SCIENCE CHINA
Life Sciences

SPECIAL TOPIC: Noncoding RNA: from dark matter to bright star
•REVIEW•

April 2020 Vol.63 No.4: 501–515
https://doi.org/10.1007/s11427-019-1658-x

Epitranscriptomic technologies and analyses

Xiaoyu Li[1], Qiao-Xia Liang[2], Jin-Ran Lin[2], Jinying Peng[1], Jian-Hua Yang[2,3*], Chengqi Yi[1,4*], Yang Yu[5*], Qiangfeng Cliff Zhang[6*] & Ke-Ren Zhou[2]

[1]*State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing 100871, China;*
[2]*MOE Key Laboratory of Gene Function and Regulation, State Key Laboratory for Biocontrol, School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, China;*
[3]*The Fifth Affiliated Hospital, Sun Yat-Sen University, Zhuhai 519000, China;*
[4]*Department of Chemical Biology and Synthetic and Functional Biomolecules Center, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China;*
[5]*Key Laboratory of RNA Biology, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101 China;*
[6]*MOE Key Laboratory of Bioinformatics, Beijing Advanced Innovation Center for Structural Biology, Center for Synthetic and Systems Biology, Tsinghua-Peking Joint Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China*

Received December 24, 2019; accepted February 12, 2020; published online March 11, 2020

RNA can interact with RNA-binding proteins (RBPs), mRNA, or other non-coding RNAs (ncRNAs) to form complex regulatory networks. High-throughput CLIP-seq, degradome-seq, and RNA-RNA interactome sequencing methods represent powerful approaches to identify biologically relevant ncRNA-target and protein-ncRNA interactions. However, assigning ncRNAs to their regulatory target genes or interacting RNA-binding proteins (RBPs) remains technically challenging. Chemical modifications to mRNA also play important roles in regulating gene expression. Investigation of the functional roles of these modifications relies highly on the detection methods used. RNA structure is also critical at nearly every step of the RNA life cycle. In this review, we summarize recent advances and limitations in CLIP technologies and discuss the computational challenges of and bioinformatics tools used for decoding the functions and regulatory networks of ncRNAs. We also summarize methods used to detect RNA modifications and to probe RNA structure.

ncRNA, bioinformatics, CLIP-seq, RNA modification quantification and locus-specific detection methods, transcriptome-wide sequencing technologies, RNA structuromes, RNA structure probing methods

Citation: Li, X., Liang, Q.X., Lin, J.R., Peng, J., Yang, J.H., Yi, C., Yu, Y., Zhang, Q.C., and Zhou, K.R. (2020). Epitranscriptomic technologies and analyses. Sci China Life Sci 63, 501–515. https://doi.org/10.1007/s11427-019-1658-x

Introduction

Eukaryotic genomes encode tens of thousands of small and large non-coding RNAs (ncRNAs), such as microRNAs (miRNAs), small interfering RNAs (siRNAs), long non-coding RNAs (lncRNAs), small nucleolar RNAs (snoRNAs), and circular RNAs (circRNAs) (Yates et al., 2013; Batista and Chang, 2013; Memczak et al., 2013). Although ncRNAs do not translate into functional proteins, most play a crucial role in biological processes. To achieve their functions, ncRNAs interact with other biomolecules such as RNA-binding proteins (RBPs), mRNA, or other ncRNAs to form complex regulatory networks (König et al., 2012).

*Corresponding authors (Chengqi Yi, email: chengqi.yi@pku.edu.cn; Jian-Hua Yang, email: yangjh7@mail.sysu.edu.cn; Yang Yu, email: yuyang@ibp.ac.cn; Qiangfeng Cliff Zhang, email: qczhang@tsinghua.edu.cn)
The authors contributed equally to this work and are arranged in alphabetic order of surnames.

Various novel experimental and computational methods have been developed to identify the functions and regulatory networks of ncRNAs. Traditional experimental approaches like RNA pull-downs identify ncRNA interactions in a low-throughput manner. With the development of next-generation sequencing, high-throughput methods like crosslinking immunoprecipitation sequencing (CLIP-seq), degradome sequencing (degradome-seq), and RNA-RNA interactome sequencing (RRI-seq) have become widely used for probing RNA functions and interactions (König et al., 2012; Lai and Meyer, 2015). The increasing amount of CLIP-seq, degradome-seq, and RRI-seq sequencing data available results in a great need to develop new computational methods and databases to explore ncRNA-target and ncRNA-protein interactions and construct new regulatory networks involving ncRNAs and proteins.

Following the experimental protocol and high-throughput sequencing, the bioinformatic analysis workflow can be divided into five parts (Figure 1): (i) preprocessing to filter out low-quality reads and PCR duplicates, (ii) mapping reads to the reference genome or transcriptome, (iii) binding site (peak calling) and interaction discovery to distinguish real signal from "noise"/background, (iv) motif finding and characterization, and (v) downstream analyses. Among these, binding site (peak calling) and interaction discovery represent the most critical analyses for identifying specific signals, which represent real RNA binding sites. Here, we discuss the types of bioinformatic tools used for analysis of CLIP-seq, degradome-seq, and RRI-seq data and their core advantages according to their primary function. We also discuss the advantages and disadvantages of current tools and the possible future computational challenges in the RNA interaction field.

In addition to regulating interactions with other biomolecules, chemical modifications to RNA are known to play important roles in regulating gene expression (Roundtree et al., 2017). Greater than 100 distinct chemical modifications to RNA have been identified to date. Most of these modifications are found in non-coding RNAs (ncRNA), such as ribosomal RNA (rRNA), transfer RNA (tRNA), and small nuclear RNA (snRNA). Only a few modifications, including N6-methyladenosine (m^6A), 5-methylcytidine (m^5C), inosine (I), pseudouridine (Ψ), N1-methyladenosine (m^1A), 5-hydroxylmethylcytidine (hm^5C), 2′-O-methylation (Nm), and N4-acetylcytidine (ac^4C), have been found in eukaryotic internal messenger RNA (mRNA) (Boccaletto et al., 2018). Investigation of the functional roles of these modifications in mRNA relies highly on the detection methods used to identify their proportions and locations. Therefore, we summarize the method used for detection of RNA modifications, including quantification methods, locus-specific detection methods, and transcriptome-wide mapping technologies.

An RNA molecule can base pair with itself to form complex secondary and tertiary structures. Genetic variations and mutations that cause aberrations in RNA structure result in many diseases, including cancers and neurodegenerative diseases (Halvorsen et al., 2010). For example, several genetic disorders (including Huntington's disease, myotonic dystrophy, and Fragile X syndrome) are associated with trinucleotide repeat expansion (Osborne and Thornton, 2006). These trinucleotide repeats can form stable RNA hairpin structures that may trigger neurological diseases by interacting with and sequestering RBPs, including important nuclear splicing factors (Krzyzosiak et al., 2012).

Substantial effort has been devoted to solving the structures of RNAs of various classes, including ribozymes, riboswitches, and RNA-protein complexes (Montange and Batey, 2006; Ramakrishnan, 2002; Scott et al., 1995; Thore et al., 2006; Will and Luhrmann, 2011). As of January 2019, ~4,200 structures in the Protein Data Bank (PDB) contain at least one RNA component (Berman et al., 2000). However, despite these efforts, this seemingly large number of RNA structures represents less than 3% of structures in the PDB. Furthermore, many of these structures are rRNAs, tRNAs, and miRNAs. One major reason for this paucity of data is the flexible and dynamic nature of RNA structures (Cruz and Westhof, 2009; Dethoff et al., 2012; Schroeder et al., 2004), which renders them difficult to solve.

Formation of RNA structures is mainly driven by base pairing between nucleotides, stacking of RNA bases, stabilization by metal ions, and modulation of trans factors, including proteins and small ligands (Klein et al., 2004; Mathews et al., 1999; Pyle, 2002; Roth and Breaker, 2009; Williamson, 2000; Woodson, 2005; Xia et al., 1998). At the most fundamental level, base pairing of nucleotides (both Watson-Crick (A-U and G-C) and wobble (G-U) pairing) defines RNA secondary structures by forming hydrogen bonds and releasing free energy. The nearest-neighbor model predicts RNA secondary structures via free energy minimization, using a set of parameters determined in optical melting experiments of short oligonucleotide structures (Xia et al., 1998). A variety of *in silico* methods (including Mfold (Zuker, 2003), the ViennaRNA package (Hofacker et al., 1994), and RNAstructure (Reuter and Mathews, 2010)) are based on these energy terms and have been the primary source of information used to develop RNA secondary structure models and hypotheses. Beyond this fundamental level, however, the RNA folding landscape is very complex, populated with many structural intermediates that prevent a folding pathway from forming the native conformation (Brion and Westhof, 1997; Herschlag, 1995; Leamy et al., 2016; Treiber and Williamson, 1999). Adding to the complexity of RNA folding, RNA structures are dynamic in nature (Cruz and Westhof, 2009; Dethoff et al., 2012; Schroeder et al., 2004). The folding landscape is easily

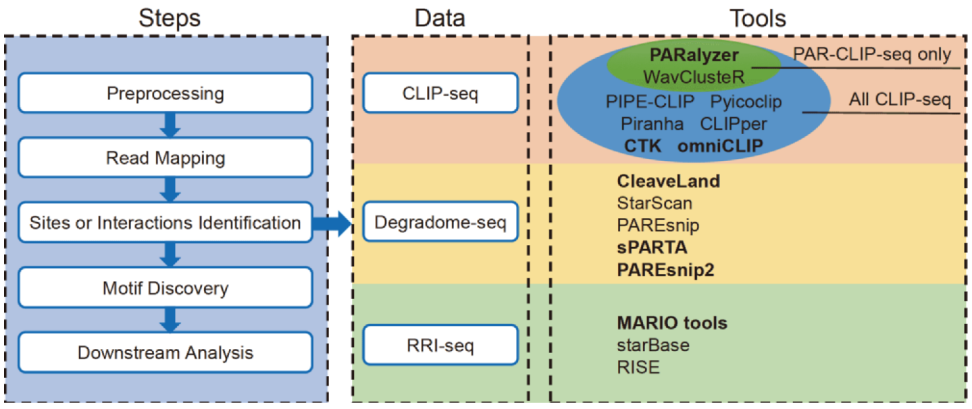

Figure 1 (Color online) The core workflow and bioinformatics tools used for decoding the functions and regulatory networks of ncRNAs. Main steps and tools of the bioinformatics workflow used to analyze CLIP-seq, degradome-seq, and RRI-seq data. The main bioinformatic steps include preprocessing of sequencing reads, read mapping, site or interaction identification, motif discovery and downstream analysis. Among these, identification of sites or interactions is the core analysis step. To identify sites or interactions from CLIP-seq data, relevant software includes PARalyzer, WavClusteR, PIPE-CLIP, Pyicoclip, Piranha, CLIPper, CTK, and omniCLIP. For analysis of degradome-seq data to identify ncRNA-induced mRNA degradations, relevant software includes CleaveLand, StarScan, PAREsnip, sPARTA, and PAREsnip2. MARIO tools and databases like starBase and RISE are used to analyze RRI-seq data.

perturbed by the cellular environment and other cellular processes. Known perturbants include temperature (Qi and Frishman, 2017; Wan et al., 2012), crowding effects (Dupuis et al., 2014; Kilburn et al., 2010), metal ions (Draper, 2004; Klein et al., 2004; Pyle, 2002; Woodson, 2005), protein and ligand binding (Roth and Breaker, 2009; Williamson, 2000), transcription (Heilman-Miller and Woodson, 2003; Kramer and Mills, 1981; Mahen et al., 2010), and as suggested by recent evidence, translation (Beaudoin et al., 2018). The structure of an RNA molecule in a test tube can be drastically different from its structure inside a cell (Leamy et al., 2016). Computational predictions are currently unable to recapitulate the diverse subcellular environments in which RNA folds.

