactions. **a** Ribozyme-RNA interaction acts as an RNA enzyme for mRNA cleavage; ribosome involves complex interactions between RNA and protein that collectively functions as an enzyme to translate message from mRNA to protein. **b** CircRNA-miRNA interaction acts as a sponge of miRNA and lncRNA-TF interaction functions as a sponge of TF to interfere

with the regulatory role of miRNA and TF, respectively, on gene expression. **c** During lncRNA facilitated nascent RNA production, lncRNA-mRNA and lncRNA-protein function both function as the scaffold. **d** SnoRNAs interact with rRNAs, where snoRNA functions as a guide; The referred RNA-RNA interactions are represented by red circles, and RNA-protein interactions are indicated by red squares. (Color figure online)

genetic modulation and in vitro affinity examination. Many methods have been developed for global-scale mapping of RNA duplexes in vitro and in vivo in a systematic global scale that couple high-throughput sequencing and RNA footprinting strategies utilizing structure-sensitive chemicals (such as structure-seq [20]) or nucleases (such as FragSeq [21], which is short for fragmentation sequencing). Detection of long-range and other tertiary RNA interactions has been achieved by structure-based approaches including nuclear magnetic resonance, X-ray crystallography and cryo-electron microscopy [22]. RNA proximity ligation following immunoprecipitation of complexes of interest given a known interacting protein partner has been launched with success for mapping several types of RNA interactions such as AGO-CLIP (crosslinking and immunoprecipitation) in identifying Argonaute-bound miRNA-mRNA interactions [23], hiCLIP (RNA hybrid and individual-nucleotide resolution UV cross-linking and immunoprecipitation) in recognizing Staufen-bound structured RNAs [24], and CLASH (cross-linking, ligation, and sequencing of hybrids) for characterizing snoRNP-bound snoRNA-rRNA interactions and miRNA interactions [25]. A RAP based method, namely RAP-RNA (RNA antisense purification-RNA), has been proposed to comprehensively characterize in vivo

intermolecular RNA-RNA interactions of a target RNA and distinguish direct and indirect interactions through differential usage of crosslinking reagents with different specificities for proteins and nucleic acids [26] (Table 2). More sophisticated approaches have been continuously developed for globally detecting RNA-RNA interactions in vivo without prior knowledge of RNAs forming interactions and their interacting proteins. Wan et al. proposed the SPLASH (sequencing of psoralen crosslinked, ligated, and selected hybrids) approach to map pairwise RNA interactions in vivo in a genome-wide scale, where EZ-Link-Psoralen-PEG3-Biotin was used to enable biotinylation of RNA via UV-light-activated intercalation of the psoralen group with thymine- and other pyrimidine-containing bases to form covalent bonds and the interacting RNA pairs were identified by sequencing (Table 2). Blencowe et al. proposed LIGR-seq (ligation of interacting RNA followed by high-throughput sequencing) to enable the global-scale mapping of RNA-RNA duplexes cross-linked in vivo, where modified psoralen derivative 4'-aminomethyltrioxalen (AMT) was used to generate AMT-cross-linked RNA duplex followed by high-throughput sequencing of RNA chimeras [27] (Table 2). Chang et al. developed the PARIS (psoralen analysis of RNA interactions and structures) method based on

reversible AMT crosslinking to globally map RNA duplexes in living cells with near base-pair resolution and capture the complexity of RNA structures through direct identification of base-paired helices [28] (Table 2). Sheng et al. presented the MARIO (mapping RNA interactome in vivo) technique to massively reveal RNA–RNA interactions from unperturbed cells by cross-linking RNAs with their bound proteins followed by ligating interacting RNAs to biotinylated RNA linker and subjecting RNA1-Linker-RNA2 to paired-end sequencing [29] (Table 2).

RNA interactions with protein

RNA binding proteins (RBPs) and ribonucleoprotein (RNP) complexes are typical functional forms of RNA–protein interactions, which extensively control, at the post-transcriptional level, the expression of genes in eukaryotes. RBPs such as PRC2 (polycomb repressive complex 2), Argonaute (the key player in all small-RNA-guided gene-silencing processes [30] and interact with approximately 20% lncRNAs), PUM2, QKI, IGF2BP1-3, TNRC61-C and Nova interact with RNAs to drive mRNA processing, nuclear export, translation, localization, turnover, etc. For example, the interaction between the Ezh2 protein domain (a RNA binding subunit of PRC2) and RepA (a 1.6 kb ncRNA residing in lncRNA Xist) recruits PRC2 to the X chromosome that leads to X inactivation; and enforced lncRNA HOTAIR expression re-allocates PRC2 to form an occupancy pattern more resembling embryonic fibroblasts, leading to altered histone H3 lysine 27 methylation, gene expression and, ultimately, increased cancer metastasis [31]. RNPs are RNA–protein complexes that play an integral part in numerous biological processes, e.g., ribosome in protein synthesis, telomerase in chromosome protection, ribonuclease P in small ncRNA transcription, hnRNP (heterogeneous nuclear ribonucleoprotein) in gene transcription and post-transcriptional modification, and snRNP in pre-mRNA splicing.