In living cells, interactions with proteins, DNA, other RNAs, and small molecules modulate RNA structures in a very complex fashion. Metabolite-sensing riboswitches are perhaps the best-studied example of how small molecules can regulate RNA structural changes (Dambach and Winkler, 2009; Serganov et al., 2006). In addition to small molecules, hundreds of RBPs interact with RNAs in a sequence- or structure-specific manner or in a promiscuous fashion to impact RNA structure. In particular, ribonucleases and helicases can directly alter RNA structures by enzymatically cleaving RNA strands or unwinding duplex base pairing. This type of modulation is tightly dependent on local concentrations of specific protein regulators and the kinetics of protein-RNA interactions. Recent studies are beginning to reveal that RNA modifications also play important roles in the regulation of RNA structures (Incarnato and Oliviero, 2017; Lewis et al., 2017). RNA modifications are predicted to affect the activity, localization, and stability of RNAs (Li and Mason, 2014; Meyer and Jaffrey, 2014; Zhao et al., 2017). One of the most intriguing functions of RNA modification is modulating RNA structure. For example, studies suggest that pseudouridylation (Ψ) and N6-methyladenosine (m^6A) modifications can affect RNA structural stability and consequently alter the RNA half-life (Lewis et al., 2017). A Ψ modification provides an extra hydrogen bond donor, which could stabilize the RNA structure (Davis, 1995; Zhou et al., 2013). Findings from a recent study suggests that the additional methyl group provided by the m^6A modification could destabilize local RNA duplex structures (Roost et al., 2015). Whether these effects are a general phenomenon *in vivo* and if they extend to other types of modifications remain to be determined. The presence of diverse conformations, the dynamic nature of RNA structures, and most importantly, the complex regulation of RNA folding makes the purely computational prediction of RNA structure an intractable task. Experimental methods to resolve RNA structures are needed to derive more physiologically relevant structure-function relationships.

Advances and limitations in CLIP technologies

CLIP normally begins with crosslinking proteins with their associated RNAs by treatment with 254 nm UV light while cells are still alive to capture endogenous protein:RNA interactions. Following RBP purifications using either an antibody or an epitope tag, the covalently crosslinked protein:RNA complexes are subjected to SDS-PAGE denaturing separation and transferred to nitrocellulose membranes (Lee and Ule, 2018). Because nitrocellulose preferentially retains proteins rather than free RNAs, any RNA recovered from nitrocellulose membranes is assumed to have been covalently attached to the RBPs. These RNAs can be cloned and identified by sequencing.

The original CLIP protocol (Ule et al., 2003) was technically challenging and suffered from the following limita-

tions: (i) UV crosslinking efficiencies; (ii) lengthy gel purification steps and use of radioactivity; and (iii) truncation of cDNAs during reverse transcription (RT) (Lee and Ule, 2018). Many modifications have been made to improve this technically challenging methodology.

First, a number of crosslinking conditions have been adopted to CLIP; these include nucleotide analogs (e.g., 4-thiouridine) and formaldehyde (Hafner et al., 2010; Kim and Kim, 2019). Use of 4-thiouridine can not only improve crosslinking efficiencies for certain proteins using long-wavelength (365 nm) UV light but also helps to resolve crosslink sites due to frequent T-to-C conversions that occur during library constructions (Hafner et al., 2010). For proteins that bind to double-stranded RNAs, a low concentration of formaldehyde has been shown to be more effective for crosslinking than UV light while minimizing unwanted protein-protein crosslinks (Hafner et al., 2010; Kim and Kim, 2019). It is worth noting here that proximity-based RNA editing by ADAR (adenosine deaminase RNA specific) has been adapted to identify RBP targets, completely omitting the requirement for any fixation steps (McMahon et al., 2016).

Second, various tags have been introduced to CLIP that allow for denaturing purifications of RBPs. Nickle column-based His-tag purification is known to be resistant to 8 mol L^{-1} urea washing conditions and when combined another specific tag (e.g., Flag or Biotin) in tandem purifications, effectively removes most, if not all, non-covalent contamination from eukaryotic cells (Sy et al., 2018). The use of covalent tags (e.g., HaloTag) greatly simplifies the purification procedure and allows for use of completely denaturing wash conditions (e.g., GoldCLIP) (Gu et al., 2018). As a result, the SDS-PAGE gel purification steps can be completely omitted and therefore, no radioactivity is needed to visualize protein:RNA complexes. As gel purification is a major source of variation in typical CLIP experiments, GoldCLIP was developed to construct highly reproducible CLIP libraries. More recently, a Flag+SpyTag double tag was used in CLIP, following the same protocol of GoldCLIP (SpyCLIP) (Zhao et al., 2019). However, like CRAC (cross-linking and analysis of cDNAs), tandem purification procedures are required to successfully isolate protein:RNA complexes, due to the inefficient binding of SpyTag.

Third, to avoid linker-linker dimers, 3′ linker ligation to the 3′ end of the RNA can be performed prior to gel purification (Ule et al., 2005). As RT often fails to read through protein:RNA adducts, a cDNA circulation method (iCLIP) was developed not only to recover truncated cDNAs but also to reduce the need for a 5′ end ligation step (Huppertz et al., 2014). An added feature of this cloning method is that the cDNA ends are usually an accurate indication of the protein:RNA crosslink sites. To further improve cloning efficiencies, various RT enzymes have been tested (irCLIP) (Zarnegar et al., 2016). Linkers containing randomized nucleotides are used to reduce PCR duplications during library preparations (eCLIP) (Van Nostrand et al., 2016). In addition, to remove free RT primer, BrdU-CLIP takes advantage of incorporation of a nucleotide analog and a specific antibody to enrich for BrdU-containing cDNAs (Weyn-Vanhentenryck et al., 2014).

Computational methods for determining protein-RNA interactions from CLIP-seq data

CLIP-seq technologies (Chi et al., 2009), including iCLIP (König et al., 2010), PAR-CLIP (Hafner et al., 2010), and eCLIP (Van Nostrand et al., 2016), are the most widely used approaches for genome-wide identification of protein-RNA interactions *in vivo*. Because of the UV crosslinking and immunoprecipitation used in these approaches, conversions, deletions, or truncations occur at or next to the protein-RNA binding sites and are subsequently detected by high-throughput sequencing. Diverse bioinformatic tools have been established to analyze data derived from different CLIP-seq technologies, such as PARalyzer (Corcoran et al., 2011) and WavClusteR (Comoglio et al., 2015) for PAR-CLIP data analysis, and Pyicoclip (Althammer et al., 2011), Piranha (Uren et al., 2012), CLIPper (Lovci et al., 2013), and PIPE-CLIP (Chen et al., 2014) to analyze all types of CLIP-seq data.

PARalyzer (Corcoran et al., 2011) is the most widely used peak calling tool for PAR-CLIP. Since PAR-CLIP uniquely introduces T→C conversions at RBP binding sites, PARalyzer employs a nonparametric kernel-density estimate classifier to identify RNA-protein interaction sites from a combination of T→C conversions and read density.

The CLIP Tool Kit (CTK) (Shah et al., 2016) is a unified software package for analysis of crosslink-induced mutation sites (CIMS analysis) and crosslink-induced truncation sites (CITS analysis). CTK uses a "valley seeking" peak-calling algorithm, in which a peak is called only when valleys of certain depths are found on both sides of the peak and it is separated from other peaks. Therefore, CTK has capacity to separate adjacent peaks without zero-tag gaps. Since CTK uses the same statistical models for both CIMS and CITS analyses, the approach allows for reproducing protein-RNA binding sites in different kinds of CLIP-seq experiments.

OmniCLIP (Drewe-Boss et al., 2018) is the latest published peak-calling tool used for analyses of all types of CLIP-seq data. Unlike CTK, the basic principle of omniCLIP's model is to identify target sites via an unsupervised segmentation of the genome. OmniCLIP learns the relevant diagnostic events directly from immense existing data based on a probabilistic model and then automatically calls

peaks. The joint model of omniCLIP introduces novel error-reducing aspects into CLIP-seq data analysis, leading to results that are more accurate. OmniCLIP models CLIP-seq data in a principled manner, leading to easy integration of other probabilistic models in omniCLIP. OmniCLIP can be easily applied to novel CLIP-seq protocols and has displayed better calibration and reliability than other existing tools.

Though many algorithms have been developed to identify RBP-RNA interactions from CLIP-seq data, analysis of CLIP-seq data still includes many challenges: (i) the proper modeling of diagnostic event types to prevent inaccurate detection of non-binding sites; (ii) a universal method to distinguish authentic interactions from background; (iii) overcoming confounding technical factors, such as RBP binding strength, RNase digestion differences, PCR biases, and most importantly, potential interaction or competition between two RBPs; (iv) using statistical methods to resolve biological variance in diverse samples; and (v) computational methods to identify binding sites at single-nucleotide resolution from various CLIP data. Current bioinformatic tools for CLIP-seq data analysis represent promising approaches to address the above challenges. However, multiple diagnostic event types may appear close to each other and may represent complex diagnostic event types. How these event types influence the joint model for CLIP-seq data analysis remains unexplored.

Bioinformatic tools for analysis of Degradome-seq data to describe ncRNA-induced mRNA degradation

Small ncRNAs (sRNA), including miRNAs and endogenous short interfering RNAs (siRNAs), are recruited into an RNA-induced silencing complex (RISC), pair with mRNAs, and catalyze endonucleolytic cleavages at binding sites/mRNA cleavage sites, the essential step for mRNA degradation. Degradome-seq (German et al., 2008) is a modified version of 5′ rapid amplification of cDNA ends (5′ RACE) combined with high-throughput sequencing to profile uncapped degradation products of mRNAs. Degradome-seq data analysis helps identify sRNA target sites/mRNA cleavage sites on a genome-wide scale and provides a comprehensive understanding of sRNA-guided mRNA degradation. Bioinformatic tools have been developed for target prediction from degradome-seq data; these include CleaveLand (Addo-Quaye et al., 2008), SeqTar (Zheng et al., 2011), PAREsnip (Folkes et al., 2012), sPARTA (Kakrana et al., 2014), StarScan (Liu et al., 2015), and PAREsnip2 (Thody et al., 2018). All of these tools use similar target prediction algorithms that are based on sRNA-mRNA complementation and targeting rules.