RNAs can function as the enzyme, sponge, scaffold, guider during RNA–protein interactions. For instance, the ribosomal RNAs convey enzymatic roles for protein production such as the catalytic peptidyl transferase activity in linking amino acids together (Fig. 3a); lncRNA PANDAR acts as a sponge of p53 protein to block its binding with the promoter region of CDKN1A gene in gastric cancer (Fig. 3b) [32]; the lncRNA Malat1 regulates nascent pre-mRNA processing through recruitment or modification of serine/arginine (SR) proteins localized to these sites [26] (Fig. 3c); several types of small RNAs function as the guider of protein enzymes to find their targets for cutting, digestion or modification such as sgRNA in guiding Cas9/dCas9 to certain genomic loci and siRNA in guiding DICER to specific mRNAs (Fig. 3d).

Several approaches have been developed to identify RNA–protein interactions. Garciblancó et al. developed the RIP (ribonucleoprotein immunoprecipitation) assay to study RNA–protein interactions in vivo (Table 2). This technique has been combined with high-throughput technologies and evolved as RIP-Chip (RIP microarray) and RIP-seq (RIP sequencing) [33] to facilitate studies on RNA–protein interactions in the genome-wide scale (Table 2). Darnell et al. developed the CLIP (UV-crosslinking and immunoprecipitation) approach to reveal the interactions between RNAs and their binding proteins through in vivo UV crosslinking and immunoprecipitation (Table 2); and later proposed the HITS-CLIP technique by combining CLIP with high-throughput sequencing to produce transcriptome-wide RNA binding maps with higher accuracy and resolution [34] (Table 2). Tuschl et al. developed a step-by-step protocol for the transcriptome-wide isolation of RNA segments bound by the protein of interest from the background un-crosslinked RNAs, namely PAR-CLIP (photoactivatable ribonucleoside-enhanced crosslinking an immunoprecipitation), via introducing photoreactive nucleosides that generate characteristic sequence alterations on crosslinking, converting isolated crosslinked RNA fragments into a cDNA library followed by deep-sequencing [35] (Table 2). Ule et al. developed individual-nucleotide resolution CLIP (iCLIP) to identify protein–RNA crosslink sites with nucleotide resolution, where the use of two cleavable adapter regions and barcode that enables PCR amplification of truncated cDNAs (the truncated sites represent the cross-linked nucleotide contributes) followed by circularization, linear cleavage and high-throughput sequencing contribute to the trick of this technology (Table 2). TRAP/RAT (tandem RNA-affinity purification/RNA affinity in tandem) does not depend on formaldehyde nor cross-linking, but is an RNA tag-based method for affinity purification of endogenously assembled RNP complexes followed by mass spectrometry protein identification (Table 2). RiboTrap is another approach of this kind that purifies targeted RNPs from cell lysates through immunoaffinity precipitation to describe in vivo endogenous assembly of ribonucleoproteins (Table 2). MS2-BioTRAP (MS2 in vivo biotin tagged RNA affinity purification) co-expresses HB-tagged bacteriophage protein MS2 and stem-loop tagged target RNAs in cells followed by HB-tag based affinity purification of authentic RNA–protein complexes, and proteins associated with target RNAs are subsequently identified and quantified using SILAC-based quantitative mass spectrometry [36] (Table 2).

Various approaches have been proposed or evolved from existing techniques to address specific scientific questions. CRAC (cross-linking and analysis of cDNAs) resolves the restrictions imposed by the use of highly-specific antibodies in CLIP for analyzing protein–RNA interactions in large RNPs containing many different proteins (Table 2).

By adapting the RNA antisense purification (RAP) method to purify a specific lncRNA complexes and identifying the interacting proteins using quantitative mass spectrometry, RAP-MS (RNA antisense purification by mass spectrometry) enables characterization of interacting proteins of a given lncRNA [37] (Table 2). RaPID (RNA purification and identification) allows for the isolation of specific mRNAs of interest and subsequent analysis of the associated proteins using mass spectrometry [38] (Table 2). A couple of methods have been developed to determine RNA–protein binding specificities. While RNA-Compete (competition between individual RNA sequences binding to proteins), SEQRS (in vitro selection, high-throughput sequencing of RNA and sequence specificity landscapes) [39], RBNS (RNA Bind-n-Seq) [40] and RNA-MITOMI (RNA-mechanically induced trapping of molecular interactions) [41] capture RNA–protein binding specificities in vitro, RNA-MaP (RNA on a massively parallel array) and HiTS-RAP (high-throughput sequencing-RNA affinity profiling) [42] achieve this in situ. RNA-Compete uses RNA–protein binding followed by microarray analysis to enable high-throughput identification of RNA binding motifs (Table 2). SEQRS is featured by being capable of identifying RNAs interacting with the protein of interest with low binding affinities and multiple binding modes in a single experiment [39] (Table 2). RBNS novels in measuring the binding affinities both quantitatively and in a high-throughput manner (Table 2). RNA-MITOMI provides a microfluidic platform to allow the identification of new binding specificities through screening the interaction of a RNA mutant library with stem-loop-binding protein [41] (Table 2). RNA-MaP repurposes a high-throughput sequencing instrument to quantitatively measure the binding and dissociation of a fluorescently labeled protein to a library of RNA targets generated on a flow cell surface by in situ transcription [43] (Table 2). HiTS-RAP allows simultaneous transcription of hundreds of millions of DNA clusters that are each derived from amplifying a single molecule with primers covalently linked to the glass flowcell, and the binding of fluorescently labeled proteins to these transcribed RNA clusters is analyzed afterwards [42] (Table 2).

RNA interactions with lipid

The primordial functions of RNAs and lipids have been hypothesized in the ‘RNA world’. The observation that the association of RNA molecules and cellular membranes is involved in forming signal recognition particles in *Escherichia Coli* and regulating cell membrane permeability in *Saccharomyces cerevisiae*, and the fact that all positive-sense (+RNA) RNA viruses remodel intracellular membranes into unique structures for viral genome replication [44] support the notion that RNA-lipid interactions exist and may be physiologically important. A recent report on the

direct interaction between lncRNA LINK-A (long intergenic noncoding RNA for kinase activation) and the phospholipid PI(3,4,5)P3 at the single nucleotide level has broken up the mystery and deciphered the signal transduction role of RNA-lipid interactions [45] (Supplementary Fig. 1a).

In fact, lipids play prominent roles in the life cycle of RNA viruses. Positive strand RNA viruses such as HCV (hepatitis C virus), DENV (dengue virus), ChikV (chikungunya virus) and CoV (coronavirus), PV (picornavirus) replicate their genomic RNA on virus-modified intracellular membranes termed replication organelles (ROs) [44]. The fusion of RNA viruses with host cell membrane, either during the entry or egress process, may imply interactions of RNAs with lipids.

The sole lipid-lncRNA interaction case so far reported was characterized using lncRNA array for RNAs extracted from the lipid-fraction of samples [45]. For instance, the MS2-TRAP (MS2-tagged RNA affinity purification) system (Table 2) was exploited to examine the LINK-A-PIP3 interaction in vivo by expressing MS2-tagged full length LINK-A or Δ PIP3 deletion mutant in cancer cells and analyzing the protein-RNA-lipid complex pulled down by GST antibodies [45]. With the incrementally recognized importance of RNA-lipid interactions and urgent demand for their characterization, fast pace and solid foundation for the development of this field, approaches identifying such interplays at the genome scale are urgently needed and forthcoming.

RNA interactions with metabolite

Riboswitches are genetic control elements found in the 5′-untranslated region of certain mRNAs that exert gene regulatory functionalities through RNA-metabolite interactions. The RNA regulates gene expression through allosteric structural reorganization on site-specific metabolite binding without protein aid (Supplementary Fig. 1b). Several distinct types of riboswitches have been discovered, such as coenzyme B12-binding RNA in the control of *btuB* gene in *Escherichia coli*, TPP (thiamine pyrophosphate) riboswitch in modulating operons in *Escherichia coli* and *Bacillus subtilis*, and FMN-dependent riboswitch in controlling FMN biosynthesis or import in *Bacillus subtilis*. Riboswitches were firstly identified as metabolite-mediated mRNA control of gene expression in prokaryotes, and later found potentially exist in eukaryotes given the presence of TPP binding RNA elements in genes of plants such as *Arabidopsis thaliana*, *Oriza sativa* (rice) and *Poa secunda* (bluegrass), and fungi such as *Neurospora crassa* and *Fusarium oxysporum*.

Metabolite binding triggers allosteric alterations in base-pairing arrangements near gene control elements in prokaryotes, and rearranges the secondary structure of mRNAs that may modulate a greater variety of RNA processing, transport and expression pathways in eukaryotes. For example, TPP

riboswitches function in *Escherichia coli* by sequestering the ribosome-binding site and transcription termination, and potentially regulate mRNA processing and stability in plants and guide RNA splicing in the fungal genome of *Neurospora crassa*.

Current approaches identifying riboswitches that imply RNA-metabolite interactions largely rely on sequencing techniques and genome comparison. NAIM-NAIS (nucleotide analog interference mapping and suppression) represents a direct RNA-metabolite interaction approach. By incorporating phosphorothioate-containing nucleotide analogs into the functional RNA of interest and optimizing a selective activity assay that physically distinguishes functional RNA molecules from inactive molecules, NAIM-NAIS has been used to investigate the interactions between glmS riboswitch and its metabolite GlcN6P (Table 2). Some computational methods have been implemented to study riboswitches, such as the computational modeling of co-transcriptional RNA-ligand interaction dynamics in I-A 2'-deoxyguanosine (a'dG)-sensing riboswitch from *Mesoplasma florum*.