CleaveLand (Addo-Quaye et al., 2008) is the oldest and most commonly used bioinformatic tool for processing of degradome-seq data. The algorithm utilizes a mismatch-based scoring scheme based on experimentally validated miRNA-target interaction data from *Arabidopsis thaliana* and complementarity rules for the seed region. CleaveLand hypothesizes a positive correlation between complementarity in the canonical seed region and the probability of authentic cleavage. However, cleavages can occur in regions with poor complementarity or even with mismatches at canonical positions, especially in animal species; thus, CleaveLand may not be reliable in certain applications. CleaveLand is also a time-consuming tool and is thus not suitable for analyzing large amounts of degradome-seq data.

Small RNA-PARE Target Analyzer (sPARTA) (Kakrana et al., 2014) is a novel tool that utilizes a built-in, plant-focused target prediction module called miRferno. Unlike CleaveLand's algorithm, miRferno utilizes two kinds of scoring schemes (standard and seed-free). The standard scoring scheme is based on complementarity rules for the seed region, similar to CleaveLand. The seed-free scoring scheme is a flexible scoring scheme based on genuine miRNA-mRNA interactions instead of strict canonical complementarity rules, helping to identify targets with weak seed-region complementarities or mismatches at canonical positions. Therefore, sPARTA allows for target searching in unannotated genomic regions and discovery of novel sRNA-mRNA interactions.

PAREsnip2 (Thody et al., 2018) employs a binary number system to encode and store sequencing data in a memory-saving fashion and improve analysis speed. The tool utilizes inverse encoding of sRNA and mRNA sequences, so that sRNA and mRNA sequences represented by the same number indicate perfect complementarity. For more precise target prediction, PAREsnip2 utilizes an optimized technique for better alignment of degradome-seq data to the reference genome to distinguish true lower abundance peaks from background degradation of the transcript.

Due to the differing mechanisms of RNA silencing in plants and mammals (mRNA cleavage in plants, translational repression in animals), the abovementioned tools for degradome-seq analysis are more suitable for study of plant miRNAs and neglect certain aspects of animal sRNAs. In addition to miRNAs and sRNAs, thousands of piRNAs participate in RNA degradation. No tools existed for prediction of piRNA cleavage targets until the development of StarScan (Liu et al., 2015). StarScan is the first tool developed for animal piRNA target identification and was designed to search for sRNA targets in multiple genomic regions, including CDSs, 3′ UTRs, exons, 5′ UTRs, introns, and intergenic regions. By integrating 100 degradome-seq datasets from 20 species, StarScan helps discover novel sRNAs' regulatory modules and complementarity rules in

addition to known sRNAs.

Decoding RNA-RNA interactions from RRI-seq data

With the development of high-throughput sequencing, experimental methods like CLASH (Helwak et al., 2013), MARIO (Nguyen et al., 2016), and LIGR-seq (Sharma et al., 2016) facilitate identification of RNA-RNA interactions at the "-omics" level. The fundamental principle of RRI-seq is to form RNA chimeras that are decoded by downstream bioinformatics analysis. Only a single public tool, MARIO tools (Nguyen et al., 2016), is currently available for RNA interactome data analysis.

Even though databases such as starBase (Li and Mason, 2014) and RISE (Gong et al., 2017) have been developed for collecting, processing, and storing RRI-seq data, sophisticated approaches for analysis of RNA chimeras and prediction of high-throughput RNA-RNA interactions have yet to be established. Taking LIGR-seq as an example, a probabilistic model assesses the significance of detected interactions utilizing observed ratios of chimeric reads. Nevertheless, one major class of technical artifacts associated with this model is mapping of intramolecular ligation products to paralogous, overlapping transcripts with different gene IDs or pseudogenes and their subsequent annotation as intermolecular interactions. Numerous complementary approaches must be developed to avoid this technical artifact.

Novel bioinformatic tools are currently being developed to analyze comprehensive RNA interactome data. Mapping RNA interactome *in vivo* (MARIO) (Nguyen et al., 2016) is a novel technology used to reveal RNA-RNA interactions from unperturbed cells. The novel method maps RNA-RNA interactions on a massive scale. MARIO identifies not only protein-assisted inter- and intra-molecular RNA interactions but also a wide variety of RNA interaction events involving mRNAs, lncRNAs, snoRNAs, miRNAs, and other RNAs. In addition, the corresponding suite of bioinformatic tools (MARIO tools) was created to analyze and visualize MARIO data. MARIO tools automate analysis steps, including removing PCR duplicates, splitting multiplexed samples, identifying the linker sequence, splitting junction reads, calling interacting RNAs, performing statistical assessments, categorizing RNA interaction types, calling interacting sites, and analyzing RNA structure as well as providing a comprehensive visualization interface. However, as MARIO tools are a bioinformatic suite specifically designed to be used with the MARIO experimental method, the applicability of MARIO tools to other experimental approaches remains to be examined.

A considerable shortcoming of existing and newly developed approaches for RRI-seq data analysis is their limited applicability to one specific experimental method. Universal algorithms for RRI-seq data analysis are not available. The novel machine learning algorithms were designed to better integrate both inter- and intra-molecular complementarities in a high-throughput way and introduce a universal probabilistic model as the foundation of downstream analysis, making target predictions more reliable.

Methods to quantify RNA modifications

High-sensitivity liquid chromatography-mass spectrometry (LC-MS) is widely used to detect and quantify modifications in mRNA. In this approach, isolated purified mRNA is fully digested into single nucleosides using nuclease P1 and phosphatase. The different nucleosides are separated using liquid chromatography and fragmented into product ions during mass spectrometry. A given nucleoside is identified according to its chromatography retention time, mass-to-charge ratio (m/z), and product ion (Gaston and Limbach, 2014). The ion intensity of nucleosides is concurrently detected using mass spectrometry, and the content of these nucleosides is quantified according to standard curves. This method is fast and sensitive and only requires small amounts of mRNA (~200 ng) to detect and quantify modified nucleotides in low-abundance mRNA. However, this method cannot provide sequence information and requires highly pure mRNA, as the mRNA abundance is low (~1%–3%) and RNA modifications are relatively abundant in ncRNAs. Hence, minimization of ncRNA contamination is a key step in detecting and quantifying mRNA chemical modifications. Two-dimensional thin-layer chromatography (2D-TLC) represents another widely used method to detect mRNA chemical modifications. In this approach, isolated RNA is first partially digested into oligonucleotides using RNase A, RNase T1, or RNase T2. Then, free 5′ OH groups are labeled using ^{32}P. The 5′ end-labeled oligonucleotides are fully digested using nuclease P1, and the ^{32}P-labeled 5-NMPs are separated according to their differing mobility using thin-layer chromatography (TLC). The modified nucleotides in RNA can be identified according to their TLC positions, and direct measurement of each spot's radioactivity enables quantification of these nucleotides (Kellner et al., 2010). This method is low-cost but displays detection bias due to bias of the RNase digestion and the disparate ^{32}P labeling efficiency of these nucleotides (Nees et al., 2014).

Locus-specific RNA modification detection methods

Primer extension is a widely used biochemical assay for

locus-specific detection of RNA modifications. Specific RNA templates are first annealed with labeled primers and then reverse-transcribed into cDNA. This method leverages the ability of certain modified nucleotides in RNA to inhibit extension by reverse transcriptase. RNA templates are reverse-transcribed into cDNA, and the truncated and full-length cDNA fragments are separated on denaturing polyacrylamide gels. The length of the truncated cDNA fragments indicates the position of the modification in the RNA (Motorin et al., 2007). However, this method can be only applied to detect RNA modifications that block base-pairing with or without chemical treatment.

Site-specific cleavage and radioactive-labeling followed by ligation-assisted extraction and TLC (SCARLET) represents another locus-specific detection method that can provide not only position information but also the stoichiometric ratio of modified sites. In this method, a chimeric oligonucleotide comprised of DNA and 2′-O-me-RNA is annealed with target RNA, and then the desired site is cleaved with RNase H. The cleaved 5′ end is labeled using ^{32}P and ligated to a DNA oligo by splint ligation. Following RNase degradation and purification, the ^{32}P-labeled nucleotides are released by nuclease P1 digestion and analyzed by TLC (Liu et al., 2013). Similar to the primer extension assay, SCARLET also requires the sequence information of the modification sites and is only suitable for high-abundance transcripts. Hence, these two locus-specific methods are usually used as orthogonal methods for detection of modified sites in mRNA, and they are used to validate results from transcriptome-wide mapping approaches.

Transcriptome-wide RNA modification sequencing technologies

The development of sequencing technologies for detection of distinct modifications not only provides a resource for detection but also tools for elucidating the functional roles of these mRNA modifications. Existing sequencing technologies are categorized into three types according to their approach: direct, antibody-based, and chemical-assisted sequencing technologies (Table 1).

Some modified nucleotides can directly block or interfere with base-pairing, leading to truncations or misincorporation during reverse transcription. Leveraging these truncations and misincorporation, researchers can directly detect these modified nucleotides in the transcriptome. For example, inosine pairs with cytidine during reverse transcription, and thus, A-to-I editing sites can be identified by comparing the genomic DNA and RNA sequencing data and detecting A-to-G mismatch sites (Bahn et al., 2012; Ju et al., 2011; Peng et al., 2012; Ramaswami et al., 2012). For these methods directly based on RT signatures, eliminating noise caused by single nucleotide polymorphisms (SNPs), somatic mutations, pseudogenes, and sequencing is vital to reducing false

Table 1 Overview of sequencing technologies used to detect mRNA modifications and distribution pattern of mRNA modifications

Modification	Sequencing technologies	Distribution pattern
m^{6}A	m^{6}A-seq (Dominissini et al., 2012)/MeRIP-seq (Meyer et al., 2012) PA-m^{6}A-seq (Chen et al., 2015)/m^{6}A-CLIP (Ke et al., 2015)/ miCLIP (Linder et al., 2015)	Enriched in 3′ UTR and near stop codon (Dominissini et al., 2012; Meyer et al., 2012); Last exon (Ke et al., 2015)
m^{6}Am	m^{6}A-seq (Dominissini et al., 2012) miCLIP (Linder et al., 2015)	Exclusively transcription start sites (Linder et al., 2015)
m^{5}C	Bisulfite-seq (Edelheit et al., 2013; Squires et al., 2012; Yang et al., 2017) Aza-IP (Khoddami and Cairns, 2013)/miCLIP (Hussain et al., 2013)	Enriched downstream of translation initiation sites and in 3′ UTR (Squires et al., 2012; Yang et al., 2017)
hm^{5}C	hMeRIP-seq (Delatte et al., 2016)	Not available
I	Direct sequencing (Bahn et al., 2012; Ju et al., 2011; Peng et al., 2012; Ramaswami et al., 2012) ICE-seq (Sakurai et al., 2014)	Mainly repetitive elements (e.g., Alu and LINE) in UTR and introns (Athanasiadis et al., 2004; Levanon et al., 2004)
Ψ	Ψ-Seq (Schwartz et al., 2014)/Pseudo-seq (Carlile et al., 2014)/ PSI-seq (Lovejoy et al., 2014) CeU-seq (Li et al., 2015)	Mainly CDS and 3′ UTR (Carlile et al., 2014; Li et al., 2015; Schwartz et al., 2014)
m^{1}A	m^{1}A-seq (Dominissini et al., 2016)/m^{1}A-ID-seq (Li et al., 2016) m^{1}A-MAP (Li et al., 2017), single base m^{1}A-seq (Safra et al., 2017)	Enriched in 5′ UTR and near start codon (Dominissini et al., 2016; Li et al., 2016; Li et al., 2017)
Nm	2OMe-seq (Incarnato et al., 2017), RiboMeth-seq (Birkedal et al., 2015; Krogh et al., 2016; Marchand et al., 2016) Nm-seq (Dai et al., 2017)	Enriched in CDS (Dai et al., 2017)
ac^{4}C	acRIP-seq (Arango et al., 2018)	Enriched in 5′ UTR and CDS (Arango et al., 2018)

positives (Wulff et al., 2011). In addition, due to lack of enrichment, this strategy is only suitable for detecting abundant modified sites in mRNA.