RNA interactions with carbohydrate

According to the 'RNA world' hypothesis, interactions between RNAs and carbohydrates are indispensable. Relatively rare has been reported in this regard. Transglycosylation is important for RNA modification and editing. Though such a process largely relies on the activity of tRNA-guanine transglycosylase and represents RNA-protein interactions, it implicates the importance of carbohydrates in manifesting RNA functionalities. Heparin sulfate (HS) and glycosaminoglycans are polysaccharides abundantly expressed on the surface of many cell types and mediate the entry of many RNA viruses into host cells [46], suggestive of RNA-carbohydrate interactions. However, direct evidence demonstrating RNA-carbohydrate interactions is still lacking, representing an understudied niche.

Functionalities derived from RNA interactions

Genetic information flow

The prominent role of RNAs in genetic information flow has long been recognized, as stated by the Central Dogma of molecular biology that the flow of genetic information moves from DNA to protein via RNA. A living entity is a complex adaptive system that differs from any chemical structure in being capable of self-information processing. Under this conception, RNA viruses have been computationally and experimentally selected to model the origin of genetic information, supporting the notion that RNAs are fundamental in genetic information flow.

As DNA and protein have later evolved to play more specialized roles, RNAs have differentiated into diverse types to interact with macromolecules of the Central Dogma, which are determinant in translating genotypes into phenotypes in a temporal and spatially controlled manner (Fig. 1).

The biological significance of mRNAs, tRNAs and rRNAs during genetic information flow has been well acknowledged, which form the three primary types in the RNA realm. Some RNAs form the critical catalytic part of the enzymes triggering the maturation process of these basic classes of RNAs. For instance, RNase P enzymes consist of both RNA and protein subunits where the RNA fraction is indispensable for catalyzing the removal of the 5' leader sequences from tRNA precursors; RNase MRP (mitochondrial RNA processing) enzymes cleave the A4 site in the ITS (internal transcribed spacer) region of pre-rRNAs to generate mature 5.8S rRNA. Some RNAs serve as the guide to specify the sites of modification for the enzyme to execute the catalytic function. For instance, site-specific synthesis of the 2'-O-methylated nucleotides and pseudouridines in rRNAs and snRNAs is guided by snoRNAs through base-pairing followed by snoRNP protein catalyzed 2'-O-methyl transfer and uridine-to-pseudouridine isomerization reactions. Some RNAs direct the associated proteins to the targeted sites to trigger conformational change of the substrate for subsequent catalysis. For example, snRNAs constitute a large family with at least 10 members; they form ribonucleoprotein complexes (snRNPs) with their associated proteins and bind specific sequences on the pre-mRNA, resulting in the exposure of nucleotides favorable for splicing.

Signal transduction

Regulatory RNAs, e.g., lncRNAs and miRNAs, also participate in the signal transduction of many biological processes.

lncRNAs transduce signals via specific interactions with proteins such as transcriptional factors and RNP complexes. For instance, lncRNA BCAR4 directly binds to two transcription factors (SNIP1, PNUTS) in response to chemokine CCL21, resulting in the release of SNIP1's inhibition on p300-dependent histone acetylation which in turn enables the binding of BCAR4-recruited PNUTS to H3K18ac and relieved inhibition of RNA Pol II via activating the PP1 phosphatase; this activates the Hedgehog/GLI2 transcriptional program that ultimately promotes cell migration [47]. The lncRNA lnc-DC binds directly to STAT3 and promotes STAT3 phosphorylation on tyrosine-705, resulting in activated STAT3 signaling that regulates dendritic cell differentiation [48]. lncRNA MRHL (meiotic recombination hot spot locus) interacts with RNPs such as p68, whose down-regulation results in tyrosine-phosphorylated p68 cytoplasmic translocation, β -catenin nuclear localization, β -catenin-TCF4 interaction, β -catenin occupancy at the

promoter regions of Wnt target genes and ultimately activated Wnt signaling in mouse [49]. The lncRNA NKILA (NF- κ B interacting lncRNA) directly interacts with the functional domains of signaling proteins NF- κ B/I κ B to form a stable complex that prevents over-activation of NF- κ B pathway in inflammation-stimulated breast epithelial cells and ultimately less tumor metastasis [50]. The cytoplasmic lncRNA LINK-A facilitates the recruitment of BRK to the EGFR:GPNMB complex and BRK kinase activation, leading to heterodimer-dependent HIF1 α phosphorylation at Tyr565 and Ser797 by BRK and LRRK2, respectively, HIF1 α stabilization, HIF1 α -p300 interaction, and HIF1 α transcriptional programs under normoxic conditions.

Physiological lncRNA-phospholipid interactions are implicated in mediating signal transduction important for homeostasis and disease. The lncRNA LINK-A has been recently shown to interact with PIP3, facilitating PIP3-AKT interactions that ultimately lead to hyperactivated AKT and poor cancer patient outcome [45].

MiRNAs have also been integrated into signaling pathways as well represented by members of the miRNA-34 family, where activated p53 directly trans-activates the miRNA-34 transcription units that are able to regulate BCL2 and induce apoptosis.

Regulation

The discovery that the genomes of complex organisms including human are largely transcribed into ncRNAs has challenged our traditional conceptions on the functionalities of RNAs and associated inter-molecular interactions.

lncRNAs have attracted much attention due to their large quantity and biological significance, which interact with proteins and function in specific genomic loci to achieve their regulatory activities. Many lncRNAs have been identified to interact with regulatory genomic elements such as promoters, enhancers, ultraconserved regions, introns and epigenomic regulators such as miRNAs [19]. Given that lncRNAs are over 200 nucleotides in length, they could fold into complex three-dimensional structures that allow them achieving the regulatory functionalities by interacting with biomolecule partners such as transcription factors, histones and chromatin-modifying proteins. Consequently, altered lncRNA expression would affect the expression of a broad spectrum of genes via protein partners and lead to profound phenotypic changes as well as severe pathological consequences.

Some small ncRNAs suppress gene expression by directing ribonucleases to the target RNA or DNA sequences. While miRNAs are endogenous interfering RNAs introducing Dicer to silence the target mRNAs, shRNAs and siRNAs are exogenous RNAs and have become standard approaches for gene modulation at the post-transcriptional

level. Similarly, CRISPR RNAs (crRNAs) and trans-activating CRISPR RNAs (transRNAs) orchestrate in leading the Cas9 endonuclease to the target site to fight against foreign genetic elements invasion in the prokaryotic immune system. This has been modified into an effective genome editing tool where crRNA and transRNA are modified into a synthetic guide RNA (gRNA) delivered together with the Cas9 nuclease for the modulation of any desired location in a cell's genome.

CircRNAs have diverse lengths and are more stable than linear RNAs. CircRNAs can function to modulate miRNA activities [51], regulate alternative splicing [52, 53], and sponge other factors such as RBPs or RNP complexes [51]. Being a miRNA activity regulator, circRNA could sponge miRNA complexes and potentially release them on cleavage. For instance, circRNA *0000515* sponges miRNA326 to promote cervical cancer progression via up-regulating *ELK1* [54], circRNA *TTN* acts as a sponge of miRNA432 to facilitate the proliferation and differentiation of myoblasts through the IGF2/PI3K/AKT pathway [55], circRNA *NRIP1* functions as a miRNA149-5p sponger to promote gastric cancer progression via AKT1/mTOR signaling [56], and circRNA *TRIM33-12* acts as the sponge of miRNA191 to suppress hepatocellular cancer progression [57]. The regulatory role of circRNA on alternative splicing is associated with circRNA biogenesis. As pre-mRNA splicing is dictated by competitions between alternative pairing of 5' and 3' splice sites and backsplicing during circRNA generation affects the splicing pattern of the remaining pre-mRNA, circRNAs modulate alternative splicing as a process regardless of the functionalities of circRNAs produced [52]. CircRNAs can also sponge RBPs or RNP complexes to prevent them from acting, provide a protective reservoir of these factors, and deliver them to particular subcellular locations [58]. Lastly, the protein-coding function of circRNAs has been recently identified but still lack sufficient evidence, and novel approaches such as the 'intron-mediated enhancement system' have been proposed to investigate circRNA protein-coding functionalities through increasing circRNA formation [59].

Perspectives

With advances in high-throughput technologies, novel classes of RNAs, their interactions with various macromolecules and the diversified RNA functionalities derived have been continuously deciphered. This has led to a paradigm shift from our conception on the RNA world to the regulatory roles of RNAs in cellular homeostasis and pathogenesis.

Unlike DNAs and proteins, the functionalities of RNAs are largely derived from their interactions with macromolecules including DNAs, RNAs, proteins, lipids, carbohydrates and metabolites. This is especially true in eukaryotes

and concurs with the hypothesis that RNAs devolved some of their functions to DNAs and proteins and pertained the mediating role. RNAs were selected as the primary regulator of many cell events due to their rapid metabolism that is required for dynamic regulation of cellular homeostasis. Though RNAs preserve the information storage capacity in prokaryotes as evidenced by RNA viruses, they are simply genetic information storage unit and need to hijack host materials for propagation. This suggests the prominent roles of macromolecular interactions in driving RNA functionalities in any organisms. Importantly, many biological activities involve multi-player interactions, e.g., RNA and protein function together to maintain chromosome stability for telomerase to take on the action mode. Thus, we cannot dissect any function from its context and leave molecular interactions aside. RNAs interact with other molecules through base pairing, and this actualizes its multiple functionalities such as information transduction and executioner guidance in the case of short regulatory RNAs, and tertiary structure formation in the case of long regulatory RNAs and mRNAs.