In transcriptome-wide sequencing methods based on antibody enrichment, purified mRNA is fragmented to ~100–150 nt in size and immunoprecipitated using an antibody that specifically recognizes the modification of interest. The enriched RNA fragments are then subjected to library construction and high-throughput sequencing. Methods based on this approach have been developed to map m^6A, m^6Am, m^1A, hm^5C, and ac^4C in a transcriptome-wide manner (Arango et al., 2018; Delatte et al., 2016; Dominissini et al., 2012; Dominissini et al., 2016; Li et al., 2016; Meyer et al., 2012). These methods are easily manageable and can be quickly adopted by many different laboratories, enabling many important RNA epigenetics discoveries. However, the resolution of these methods is low (~100–200 nt), limiting the functional studies of modifications that do not have a strong consensus sequence. In addition, these methods rely highly on the specificity and efficiency of antibodies; hence, antibody validation is essential to ensure sensitivity and reduce false positives caused by intrinsic bias due to RNA sequences or antibody structure. To improve the detection resolution of antibody-based sequencing methods, researchers have altered the modification-containing RNA immunoprecipitation protocol by adopting a UV-induced RNA-antibody crosslinking strategy. In this strategy, the protein-RNA crosslinking sites result in truncation or misincorporation during reverse transcription to allow for precise detection of modified sites. Using this strategy, researchers have developed a photo-crosslinking assisted m^6A sequencing strategy (PA-m^6A-seq) (Chen et al., 2015) and an m^6A individual-nucleotide-resolution crosslinking and immunoprecipitation (miCLIP) detection method to obtain base-resolution m^6A methylation profiles (Ke et al., 2015; Linder et al., 2015). For those modified nucleotides that block or interfere with base-pairing, detection methods that combine antibody enrichment and RT signatures not only allow single base-resolution but are also suitable for detection of low-abundance modified sites. For example, researchers have developed m^1A-MAP, which couples antibody-mediated pre-enrichment with m^1A-induced misincorporation detection along with an additional *in vitro* demethylation step (Figure 2). This method not only allows for sensitive detection of m^1A methylation at single-base resolution but also minimizes false signals caused by SNPs or other modifications (Li et al., 2017). In addition, exploiting catalytic mechanisms of the modification enzymes facilitates identification of their direct targets at single-base resolution. For instance, 5-azacytidine-mediated RNA immunoprecipitation (Aza-IP) and methylation iCLIP (miCLIP) were used to successfully identify the direct targets of two m^5C methyltransferases, NSUN2 and DNMT2 (Hussain et al., 2013; Khoddami and Cairns, 2013).

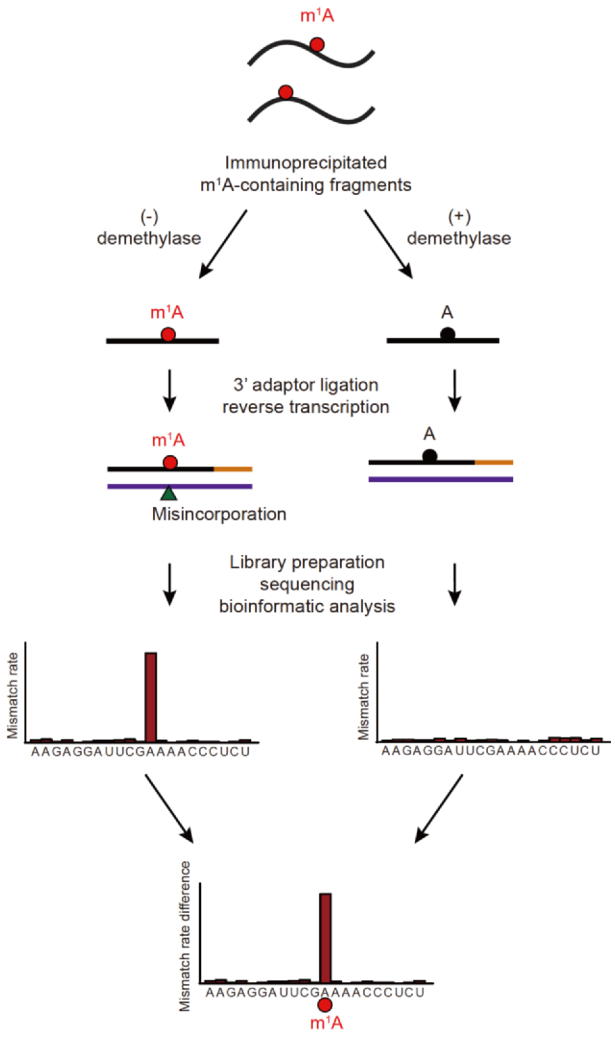

Figure 2 Schematic of m^1A-MAP. The m^1A antibody pre-enriches the m^1A-containing RNA fragments, and RT conditions are optimized to allow for efficient misincorporation during cDNA synthesis. Demethylase treatment is used to improve the confidence of detection. An m^1A methylation site is identified based on the difference and fold change of the mismatch rate between the (–)demethylase and (+)demethylase samples.

For some modified nucleotides, chemical treatment can alter their base-pairing and cause cDNA truncation or misincorporation of nucleotides during reverse transcription. Researchers have developed several chemical-assisted sequencing technologies to detect these nucleotides. For example, the conventional direct sequencing approach for detection of inosine sites is affected by background noise. The compound acrylonitrile selectively reacts with inosine in RNA to form N^1-cyanoethylinosine (ce^1I), which blocks reverse transcription and results in truncation of cDNA. Coupling this chemical reaction to high-throughput sequencing, ICE-seq can profile inosine sites in a transcriptome-wide manner (Sakurai et al., 2014). Bisulfite treatment is a conventional strategy to detect m^5dC in DNA. Combining a modified bisulfite treatment and high-throughput sequen-

cing, researchers have produced transcriptome-wide base-resolution profiles of m^5C (Edelheit et al., 2013; Squires et al., 2012; Yang et al., 2017). Ψ is an isomer of uridine that is able to pair with adenosine during reverse transcription. A chemical called CMCT (1-cyclohexyl-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate) reacts with Ψ, and the resulting CMC-Ψ adducts stall reverse transcription, leading to truncation of cDNA. Combining this chemical reaction with high-throughput sequencing, researchers have developed Ψ-Seq, Pseudo-seq, and PSI-seq to map Ψ transcriptome-wide (Carlile et al., 2014; Lovejoy et al., 2014; Schwartz et al., 2014). However, these methods can only detect highly abundant Ψ sites. A chemically synthesized CMC derivative, azido-CMC (N_3-CMC), can not only selectively label Ψ but also be conjugated to a biotin molecule in a subsequent click reaction. In this approach, named CeU-seq, Ψ-containing RNA can be pre-enriched prior to sequencing (Li et al., 2015), which allows for identification of low-abundance Ψ sites. While most RNA modifications are made to the base, 2′-O-methylation is a type of ribose methylation. Three sequencing technologies have been developed (2OMe-seq, RiboMeth-seq, and Nm-seq) that leverage the different chemical properties of 2′-O-methylation. 2OMe-seq relies on the fact that low amounts of dNTPs stall reverse transcription (Incarnato et al., 2017), while RiboMeth-seq depends on the resistance of 2′-O-methylation to alkaline hydrolysis, compared to normal ribose (Birkedal et al., 2015; Krogh et al., 2016; Marchand et al., 2016). However, these two methods are only suitable for identifying highly abundant methylation sites. In Nm-seq, 2′-O-methylation is resistant to periodate oxidation treatment, and hence, 2′-O-methylation is enriched after several cycles of oxidation–elimination–dephosphorylation steps by removing the unmodified ribose. Using this strategy, low-abundance 2′-O-methylation sites can be detected at single-base resolution.

Experimental methods for probing the RNA structurome

For decades, nucleases and small molecules that modify RNAs in a structure-specific fashion have been used to probe RNA secondary structures. With the advent of sequencing technology, it is now possible to combine these biochemical methods with deep sequencing to assess RNA secondary structures in a massively parallel fashion, allowing for evaluation of the whole transcriptome (i.e., the RNA structurome) in a single experiment. Such studies have revealed many novel insights into RNA structure-function relationships (Bevilacqua et al., 2016; Kubota et al., 2015; Kwok, 2016; Kwok et al., 2016; Mortimer et al., 2014; Nguyen et al., 2016; Piao et al., 2017; Qian et al., 2019; Silverman et al., 2016; Wan et al., 2011).

The first generation of high-throughput methods for RNA structural analysis was developed in the 2010s using enzymatic cleavage. Certain ribonucleases (RNases) have been found to be structure-specific (Chang and Rajbhandary, 1968). For example, RNases S1, A, P1, T1, and U2 cleave RNA single-stranded regions (ssRNA), whereas RNase V1 only cuts double-stranded regions (dsRNA). Each resulting RNA fragment carries one bit of structural information on its 5′ end that is decoded via high-throughput sequencing and subsequent bioinformatic analysis. PARS (Del Campo et al., 2015; Kertesz et al., 2010; Wan et al., 2012; Wan et al., 2014), FragSeq (Underwood et al., 2010), and ssRNA/dsRNA-seq (Gosai et al., 2015; Li et al., 2012a; Li et al., 2012b; Zheng et al., 2010) were among the first techniques developed to assess RNA structure from sequencing data and have been successfully used to determine *in vitro* structuromes for a variety of species. The results from these studies suggest that, even in the absence of other cellular components, most RNA molecules maintain the ability to form stable secondary structures (Kertesz et al., 2010). These approaches have revealed specific structural signatures of many functional sites associated with the regulation of splicing, translation, and miRNA targeting (Kertesz et al., 2010; Wan et al., 2014). However, these techniques have several limitations: RNA needs to be extracted and folded in test tubes before structures are probed, and thus, the resulting observations are based on RNA structures that may differ from their functional *in vivo* state.

Second-generation methods for probing RNA structure involve small molecule modifications (Ding et al., 2014; Rouskin et al., 2014; Spitale et al., 2015). Certain chemicals can modify RNA nucleotides in a specific structural context; these modifications can be decoded as truncations or mutations during reverse transcription. Small molecule modifications can be divided into two types: base modifications and sugar-ring or backbone modifications. DMS, CMCT, and Kethoxal are commonly used to modify certain types of unpaired bases in RNA. By contrast, SHAPE reagents, such as 1M7, NMIA, and NAI, acylate the unpaired ribose without nucleotide bias. Structure-seq (Ding et al., 2014) and DMS-seq (Rouskin et al., 2014) are among the first techniques to combine DMS probing with high-throughput sequencing to decipher RNA structuromes both *in vivo* and *in vitro*. Other methods, such as ChemModSeq (Hector et al., 2014), MAP-seq (Seetin et al., 2014), and CIRS-seq (Incarnato et al., 2014) use DMS to probe only *in vitro* or deproteinized *ex vivo* RNA structures. However, notably, these three methods combine different reagents to increase base coverage. SHAPE-seq (Loughrey et al., 2014; Mortimer et al., 2012) and icSHAPE (Spitale et al., 2015) use 1M7 and NAIto, respectively, to obtain unbiased RNA structuromes of all four bases. As reverse transcribing modified RNAs often

result in mismatch, DMS-MaPseq (Zubradt et al., 2017) and SHAPE-MaP (Siegfried et al., 2014) were developed to detect mismatches to identify RNA structural information. LASER (Feng et al., 2018) uses a light-activated chemical reagent (NAz) to reveal structural changes to adenosine and guanosine. Small molecule probing has allowed for the *in vivo* characterization of RNA structuromes and illuminated the role of RNA structures in transcriptional and post-transcriptional regulation inside cells. In addition, thanks to their smaller size and reduced steric hindrance, small molecule probes can achieve higher structural coverage of transcripts. However, as no existing chemical is known to modify double-stranded RNA regions, we can only probe and obtain structural information for single-stranded regions and are currently unable to distinguish double-stranded nucleotides from missing data.

RNA molecules can fold into complex structural motifs, such as the triplex (Devi et al., 2015) and G-quadruplex (Fay et al., 2017). Formed by stacking G-quartets into special topologies, G-quadruplexes play important roles in RNA regulation. The formation of G-quadruplexes can cause reverse transcriptase to stop. Kwok and Balasubramanian designed *in vitro* assays to identify RT stop positions and hence, the locations of G-quadruplexes (Kwok and Balasubramanian, 2015; Kwok et al., 2016). Their results indicated that G-quadruplexes are significantly enriched around microRNA target sites and polyadenylation signals. Recently, Guo and Bartel combined DMS probing with G-quadruplex RT stop profiling and reported G-quadruplex maps both *in vitro* and *in vivo* (Guo and Bartel, 2016). This study suggested that regions forming G-quadruplexes *in vitro* are largely unfolded *in vivo* in both mouse and yeast cells. Recently, Yang et al. combined affinity capture of G-quadruplexes with sequencing and reported the existence of transient G-quadruplexes (Yang et al., 2018).

Enzymatic cleavage and small molecule modification can only reveal the tendency of a nucleotide to be single-stranded or double-stranded. RNA proximity ligation offers a possible path to direct assessment of base-pairing, as identified through sequencing chimeric reads. Methods vary in the techniques used for the generation and purification of ligated RNA duplexes. CLASH (Helwak et al., 2013; Kudla et al., 2011) and hiCLIP (Sugimoto et al., 2017; Sugimoto et al., 2015) use UV to crosslink protein-RNA complexes and purify RNA duplexes associated with proteins. RPL (RNA proximity ligation) includes *in situ* RNase digestion, proximity ligation, and purification of the whole transcriptome (Ramani et al., 2015). MARIO isolates RNAs crosslinked to proteins and includes proximity ligation to identify protein-dependent RNA-RNA interactions (Nguyen et al., 2016). RAP-RNA uses AMT (4′-aminomethyl trioxsalen) crosslinking to link RNA duplexes and antisense oligos to pull down target RNAs (Engreitz et al., 2015; Kretz et al., 2013).

Recently, a new generation of technologies has been developed that exploits crosslinking agents (e.g., psoralen and its derivatives) to obtain global intra- and inter-molecular RNA duplexes (Nilsen, 2014). PARIS (Lu et al., 2016), SPLASH (Aw et al., 2016), LIGR-seq (Sharma et al., 2016), and COMRADES (Ziv et al., 2018) all use different strategies to enrich crosslinked duplexes, which can then be ligated for deep sequencing and bioinformatics. Important structural information, such as the presence of pseudo-knots, long-range interactions, and alternative structures, can be unveiled using crosslinking techniques (Lu et al., 2016). Crosslinking studies have identified conserved RNA structures across species (Lu et al., 2016), organization of important long non-coding RNAs (e.g., XIST) (Lu et al., 2016), structural bases that affect translation efficiency (Aw et al., 2016), and structural elements affecting virus infection (Li et al., 2018). These methods are still limited, however, in terms of their coverage and resolution.

Discussion

Bioinformatic tools are being rapidly developed for analysis of data from high-throughput sequencing methods for detection of ncRNA interactions with RBPs, mRNAs, and other ncRNAs. However, in most cases, due to the diversity of experimental methods and the complexity of factors that influence RNA interactions, researchers must utilize different methods to preprocess the data input and various models to describe it. Several kinds of software mentioned above attempt to address this problem. The tools include two primary methods. First, some tools model interaction events based on a universal statistical model. This method identifies the most reliable candidates and calculates their respective confidence probabilities for different interaction event types. The use of a core statistical model not only allows for a unified analysis workflow but also provides system metrics (e.g., scores, *P*-values) for assessment and comparison of different types of ncRNA interaction events. Second, other approaches integrate ncRNA interaction analysis by utilizing previous interaction datasets, helping to identify genuine interactions not deemed significant using the above statistical analyses.

Despite these advances, technical artifacts in computational pipelines remain a challenge in the accurate and sensitive detection of RNA interactions. Although degradome-seq analysis allows for precise miRNA-target detection for plant miRNAs, a more accurate analysis approach to profile miRNA-mRNA interactions in animal cells remains to be explored. High-throughput calculation of the resulting data is also necessary; reliable, time-saving algorithms and analysis pipelines are urgently needed for the ever-increasing amount of sequencing data being generated, which presents both

challenges and opportunities. It has long been possible to analyze and visualize the data, but future research will require more efficient data structures and analysis methods suitable for large-scale data. In addition to developing methods for analyzing these interaction data, methods that facilitate biological interpretation and annotation by integrating gene ontologies and cellular pathways are also required.

In the past several years, rapid technological advances have enabled the study of RNA structures at the systems biology level. We hope that, in the near future, the integration of novel chemistry, biotechnology, and computational approaches will result in more advanced structure-probing methods with the power to comprehensively and precisely elucidate RNA structuromes and their functional roles, *in vivo* or *in situ*. We must develop methods that require only small quantities of RNA, allowing the use of limited amounts of tissue. Novel technologies that probe a defined subset of RNAs (e.g., RNAs in specific subcellular localizations, bound by certain proteins, or with the same modifications) are highly desirable. Notably, current studies of the structural signatures of protein-RNA interactions and RNA modifications averaged over ensembles include only a fraction of RNA copies with bound or modified sites (Spitale et al., 2015). To identify accurate structural signatures, a technology that combines the discovery of RBP binding or RNA modifications with structural probing in a single experiment is needed to specifically measure the structures of bound or modified sites.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Author contributions *Prof. Chengqi Yi's team wrote RNA modifications, Prof. Jian-Hua Yang's team wrote bioinformatics methods, Prof. Qiangfeng Zhang's team wrote RNA structures, and Prof. Yang Yu's team wrote biological methods for RNA-protein interaction.*

References

Addo-Quaye, C., Miller, W., and Axtell, M.J. (2008). CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. Bioinformatics 25, 130–131.

Althammer, S., González-Vallinas, J., Ballaré, C., Beato, M., and Eyras, E. (2011). Pyicos: a versatile toolkit for the analysis of high-throughput sequencing data. Bioinformatics 27, 3333–3340.

Andronescu, M., Zhang, Z.C., and Condon, A. (2005). Secondary structure prediction of interacting RNA molecules. J Mol Biol 345, 987–1001.

Arango, D., Sturgill, D., Alhusaini, N., Dillman, A.A., Sweet, T.J., Hanson, G., Hosogane, M., Sinclair, W.R., Nanan, K.K., Mandler, M.D., et al. (2018). Acetylation of cytidine in mRNA promotes translation efficiency. Cell 175, 1872–1886.e24.

Athanasiadis, A., Rich, A., and Maas, S. (2004). Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. PLoS Biol 2, e391.

Aw, J.G.A., Shen, Y., Wilm, A., Sun, M., Lim, X.N., Boon, K.L., Tapsin, S., Chan, Y.S., Tan, C.P., Sim, A.Y.L., et al. (2016). *In vivo* mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. Mol Cell 62, 603–617.

Bahn, J.H., Lee, J.H., Li, G., Greer, C., Peng, G., and Xiao, X. (2012). Accurate identification of A-to-I RNA editing in human by transcriptome sequencing. Genome Res 22, 142–150.

Batista, P.J., and Chang, H.Y. (2013). Long noncoding RNAs: cellular address codes in development and disease. Cell 152, 1298–1307.

Beaudoin, J.D., Novoa, E.M., Vejnar, C.E., Yartseva, V., Takacs, C.M., Kellis, M., and Giraldez, A.J. (2018). Analyses of mRNA structure dynamics identify embryonic gene regulatory programs. Nat Struct Mol Biol 25, 677–686.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The protein data bank. Nucleic Acids Res 28, 235–242.

Bernhart, S.H., Tafer, H., Mückstein, U., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2006). Partition function and base pairing probabilities of RNA heterodimers. Algorithms Mol Biol 1, 3.

Bevilacqua, P.C., Ritchey, L.E., Su, Z., and Assmann, S.M. (2016). Genome-wide analysis of RNA secondary structure. Annu Rev Genet 50, 235–266.

Birkedal, U., Christensen-Dalsgaard, M., Krogh, N., Sabarinathan, R., Gorodkin, J., and Nielsen, H. (2015). Profiling of ribose methylations in RNA by high-throughput sequencing. Angew Chem 127, 461–465.

Boccaletto, P., Machnicka, M.A., Purta, E., Piątkowski, P., Bagiński, B., Wirecki, T.K., de Crécy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A., et al. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 46, D303–D307.

Brion, P., and Westhof, E. (1997). Hierarchy and dynamics of RNA folding. Annu Rev Biophys Biomol Struct 26, 113–137.

Busch, A., Richter, A.S., and Backofen, R. (2008). IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. Bioinformatics 24, 2849–2856.

Carlile, T.M., Rojas-Duran, M.F., Zinshteyn, B., Shin, H., Bartoli, K.M., and Gilbert, W.V. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature 515, 143–146.

Chang, S.H., and Rajbhandary, U.L. (1968). Studies on polynucleotides. LXXXI. Yeast phenylalanine transfer ribonucleic acid: partial digestion with pancreatic ribonuclease. J Biol Chem 243, 592–597.