RNAs can achieve various roles to support the life cycle of parasites. RNA viruses have RNA as their genetic materials that can be translated on cell infection through interacting with host cellular system. For instance, positive strand RNA viruses perform many functions during viral replication such as viral RNA recruitment for replication, membrane-bound replicase complex assembly, replicase activity regulation, and minus- and plus- strand RNA synthesis through interacting with various host RBPs, and these RBPs include eEF1A, hnRNP proteins and the Lsm 1-7 complex that are conserved across diverse positive strand RNA viruses [60]. Viroids are small nonencapsulated circular ssRNAs that are not protein-coding but replicate in plants as free RNA molecules through RNA structure-based interactions with host cellular factors for functionalities and rely on host DNA-dependent RNA polymerase II for transcription. For instance, potato spindle tuber viroid modulates its replication through directly interacting with a negative splicing regulator RPL5 [61]. CrRNAs constitute prokaryotic immune system by assisting Cas proteins recognize and cut exogenous DNA. On phage infection, proteo-spacers are lysed by bacteria Cas protein complexes and integrated into the 5' end of the CRISPR sites in host genome which are then transcribed into crRNAs and function as the template against phage proto-spacers once infected again [62–64].

Since proteins have long been conceived as principle elements manifesting phenotypes, our focus has been attracted to protein-coding RNAs and interactions of RNAs with primary elements comprising the Central Dogma. Recent discoveries on physiological RNA-phospholipid interactions have challenged our conceptions on RNA functionalities and raised our interest in their involvement in signal transduction and essential cellular processes. Far less effort

has been laid on communications between RNAs and carbohydrates except for the glycosylation potential of tRNAs [65], as well as interactions between RNAs and metabolites except for riboswitches. The story does not stop here and we cannot help ask whether other molecules, such as metallic ions, hormones and ATP interact with RNAs? What are the exact roles played by RNAs if such interplays existed? Do these interactions represent a distinct class of RNA function besides information flow, genetic and epigenetic regulation, as well as signal transduction as afore-discussed?

Conclusion

RNAs had devolved most of its functionalities to other elements such as DNA and protein, and pertained the regulatory roles that are derived from their interplays with diverse macromolecules. Thus, we can presume that while DNAs store inheritable information and proteins take action, RNAs constitute as the primary controller of cell fate as the connectivity of a network determines the topology and consequently the cell state.

Despite the incrementally acknowledged importance of RNAs in life and disease control as well as our accumulated understandings on RNA functionalities, still much remains to be explored. These include less well-studied RNA-macromolecule interactions, working mechanisms of novel RNAs, and experimental techniques facilitating these investigations.

Author contributions XFD conceptualised the ideas, conducted the literature search, prepared the figures, tables, and drafted the manuscript. SZ and KZR contributed to literature search, analysis and preparation of figures and tables. All authors have read and approved the manuscript.

Funding This study was funded by the National Natural Science Foundation of China (Grant No. 81972789), the National Science and Technology Major Project (Grant No. 2018ZX10302205-004-002), the Six Talent Peaks Project in Jiangsu Province (Grant No. SWYY-128), Technology Development Funding of Wuxi (Grant No. WX18IVJN017), Major Project of Science and Technology in Henan Province (Grant No.161100311400). These funding sources have no role in the writing of the manuscript or the decision to submit it for publication.

Compliance with ethical standards

Conflict of interest The authors declares that they have no conflict of interest.

References

1. Djebali S et al (2012) Landscape of transcription in human cells. *Nature* 489(7414):101–108

2. Xie ZY et al (2019) Long non-coding RNA: the functional regulator of mesenchymal stem cells. *World J Stem Cells* 11(3):167–179
3. Kok FO, Baker AH (2019) The function of long non-coding RNAs in vascular biology and disease. *Vascul Pharmacol* 114:23–30
4. Natoli G, Andrau JC (2012) Noncoding transcription at enhancers: general principles and functional models. *Annu Rev Genet* 46:1–19
5. Zhou F et al (2019) Regulation of long non-coding RNAs and circular RNAs in spermatogonial stem cells. *Reproduction*. <https://doi.org/10.1530/REP-18-0517>
6. Lin A et al (2016) The LINK-A lncRNA activates normoxic HIF1 α signalling in triple-negative breast cancer. *Nat Cell Biol* 18(2):213–224
7. Pamudurti NR et al (2017) Translation of CircRNAs. *Mol Cell* 66(1):9–21
8. Engreitz JM et al (2013) The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 341(6147):1237973
9. Koziol MJ, Rinn JL (2010) RNA traffic control of chromatin complexes. *Curr Opin Genet Dev* 20(2):142–148
10. Wang KC et al (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472(7341):120–124
11. Wang Y et al (2015) The long noncoding RNA lncTCF7 promotes self-renewal of human liver cancer stem cells through activation of Wnt signaling. *Cell Stem Cell* 16(4):413–425
12. Jividen K, Li H (2014) Chimeric RNAs generated by intergenic splicing in normal and cancer cells. *Genes Chromosom Cancer* 53(12):963–971
13. Qi X et al (2012) RNA/DNA hybrid binding affinity determines telomerase template-translocation efficiency. *EMBO J* 31(1):150–161
14. Skourti-Stathaki K, Proudfoot NJ (2014) A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev* 28(13):1384–1396
15. Sun Q et al (2013) R-loop stabilization represses antisense transcription at the Arabidopsis FLC locus. *Science* 340(6132):619–621
16. Li Y, Syed J, Sugiyama H (2016) RNA-DNA triplex formation by long noncoding RNAs. *Cell Chem Biol* 23(11):1325–1333
17. Plum GE (2015) Thermodynamics of oligonucleotide triple helices. *Biopolymers* 44(3):241–256
18. Simon MD (2013) Capture hybridization analysis of RNA targets (CHART). *Curr Protoc Mol Biol* 101(1):21–25
19. Zhu Y et al (2018) Long non-coding RNA FOXD2-AS1 contributes to colorectal cancer proliferation through its interaction with microRNA-185-5p. *Cancer Sci* 109(7):2235–2242
20. Ding Y et al (2014) In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* 505(7485):696–700
21. Underwood JG et al (2010) FragSeq: transcriptome-wide RNA structure probing using high-throughput sequencing. *Nat Methods* 7(12):995–1001
22. Feigon J (2015) Back to the future of RNA structure. *RNA* 21(4):611–612
23. Grosswendt S et al (2014) Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. *Mol Cell* 54(6):1042–1054
24. Sugimoto Y et al (2015) hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. *Nature* 519(7544):491–494
25. Helwak A, Tollervey D (2014) Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH). *Nat Protoc* 9(3):711–728
26. Engreitz JM et al (2014) RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent Pre-mRNAs and chromatin sites. *Cell* 159(1):188–199
27. Sharma E et al (2016) Global mapping of human RNA-RNA interactions. *Mol Cell* 62(4):618–626
28. Lu Z et al (2016) RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* 165(5):1267–1279
29. Nguyen TC et al (2016) Mapping RNA-RNA interactome and RNA structure in vivo by MARIO. *Nat Commun* 7:12023
30. Meister G (2013) Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet* 14(7):447–459
31. Gupta RA et al (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464(7291):1071–1076
32. Liu J et al (2018) Long noncoding RNA PANDAR blocks CDKN1A gene transcription by competitive interaction with p53 protein in gastric cancer. *Cell Death Dis* 9(2):168
33. Zhao J et al (2010) Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol Cell* 40(6):939–953
34. Moore MJ et al (2014) Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. *Nat Protoc* 9(2):263–293
35. Spitzer J et al (2014) PAR-CLIP (photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation): a step-by-step protocol to the transcriptome-wide identification of binding sites of RNA-binding proteins. *Methods Enzymol* 539:113–161
36. Tsai BP et al (2011) Quantitative profiling of in vivo-assembled RNA-protein complexes using a novel integrated proteomic approach. *Mol Cell Proteomics* 10(4):M110
37. McHugh CA et al (2015) The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* 521(7551):232–236
38. Ramanathan M et al (2018) RNA-protein interaction detection in living cells. *Nat Methods* 15(3):207–212
39. Campbell ZT et al (2012) Cooperativity in RNA-protein interactions: global analysis of RNA binding specificity. *Cell Rep* 1(5):570–581
40. Lambert N et al (2014) RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol Cell* 54(5):887–900
41. Martin L et al (2012) Systematic reconstruction of RNA functional motifs with high-throughput microfluidics. *Nat Methods* 9(12):1192–1194
42. Tome JM et al (2014) Comprehensive analysis of RNA-protein interactions by high-throughput sequencing-RNA affinity profiling. *Nat Methods* 11(6):683–688
43. Buenrostro JD et al (2014) Quantitative analysis of RNA-protein interactions on a massively parallel array reveals biophysical and evolutionary landscapes. *Nat Biotechnol* 32(6):562–568
44. Dorobantu CM et al (2015) Modulation of the host lipid landscape to promote RNA virus replication: the picornavirus encephalomyocarditis virus converges on the pathway used by hepatitis C virus. *PLoS Pathog* 11(9):e1005185
45. Lin A et al (2017) The LINK-A lncRNA interacts with PtdIns(3,4,5)P3 to hyperactivate AKT and confer resistance to AKT inhibitors. *Nat Cell Biol* 19(3):238–251
46. Mecham JO, McHolland LE (2010) Measurement of blue-tongue virus binding to a mammalian cell surface receptor by an in situ immune fluorescent staining technique. *J Virol Methods* 165(1):112–115
47. Xing Z et al (2014) lncRNA directs cooperative epigenetic regulation downstream of chemokine signals. *Cell* 159(5):1110–1125
48. Wang P et al (2014) The STAT3-binding long noncoding RNA lnc-DC controls human dendritic cell differentiation. *Science* 344(6181):310–313