Chen, B., Yun, J., Kim, M., Mendell, J.T., and Xie, Y. (2014). PIPE-CLIP: a comprehensive online tool for CLIP-seq data analysis. Genome Biol 15, R18.

Chen, K., Lu, Z., Wang, X., Fu, Y., Luo, G.Z., Liu, N., Han, D., Dominissini, D., Dai, Q., Pan, T., et al. (2015). High-resolution N^6-methyladenosine (m^6A) map using photo-crosslinking-assisted m^6A sequencing. Angew Chem Int Ed 54, 1587–1590.

Chi, S.W., Zang, J.B., Mele, A., and Darnell, R.B. (2009). Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps. Nature 460, 479–486.

Comoglio, F., Sievers, C., and Paro, R. (2015). Sensitive and highly resolved identification of RNA-protein interaction sites in PAR-CLIP data. BMC Bioinformatics 16, 32.

Corcoran, D.L., Georgiev, S., Mukherjee, N., Gottwein, E., Skalsky, R.L., Keene, J.D., and Ohler, U. (2011). PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data. Genome Biol 12, R79.

Cruz, J.A., and Westhof, E. (2009). The dynamic landscapes of RNA architecture. Cell 136, 604–609.

Dai, Q., Moshitch-Moshkovitz, S., Han, D., Kol, N., Amariglio, N., Rechavi, G., Dominissini, D., and He, C. (2017). Nm-seq maps 2′-O-methylation sites in human mRNA with base precision. Nat Methods 14, 695–698.

Dambach, M.D., and Winkler, W.C. (2009). Expanding roles for metabolite-sensing regulatory RNAs. Curr Opin Microbiol 12, 161–169.

Davis, D.R. (1995). Stabilization of RNA stacking by pseudouridine. Nucl Acids Res 23, 5020–5026.

Del Campo, C., Bartholomaus, A., Fedyunin, I., and Ignatova, Z. (2015).

Secondary structure across the bacterial transcriptome reveals versatile roles in mRNA regulation and function. PLoS Genet 11, e1005613.
Delatte, B., Wang, F., Ngoc, L.V., Collignon, E., Bonvin, E., Deplus, R., Calonne, E., Hassabi, B., Putmans, P., Awe, S., et al. (2016). Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. Science 351, 282–285.
Dethoff, E.A., Chugh, J., Mustoe, A.M., and Al-Hashimi, H.M. (2012). Functional complexity and regulation through RNA dynamics. Nature 482, 322–330.
Devi, G., Zhou, Y., Zhong, Z., Toh, D.F.K., and Chen, G. (2015). RNA triplexes: from structural principles to biological and biotech applications. Wires RNA 6, 111–128.
Ding, Y., Tang, Y., Kwok, C.K., Zhang, Y., Bevilacqua, P.C., and Assmann, S.M. (2014). *In vivo* genome-wide profiling of RNA secondary structure reveals novel regulatory features. Nature 505, 696–700.
Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., et al. (2012). Topology of the human and mouse m^6A RNA methylomes revealed by m^6A-seq. Nature 485, 201–206.
Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M.S., Dai, Q., Di Segni, A., Salmon-Divon, M., Clark, W.C., et al. (2016). The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. Nature 530, 441–446.
Draper, D.E. (2004). A guide to ions and RNA structure. RNA 10, 335–343.
Drewe-Boss, P., Wessels, H.H., and Ohler, U. (2018). omniCLIP: probabilistic identification of protein-RNA interactions from CLIP-seq data. Genome Biol 19, 183.
Dupuis, N.F., Holmstrom, E.D., and Nesbitt, D.J. (2014). Molecular-crowding effects on single-molecule RNA folding/unfolding thermodynamics and kinetics. Proc Natl Acad Sci USA 111, 8464–8469.
Edelheit, S., Schwartz, S., Mumbach, M.R., Wurtzel, O., and Sorek, R. (2013). Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m^5C within archaeal mRNAs. PLoS Genet 9, e1003602.
Engreitz, J., Lander, E.S., and Guttman, M. (2015). RNA antisense purification (RAP) for mapping RNA interactions with chromatin. Methods Mol Biol 1262, 183–197.
Fay, M.M., Lyons, S.M., and Ivanov, P. (2017). RNA G-quadruplexes in biology: principles and molecular mechanisms. J Mol Biol 429, 2127–2147.
Feng, C., Chan, D., Joseph, J., Muuronen, M., Coldren, W.H., Dai, N., Corrêa Jr, I.R., Furche, F., Hadad, C.M., and Spitale, R.C. (2018). Light-activated chemical probing of nucleobase solvent accessibility inside cells. Nat Chem Biol 14, 276–283.
Folkes, L., Moxon, S., Woolfenden, H.C., Stocks, M.B., Szittya, G., Dalmay, T., and Moulton, V. (2012). PAREsnip: a tool for rapid genome-wide discovery of small RNA/target interactions evidenced through degradome sequencing. Nucleic Acids Res 40, e103.
Gaston, K.W., and Limbach, P.A. (2014). The identification and characterization of non-coding and coding RNAs and their modified nucleosides by mass spectrometry. RNA Biol 11, 1568–1585.
German, M.A., Pillay, M., Jeong, D.H., Hetawal, A., Luo, S., Janardhanan, P., Kannan, V., Rymarquis, L.A., Nobuta, K., German, R., et al. (2008). Global identification of microRNA–target RNA pairs by parallel analysis of RNA ends. Nat Biotechnol 26, 941–946.
Gong, J., Shao, D., Xu, K., Lu, Z., Lu, Z.J., Yang, Y.T., and Zhang, Q.C. (2017). RISE: a database of RNA interactome from sequencing experiments. Nucleic Acids Res 46, D194–D201.
Gosai, S.J., Foley, S.W., Wang, D., Silverman, I.M., Selamoglu, N., Nelson, A.D.L., Beilstein, M.A., Daldal, F., Deal, R.B., and Gregory, B.D. (2015). Global analysis of the RNA-protein interaction and RNA secondary structure landscapes of the *Arabidopsis* nucleus. Mol Cell 57, 376–388.
Gu, J., Wang, M., Yang, Y., Qiu, D., Zhang, Y., Ma, J., Zhou, Y., Hannon, G.J., and Yu, Y. (2018). GoldCLIP: gel-omitted ligation-dependent CLIP. Genomics Proteomics Bioinformatics 16, 136–143.
Guo, J.U., and Bartel, D.P. (2016). RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. Science 353, aaf5371.
Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano Jr., M., Jungkamp, A.C., Munschauer, M., et al. (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141, 129–141.
Halvorsen, M., Martin, J.S., Broadaway, S., and Laederach, A. (2010). Disease-associated mutations that alter the RNA structural ensemble. PLoS Genet 6, e1001074.
Hector, R.D., Burlacu, E., Aitken, S., Bihan, T.L., Tuijtel, M., Zaplatina, A., Cook, A.G., and Granneman, S. (2014). Snapshots of pre-rRNA structural flexibility reveal eukaryotic 40S assembly dynamics at nucleotide resolution. Nucleic Acids Res 42, 12138–12154.
Heilman-Miller, S.L., and Woodson, S.A. (2003). Effect of transcription on folding of the Tetrahymena ribozyme. RNA 9, 722–733.
Helwak, A., Kudla, G., Dudnakova, T., and Tollervey, D. (2013). Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell 153, 654–665.
Herschlag, D. (1995). RNA chaperones and the RNA folding problem. J Biol Chem 270, 20871–20874.
Hofacker, I.L., Fontana, W., Stadler, P.F., Bonhoeffer, L.S., Tacker, M., and Schuster, P. (1994). Fast folding and comparison of RNA secondary structures. Monatsh Chem 125, 167–188.
Huppertz, I., Attig, J., D'Ambrogio, A., Easton, L.E., Sibley, C.R., Sugimoto, Y., Tajnik, M., König, J., and Ule, J. (2014). iCLIP: Protein–RNA interactions at nucleotide resolution. Methods 65, 274–287.
Hussain, S., Sajini, A.A., Blanco, S., Dietmann, S., Lombard, P., Sugimoto, Y., Paramor, M., Gleeson, J.G., Odom, D.T., Ule, J., et al. (2013). NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. Cell Rep 4, 255–261.
Incarnato, D., and Oliviero, S. (2017). The RNA epistructurome: uncovering RNA function by studying structure and post-transcriptional modifications. Trends Biotech 35, 318–333.
Incarnato, D., Anselmi, F., Morandi, E., Neri, F., Maldotti, M., Rapelli, S., Parlato, C., Basile, G., and Oliviero, S. (2017). High-throughput single-base resolution mapping of RNA 2′-O-methylated residues. Nucleic Acids Res 45, 1433–1441.
Incarnato, D., Neri, F., Anselmi, F., and Oliviero, S. (2014). Genome-wide profiling of mouse RNA secondary structures reveals key features of the mammalian transcriptome. Genome Biol 15, 491.
Ju, Y.S., Kim, J.I., Kim, S., Hong, D., Park, H., Shin, J.Y., Lee, S., Lee, W. C., Kim, S., Yu, S.B., et al. (2011). Extensive genomic and transcriptional diversity identified through massively parallel DNA and RNA sequencing of eighteen Korean individuals. Nat Genet 43, 745–752.
Kakrana, A., Hammond, R., Patel, P., Nakano, M., and Meyers, B.C. (2014). sPARTA: a parallelized pipeline for integrated analysis of plant miRNA and cleaved mRNA data sets, including new miRNA target-identification software. Nucleic Acids Res 42, e139.
Ke, S., Alemu, E.A., Mertens, C., Gantman, E.C., Fak, J.J., Mele, A., Haripal, B., Zucker-Scharff, I., Moore, M.J., Park, C.Y., et al. (2015). A majority of m^6A residues are in the last exons, allowing the potential for 3′ UTR regulation. Genes Dev 29, 2037–2053.
Kellner, S., Burhenne, J., and Helm, M. (2010). Detection of RNA modifications. RNA Biol 7, 237–247.
Kertesz, M., Wan, Y., Mazor, E., Rinn, J.L., Nutter, R.C., Chang, H.Y., and Segal, E. (2010). Genome-wide measurement of RNA secondary structure in yeast. Nature 467, 103–107.
Khoddami, V., and Cairns, B.R. (2013). Identification of direct targets and modified bases of RNA cytosine methyltransferases. Nat Biotechnol 31, 458–464.
Kilburn, D., Roh, J.H., Guo, L., Briber, R.M., and Woodson, S.A. (2010). Molecular crowding stabilizes folded RNA structure by the excluded

volume effect. J Am Chem Soc 132, 8690–8696.
Kim, B., and Kim, V.N. (2019). fCLIP-seq for transcriptomic footprinting of dsRNA-binding proteins: Lessons from DROSHA. Methods 152, 3–11.
Klein, D.J., Moore, P.B., and Steitz, T.A. (2004). The contribution of metal ions to the structural stability of the large ribosomal subunit. RNA 10, 1366–1379.
König, J., Zarnack, K., Luscombe, N.M., and Ule, J. (2012). Protein–RNA interactions: new genomic technologies and perspectives. Nat Rev Genet 13, 77–83.
König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol 17, 909–915.
Kramer, F.R., and Mills, D.R. (1981). Secondary structure formation during RNA synthesis. Nucl Acids Res 9, 5109–5124.
Kretz, M., Siprashvili, Z., Chu, C., Webster, D.E., Zehnder, A., Qu, K., Lee, C.S., Flockhart, R.J., Groff, A.F., Chow, J., et al. (2013). Control of somatic tissue differentiation by the long non-coding RNA TINCR. Nature 493, 231–235.
Krogh, N., Jansson, M.D., Häfner, S.J., Tehler, D., Birkedal, U., Christensen-Dalsgaard, M., Lund, A.H., and Nielsen, H. (2016). Profiling of 2′-*O*-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. Nucleic Acids Res 44, 7884–7895.
Krzyzosiak, W.J., Sobczak, K., Wojciechowska, M., Fiszer, A., Mykowska, A., and Kozlowski, P. (2012). Triplet repeat RNA structure and its role as pathogenic agent and therapeutic target. Nucleic Acids Res 40, 11–26.
Kubota, M., Chan, D., and Spitale, R.C. (2015). RNA structure: merging chemistry and genomics for a holistic perspective. Bioessays 37, 1129–1138.
Kudla, G., Granneman, S., Hahn, D., Beggs, J.D., and Tollervey, D. (2011). Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. Proc Natl Acad Sci USA 108, 10010–10015.
Kwok, C.K. (2016). Dawn of the *in vivo* RNA structurome and interactome. Biochem Soc Trans 44, 1395–1410.
Kwok, C.K., and Balasubramanian, S. (2015). Targeted detection of G-quadruplexes in cellular RNAs. Angew Chem Int Ed 54, 6751–6754.
Kwok, C.K., Marsico, G., Sahakyan, A.B., Chambers, V.S., and Balasubramanian, S. (2016). rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome. Nat Methods 13, 841–844.
Lai, D., and Meyer, I.M. (2015). A comprehensive comparison of general RNA–RNA interaction prediction methods. Nucleic Acids Res 44, e61.
Leamy, K.A., Assmann, S.M., Mathews, D.H., and Bevilacqua, P.C. (2016). Bridging the gap between *in vitro* and *in vivo* RNA folding. Quart Rev Biophys 49, e10.
Lee, F.C.Y., and Ule, J. (2018). Advances in CLIP technologies for studies of protein-RNA interactions. Mol Cell 69, 354–369.
Levanon, E.Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., Shemesh, R., Fligelman, Z.Y., Shoshan, A., Pollock, S.R., Sztybel, D., et al. (2004). Systematic identification of abundant A-to-I editing sites in the human transcriptome. Nat Biotechnol 22, 1001–1005.
Lewis, C.J.T., Pan, T., and Kalsotra, A. (2017). RNA modifications and structures cooperate to guide RNA–protein interactions. Nat Rev Mol Cell Biol 18, 202–210.
Li, F., Zheng, Q., Ryvkin, P., Dragomir, I., Desai, Y., Aiyer, S., Valladares, O., Yang, J., Bambina, S., Sabin, L.R., et al. (2012a). Global analysis of RNA secondary structure in two metazoans. Cell Rep 1, 69–82.
Li, F., Zheng, Q., Vandivier, L.E., Willmann, M.R., Chen, Y., and Gregory, B.D. (2012b). Regulatory impact of RNA secondary structure across the *Arabidopsis* transcriptome. Plant Cell 24, 4346–4359.
Li, J.H., Liu, S., Zhou, H., Qu, L.H., and Yang, J.H. (2013). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucl Acids Res 42, D92–D97.
Li, P., Wei, Y., Mei, M., Tang, L., Sun, L., Huang, W., Zhou, J., Zou, C., Zhang, S., Qin, C.F., et al. (2018). Integrative analysis of Zika virus genome RNA structure reveals critical determinants of viral infectivity. Cell Host Microbe 24, 875–886.e5.
Li, S., and Mason, C.E. (2014). The pivotal regulatory landscape of RNA modifications. Annu Rev Genom Hum Genet 15, 127–150.
Li, X., Xiong, X., Wang, K., Wang, L., Shu, X., Ma, S., and Yi, C. (2016). Transcriptome-wide mapping reveals reversible and dynamic N^1-methyladenosine methylome. Nat Chem Biol 12, 311–316.
Li, X., Xiong, X., Zhang, M., Wang, K., Chen, Y., Zhou, J., Mao, Y., Lv, J., Yi, D., Chen, X.W., et al. (2017). Base-resolution mapping reveals distinct m^1A methylome in nuclear- and mitochondrial-encoded transcripts. Mol Cell 68, 993–1005.e9.
Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F., and Yi, C. (2015). Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol 11, 592–597.
Linder, B., Grozhik, A.V., Olarerin-George, A.O., Meydan, C., Mason, C.E., and Jaffrey, S.R. (2015). Single-nucleotide-resolution mapping of m^6A and m^6Am throughout the transcriptome. Nat Methods 12, 767–772.
Liu, N., Parisien, M., Dai, Q., Zheng, G., He, C., and Pan, T. (2013). Probing N^6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA 19, 1848–1856.
Liu, S., Li, J.H., Wu, J., Zhou, K.R., Zhou, H., Yang, J.H., and Qu, L.H. (2015). StarScan: a web server for scanning small RNA targets from degradome sequencing data. Nucleic Acids Res 43, W480–W486.
Lorenz, R., Bernhart, S.H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol Biol 6, 26.
Loughrey, D., Watters, K.E., Settle, A.H., and Lucks, J.B. (2014). SHAPE-Seq 2.0: systematic optimization and extension of high-throughput chemical probing of RNA secondary structure with next generation sequencing. Nucleic Acids Res 42, e165.
Lovci, M.T., Ghanem, D., Marr, H., Arnold, J., Gee, S., Parra, M., Liang, T.Y., Stark, T.J., Gehman, L.T., Hoon, S., et al. (2013). Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. Nat Struct Mol Biol 20, 1434–1442.
Lovejoy, A.F., Riordan, D.P., and Brown, P.O. (2014). Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in *S. cerevisiae*. PLoS ONE 9, e110799.
Lu, Z., Zhang, Q.C., Lee, B., Flynn, R.A., Smith, M.A., Robinson, J.T., Davidovich, C., Gooding, A.R., Goodrich, K.J., Mattick, J.S., et al. (2016). RNA duplex map in living cells reveals higher-order transcriptome structure. Cell 165, 1267–1279.
Mahen, E.M., Watson, P.Y., Cottrell, J.W., and Fedor, M.J. (2010). mRNA secondary structures fold sequentially but exchange rapidly *in vivo*. PLoS Biol 8, e1000307.
Maragkakis, M., Alexiou, P., Nakaya, T., and Mourelatos, Z. (2016). CLIPSeqTools—a novel bioinformatics CLIP-seq analysis suite. RNA 22, 1–9.
Marchand, V., Blanloeil-Oillo, F., Helm, M., and Motorin, Y. (2016). Illumina-based RiboMethSeq approach for mapping of 2′-*O*-Me residues in RNA. Nucleic Acids Res 44, e135.
Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol 288, 911–940.
McMahon, A.C., Rahman, R., Jin, H., Shen, J.L., Fieldsend, A., Luo, W., and Rosbash, M. (2016). TRIBE: Hijacking an RNA-editing enzyme to identify cell-specific targets of RNA-binding proteins. Cell 165, 742–753.
Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. Nature 495, 333–338.
Meyer, K.D., and Jaffrey, S.R. (2014). The dynamic epitranscriptome: N^6-methyladenosine and gene expression control. Nat Rev Mol Cell Biol