49. Arun G et al (2012) mrhl RNA, a long noncoding RNA, negatively regulates Wnt signaling through its protein partner Ddx5/p68 in mouse spermatogonial cells. *Mol Cell Biol* 32(15):3140–3152
50. Liu B et al (2015) A cytoplasmic NF-kappaB interacting long noncoding RNA blocks IkappaB phosphorylation and suppresses breast cancer metastasis. *Cancer Cell* 27(3):370–381
51. Memczak S et al (2013) Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495(7441):333–338
52. Ashwal-Fluss R et al (2014) circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* 56(1):55–66
53. Zhang XO et al (2014) Complementary sequence-mediated exon circularization. *Cell* 159(1):134–147
54. Tang Q, Chen Z, Zhao L (2019) Circular RNA hsa_circ_0000515 acts as a miR-326 sponge to promote cervical cancer progression through up-regulation of ELK1. *Aging* 11(22):9982–9999
55. Wang X et al (2019) Circular RNA TTN acts as a miR-432 sponge to facilitate proliferation and differentiation of myoblasts via the IGF2/PI3K/AKT signaling pathway. *Mol Ther Nucleic Acids* 18:966–980
56. Zhang X et al (2019) Circular RNA circNRIP1 acts as a microRNA-149-5p sponge to promote gastric cancer progression via the AKT1/mTOR pathway. *Mol Cancer* 18(1):20
57. Zhang PF et al (2019) Circular RNA circTRIM33-12 acts as the sponge of MicroRNA-191 to suppress hepatocellular carcinoma progression. *Mol Cancer* 18(1):105
58. Du WW et al (2017) Identifying and characterizing circRNA-protein interaction. *Theranostics* 7(17):4183–4191
59. Mo D et al (2019) A universal approach to investigate circRNA protein coding function. *Sci Rep* 9(1):11684
60. Li Z, Nagy PD (2011) Diverse roles of host RNA binding proteins in RNA virus replication. *RNA Biol* 8(2):305–315
61. Jiang J et al (2018) Potato spindle tuber viroid modulates its replication through a direct interaction with a splicing regulator. *J Virol* 92(20):e01004
62. Pan S et al (2019) A seed motif for target RNA capture enables efficient immune defence by a type III-B CRISPR-Cas system. *RNA Biol* 16(9):1166–1178
63. Zhang T et al (2019) Establishing CRISPR/Cas13a immune system conferring RNA virus resistance in both dicot and monocot plants. *Plant Biotechnol J* 17(7):1185–1187
64. Majumdar S, Terns MP (2019) CRISPR RNA-guided DNA cleavage by reconstituted Type I-A immune effector complexes. *Extremophiles* 23(1):19–33
65. Suzuki T, Suzuki T (2014) A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs. *Nucleic Acids Res* 42(11):7346–7357
66. Kudla G et al (2011) Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. *Proc Natl Acad Sci USA* 108(24):10010–10015
67. Aw JG et al (2016) In vivo mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. *Mol Cell* 62(4):603–617
68. Chu C, Quinn J, Chang HY (2012) Chromatin isolation by RNA purification (ChIRP). *J Vis Exp* 61:e3912
69. Sridhar B et al (2017) Systematic mapping of RNA-chromatin interactions in vivo. *Curr Biol* 27(4):610–612
70. Niranjanakumari S et al (2002) Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions in vivo. *Methods* 26(2):182–190
71. Keene JD, Komisarow JM, Friedersdorf MB (2006) RIP-chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat Protoc* 1(1):302–307
72. Ule J et al (2003) CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302(5648):1212–1215
73. Huppertz I et al (2014) iCLIP: protein-RNA interactions at nucleotide resolution. *Methods* 65(3):274–287
74. Hogg JR, Collins K (2007) RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation. *RNA* 13(6):868–880
75. Beach DL, Keene JD (2008) Ribotrap: targeted purification of RNA-specific RNPs from cell lysates through immunoaffinity precipitation to identify regulatory proteins and RNAs. *Methods Mol Biol* 419:69–91
76. Granneman S et al (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc Natl Acad Sci USA* 106(24):9613–9618
77. Slobodin B, Gerst JE (2011) RaPID: an aptamer-based mRNA affinity purification technique for the identification of RNA and protein factors present in ribonucleoprotein complexes. *Methods Mol Biol* 714:387–406
78. Ray D et al (2009) Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. *Nat Biotechnol* 27(7):667–670
79. Soukup JK, Soukup GA (2009) Identification of metabolite-riboswitch interactions using nucleotide analog interference mapping and suppression. *Methods Mol Biol* 540:193–206

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