15, 313–326.
Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., and Jaffrey, S.R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell 149, 1635–1646.
Montange, R.K., and Batey, R.T. (2006). Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. Nature 441, 1172–1175.
Mortimer, S.A., Kidwell, M.A., and Doudna, J.A. (2014). Insights into RNA structure and function from genome-wide studies. Nat Rev Genet 15, 469–479.
Mortimer, S.A., Trapnell, C., Aviran, S., Pachter, L., and Lucks, J.B. (2012). SHAPE-Seq: High-throughput RNA structure analysis. Curr Protoc Chem Biol 4, 275–297.
Motorin, Y., Muller, S., Behm-Ansmant, I., and Branlant, C. (2007). Identification of modified residues in RNAs by reverse transcription-based methods. Methods Enzymol 425, 21–53.
Nees, G., Kaufmann, A., and Bauer, S. (2014). Detection of RNA modifications by HPLC analysis and competitive ELISA. Methods Mol Biol 1169, 3–14.
Nguyen, T.C., Cao, X., Yu, P., Xiao, S., Lu, J., Biase, F.H., Sridhar, B., Huang, N., Zhang, K., and Zhong, S. (2016). Mapping RNA–RNA interactome and RNA structure *in vivo* by MARIO. Nat Commun 7, 12023.
Nicholson, C.O., Friedersdorf, M.B., Bisogno, L.S., and Keene, J.D. (2017). DO-RIP-seq to quantify RNA binding sites transcriptome-wide. Methods 118-119, 16–23.
Nilsen, T.W. (2014). Detecting RNA-RNA interactions using psoralen derivatives. Cold Spring Harbor Protocols 2014(9), pdb.prot080861.
Osborne, R.J., and Thornton, C.A. (2006). RNA-dominant diseases. Hum Mol Genet 15, R162–R169.
Peng, Z., Cheng, Y., Tan, B.C.M., Kang, L., Tian, Z., Zhu, Y., Zhang, W., Liang, Y., Hu, X., Tan, X., et al. (2012). Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. Nat Biotechnol 30, 253–260.
Piao, M., Sun, L., and Zhang, Q.C. (2017). RNA regulations and functions decoded by transcriptome-wide RNA structure probing. Genomics Proteomics Bioinformatics 15, 267–278.
Pyle, A. (2002). Metal ions in the structure and function of RNA. J Biol Inorg Chem 7, 679–690.
Qi, F., and Frishman, D. (2017). Melting temperature highlights functionally important RNA structure and sequence elements in yeast mRNA coding regions. Nucleic Acids Res 45, 6109–6118.
Qian, X., Zhao, J., Yeung, P.Y., Zhang, Q.C., and Kwok, C.K. (2019). Revealing lncRNA structures and interactions by sequencing-based approaches. Trends Biochem Sci 44, 33–52.
Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. Cell 108, 557–572.
Ramanathan, M., Porter, D.F., and Khavari, P.A. (2019). Methods to study RNA–protein interactions. Nat Methods 16, 225–234.
Ramani, V., Qiu, R., and Shendure, J. (2015). High-throughput determination of RNA structure by proximity ligation. Nat Biotechnol 33, 980–984.
Ramaswami, G., Lin, W., Piskol, R., Tan, M.H., Davis, C., and Li, J.B. (2012). Accurate identification of human Alu and non-Alu RNA editing sites. Nat Methods 9, 579–581.
Reuter, J.S., and Mathews, D.H. (2010). RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics 11, 129.
Roost, C., Lynch, S.R., Batista, P.J., Qu, K., Chang, H.Y., and Kool, E.T. (2015). Structure and thermodynamics of N^6-methyladenosine in RNA: A spring-loaded base modification. J Am Chem Soc 137, 2107–2115.
Roth, A., and Breaker, R.R. (2009). The structural and functional diversity of metabolite-binding riboswitches. Annu Rev Biochem 78, 305–334.
Roundtree, I.A., Evans, M.E., Pan, T., and He, C. (2017). Dynamic RNA modifications in gene expression regulation. Cell 169, 1187–1200.
Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., and Weissman, J.S. (2014). Genome-wide probing of RNA structure reveals active unfolding of mRNA structures *in vivo*. Nature 505, 701–705.
Safra, M., Sas-Chen, A., Nir, R., Winkler, R., Nachshon, A., Bar-Yaacov, D., Erlacher, M., Rossmanith, W., Stern-Ginossar, N., and Schwartz, S. (2017). The m^1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature 551, 251–255.
Sakurai, M., Ueda, H., Yano, T., Okada, S., Terajima, H., Mitsuyama, T., Toyoda, A., Fujiyama, A., Kawabata, H., and Suzuki, T. (2014). A biochemical landscape of A-to-I RNA editing in the human brain transcriptome. Genome Res 24, 522–534.
Schroeder, R., Barta, A., and Semrad, K. (2004). Strategies for RNA folding and assembly. Nat Rev Mol Cell Biol 5, 908–919.
Schwartz, S., Bernstein, D.A., Mumbach, M.R., Jovanovic, M., Herbst, R.H., León-Ricardo, B.X., Engreitz, J.M., Guttman, M., Satija, R., Lander, E.S., et al. (2014). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell 159, 148–162.
Scott, W.G., Finch, J.T., and Klug, A. (1995). The crystal structure of an AII-RNAhammerhead ribozyme: A proposed mechanism for RNA catalytic cleavage. Cell 81, 991–1002.
Seetin, M.G., Kladwang, W., Bida, J.P., and Das, R. (2014). Massively parallel RNA chemical mapping with a reduced bias MAP-Seq protocol. Methods Mol Biol 1086, 95–117.
Serganov, A., Polonskaia, A., Phan, A.T., Breaker, R.R., and Patel, D.J. (2006). Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. Nature 441, 1167–1171.
Shah, A., Qian, Y., Weyn-Vanhentenryck, S.M., and Zhang, C. (2016). CLIP Tool Kit (CTK): a flexible and robust pipeline to analyze CLIP sequencing data. Bioinformatics 33, btw653.
Sharma, E., Sterne-Weiler, T., O'Hanlon, D., and Blencowe, B.J. (2016). Global mapping of human RNA-RNA interactions. Mol Cell 62, 618–626.
Siegfried, N.A., Busan, S., Rice, G.M., Nelson, J.A.E., and Weeks, K.M. (2014). RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). Nat Methods 11, 959–965.
Silverman, I.M., Berkowitz, N.D., Gosai, S.J., and Gregory, B.D. (2016). Genome-wide approaches for RNA structure probing. Adv Exp Med Biol 907, 29–59.
Spitale, R.C., Flynn, R.A., Zhang, Q.C., Crisalli, P., Lee, B., Jung, J.W., Kuchelmeister, H.Y., Batista, P.J., Torre, E.A., Kool, E.T., et al. (2015). Structural imprints *in vivo* decode RNA regulatory mechanisms. Nature 519, 486–490.
Squires, J.E., Patel, H.R., Nousch, M., Sibbritt, T., Humphreys, D.T., Parker, B.J., Suter, C.M., and Preiss, T. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40, 5023–5033.
Sugimoto, Y., Chakrabarti, A.M., Luscombe, N.M., and Ule, J. (2017). Using hiCLIP to identify RNA duplexes that interact with a specific RNA-binding protein. Nat Protoc 12, 611–637.
Sugimoto, Y., Vigilante, A., Darbo, E., Zirra, A., Militti, C., D'Ambrogio, A., Luscombe, N.M., and Ule, J. (2015). hiCLIP reveals the *in vivo* atlas of mRNA secondary structures recognized by Staufen 1. Nature 519, 491–494.
Sy, B., Wong, J., Granneman, S., Tollervey, D., Gally, D., and Tree, J. (2018). High-resolution, high-throughput analysis of Hfq-binding sites using UV crosslinking and analysis of cDNA (CRAC). Methods Mol Biol 1737, 251–272.
Tafer, H., and Hofacker, I.L. (2008). RNAplex: a fast tool for RNA–RNA interaction search. Bioinformatics 24, 2657–2663.
Thody, J., Folkes, L., Medina-Calzada, Z., Xu, P., Dalmay, T., and Moulton, V. (2018). PAREsnip2: a tool for high-throughput prediction of small RNA targets from degradome sequencing data using configurable targeting rules. Nucleic Acids Res 46, 8730–8739.
Thore, S., Leibundgut, M., and Ban, N.N. (2006). Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. Science 312, 1208–1211.
Treiber, D.K., and Williamson, J.R. (1999). Exposing the kinetic traps in

RNA folding. Curr Opin Struct Biol 9, 339–345.
Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. Science 302, 1212–1215.
Ule, J., Jensen, K., Mele, A., and Darnell, R.B. (2005). CLIP: A method for identifying protein–RNA interaction sites in living cells. Methods 37, 376–386.
Underwood, J.G., Uzilov, A.V., Katzman, S., Onodera, C.S., Mainzer, J.E., Mathews, D.H., Lowe, T.M., Salama, S.R., and Haussler, D. (2010). FragSeq: transcriptome-wide RNA structure probing using high-throughput sequencing. Nat Methods 7, 995–1001.
Uren, P.J., Bahrami-Samani, E., Burns, S.C., Qiao, M., Karginov, F.V., Hodges, E., Hannon, G.J., Sanford, J.R., Penalva, L.O.F., and Smith, A. D. (2012). Site identification in high-throughput RNA–protein interaction data. Bioinformatics 28, 3013–3020.
Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Methods 13, 508–514.
Wan, Y., Kertesz, M., Spitale, R.C., Segal, E., and Chang, H.Y. (2011). Understanding the transcriptome through RNA structure. Nat Rev Genet 12, 641–655.
Wan, Y., Qu, K., Ouyang, Z., Kertesz, M., Li, J., Tibshirani, R., Makino, D. L., Nutter, R.C., Segal, E., and Chang, H.Y. (2012). Genome-wide measurement of RNA folding energies. Mol Cell 48, 169–181.
Wan, Y., Qu, K., Zhang, Q.C., Flynn, R.A., Manor, O., Ouyang, Z., Zhang, J., Spitale, R.C., Snyder, M.P., Segal, E., et al. (2014). Landscape and variation of RNA secondary structure across the human transcriptome. Nature 505, 706–709.
Wenzel, A., Akbaşli, E., and Gorodkin, J. (2012). RIsearch: fast RNA–RNA interaction search using a simplified nearest-neighbor energy model. Bioinformatics 28, 2738–2746.
Weyn-Vanhentenryck, S.M., Mele, A., Yan, Q., Sun, S., Farny, N., Zhang, Z., Xue, C., Herre, M., Silver, P.A., Zhang, M.Q., et al. (2014). HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. Cell Rep 6, 1139–1152.
Will, C.L., and Luhrmann, R. (2011). Spliceosome structure and function. Cold Spring Harb Perspect Biol 3, a003707.
Williamson, J.R. (2000). Induced fit in RNA-protein recognition. Nat Struct Biol 7, 834–837.
Woodson, S.A. (2005). Metal ions and RNA folding: a highly charged topic with a dynamic future. Curr Opin Chem Biol 9, 104–109.
Wulff, B.E., Sakurai, M., and Nishikura, K. (2011). Elucidating the inosinome: global approaches to adenosine-to-inosine RNA editing. Nat Rev Genet 12, 81–85.
Xia, T., SantaLucia Jr., J., Burkard, M.E., Kierzek, R., Schroeder, S.J., Jiao, X., Cox, C., and Turner, D.H. (1998). Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson−Crick base pairs. Biochemistry 37, 14719–14735.
Yang, S.Y., Lejault, P., Chevrier, S., Boidot, R., Robertson, A.G., Wong, J. M.Y., and Monchaud, D. (2018). Transcriptome-wide identification of transient RNA G-quadruplexes in human cells. Nat Commun 9, 4730.
Yang, X., Yang, Y., Sun, B.F., Chen, Y.S., Xu, J.W., Lai, W.Y., Li, A., Wang, X., Bhattarai, D.P., Xiao, W., et al. (2017). 5-methylcytosine promotes mRNA export—NSUN2 as the methyltransferase and ALYREF as an m^5C reader. Cell Res 27, 606–625.
Yates, L.A., Norbury, C.J., and Gilbert, R.J.C. (2013). The long and short of microRNA. Cell 153, 516–519.
Zarnegar, B.J., Flynn, R.A., Shen, Y., Do, B.T., Chang, H.Y., and Khavari, P.A. (2016). irCLIP platform for efficient characterization of protein–RNA interactions. Nat Methods 13, 489–492.
Zhao, B.S., Roundtree, I.A., and He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. Nat Rev Mol Cell Biol 18, 31–42.
Zhao, Y., Zhang, Y., Teng, Y., Liu, K., Liu, Y., Li, W., and Wu, L. (2019). SpyCLIP: an easy-to-use and high-throughput compatible CLIP platform for the characterization of protein–RNA interactions with high accuracy. Nucleic Acids Res 47, e33.
Zheng, Q., Ryvkin, P., Li, F., Dragomir, I., Valladares, O., Yang, J., Cao, K., Wang, L.S., and Gregory, B.D. (2010). Genome-wide double-stranded RNA sequencing reveals the functional significance of base-paired RNAs in *Arabidopsis*. Plos Genet 6, e1001141.
Zheng, Y., Li, Y.F., Sunkar, R., and Zhang, W. (2011). SeqTar: an effective method for identifying microRNA guided cleavage sites from degradome of polyadenylated transcripts in plants. Nucleic Acids Res 40, e28.
Zhou, Y., Kierzek, E., Loo, Z.P., Antonio, M., Yau, Y.H., Chuah, Y.W., Geifman-Shochat, S., Kierzek, R., and Chen, G. (2013). Recognition of RNA duplexes by chemically modified triplex-forming oligonucleotides. Nucleic Acids Res 41, 6664–6673.
Ziv, O., Gabryelska, M.M., Lun, A.T.L., Gebert, L.F.R., Sheu-Gruttadauria, J., Meredith, L.W., Liu, Z.Y., Kwok, C.K., Qin, C.F., MacRae, I.J., et al. (2018). COMRADES determines in vivo RNA structures and interactions. Nat Methods 15, 785–788.
Zubradt, M., Gupta, P., Persad, S., Lambowitz, A.M., Weissman, J.S., and Rouskin, S. (2017). DMS-MaPseq for genome-wide or targeted RNA structure probing *in vivo*. Nat Methods 14, 75–82.
Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31, 3406–3415.