SCIENCE CHINA
Life Sciences

• REVIEW •

January 2015 Vol.58 No.1: 75–88
doi: 10.1007/s11427-014-4764-5

CLIP: viewing the RNA world from an RNA-protein interactome perspective

ZHANG Yin, XIE ShuJuan, XU Hui & QU LiangHu*

Key Laboratory of Gene Engineering of Ministry of Education, State Key Laboratory for Biocontrol, Sun Yat-Sen University, Guangzhou 510275, China

Received July 14, 2014; accepted August 13, 2014

The pervasive transcription of the genome creates many types of non-coding RNAs (ncRNAs). However, we know very little regarding the functions and the regulatory mechanisms of these ncRNAs. Exploring the interactions of RNA and RNA binding proteins (RBPs) is vital because it can allow us to truly understand how these ncRNAs behave *in vivo*. High-throughput sequencing of RNA isolated by cross-linking immunoprecipitation (HITS-CLIP or CLIP-seq) and its variants have been successfully used as systemic techniques to study RBP binding sites. In this review, we will explain the major differences between the CLIP techniques, summarize successful applications of these techniques, discuss limitations of CLIP, present some suggested solutions and project their promising future roles in studying the RNA world.

CLIP, CLASH, RPBs, ncRNAs, functional RNomics

Citation: Zhang Y, Xie SJ, Xu H, Qu LH. CLIP: viewing the RNA world from an RNA-protein interactome perspective. Sci China Life Sci, 2015, 58: 75−88, doi: 10.1007/s11427-014-4764-5

The ENCODE (Encyclopedia of DNA Elements) project indicated that approximately 75% of the genome are covered by primary transcripts and 62% by processed transcripts [1], while only 2% encode for proteins [2]. This vast disparity indicates that there are many types of non-coding RNAs (ncRNAs). Recently, some small ncRNAs, such as microRNA, piRNA and siRNA, have attracted attention and some progress has occurred. The function and the regulatory mechanisms of these ncRNAs have been illustrated. However, for some types of ncRNA, such as long non-coding RNAs (lncRNAs), their roles and functional mechanisms are poorly understood. Moreover, due to a lack of similar mechanisms of action and sequence conservation, it is very challenging to fully master these ncRNAs. Investigations of these and other newly identified ncRNAs will be important work in the next decade.

The biogenesis and the maturation of RNAs are regulated by many proteins. Moreover, RNA and RNA binding proteins (RBPs) usually form RNA-protein complexes to function together. Some types of RNAs, for example, piRNA (Piwi-interacting RNA), are named after the corresponding binding protein. Therefore, investigating the physiological functions of RBPs and their interactions with RNAs will be of help in understanding the RNA world. The dynamic combination, competition and collaboration between RNA and RBPs offers many opportunities to explore the functions and mechanisms of action of these RNAs [3]. Many biochemical technologies are invented to explore interactions between RNAs and RBPs. Recently, a method known as cross-linking immunoprecipitation (CLIP), and modified forms of this method, such as photoactivatable-ribonucleoside-enhanced cross-linking and immunoprecipitation (PAR-CLIP), individual-nucleotide resolution cross-linking and immunoprecipitation (iCLIP), cross-linking and analysis of cDNAs (CRAC) and cross-linking, ligation, and sequencing of hybrids (CLASH), have been successfully used

*Corresponding author (email: lssqlh@mail.sysu.edu.cn)

in decoding RNA-protein interactions. In this review, we describe the procedures, discuss the differences between each protocol, illustrate some classical cases and discuss possible improvements and applications of these methods.

1 The development and modifications of CLIP

Many biochemical technologies are invented to study the interactions between RNAs and RBPs. The most widely used strategy is purification of RNA-protein complexes through different methods by pulling down either RNAs or proteins. RNA immunoprecipitation (RIP) is extensively employed to identify the interactions between RNAs and proteins, and many commercial kits are available. However, this method also has many disadvantages; for example, producing high false positive products and an inability to pinpoint the sites of RNA-protein interactions [4,5].

To overcome the intrinsic disadvantages of RIP, Darnell and co-workers [6,7] developed a complex protocol based on RIP, named UV cross-linking and immunoprecipitation (CLIP). Then, they combined emerging high-throughput sequencing with CLIP and named the technique HITS-CLIP [8,9]. They improved the RIP protocol in three ways. First, the samples, either tissues or cells, are cross-linked by UV irradiation. UV cross-linking produces covalent bonds between RNA and protein, which allows stringent purification and increasing the signal-to-noise ratio in the following steps. Compared to formaldehyde cross-linking, which is usually used in chromatin immunoprecipitation (ChIP), the UV cross-linking does not induce large multi-molecular chemical bridges and only cross-links direct nucleic acid-protein interactions (within Ångstrom distances) [10]. Second, RNase digestion was introduced to the protocol. Because free RNAs are much more sensitive to RNase than protein-bound RNAs, moderate RNase digestion can digest the unbound parts of RNA and reserve the protein binding sites of RNA. The fragmentation of RNAs makes it convenient to perform sequencing and motif analysis, in addition to avoiding pulling down undesirable protein-RNA-protein complexes. Third, rigorous purification, such as multiple washings between each reaction, SDS-PAGE and transfer to nitrocellulose membrane were added. These steps can help to remove the non-specific binding proteins for the antibody, and eliminate unbound and non-cross-linked RNAs, which will decrease the background of the data.

After successfully using HITS-CLIP to decode the RNA-protein interactions of Nova and Argonaute [6,8,9], many modified versions of CLIP have been developed. Among them, PAR-CLIP is the most widely used. The PAR-CLIP protocol employs ribonucleoside analogs, such as 4-thiouridine (4SU) and 6-thioguanosine (6SG), which can be selectively photoactivated at long wavelength UV (>320 nm), to enhance cross-linking between nucleic acids and proteins [11]. 4SU is the preferred option because it has higher cross-linking efficiency [12] and uridines are the most frequent cross-linking sites [13]. The incorporation of 4SU into RNAs results in thymidine (T) to cytidine (C) transitions during the RT reaction, which is a good indicator of binding sites and can be used to subtract non-cross-linking background signals [14]. The addition of 4SU to PAR-CLIP also results in some limitations. To incorporate 4SU into nascent RNAs, it must be added to the medium, which is not possible for most model animals or clinical specimens. Moreover, although the mRNA profile is almost unchanged after adding 4SU [12], there are some reports showing that 4SU is toxic to cells [15–17].

Another variant is individual-nucleotide resolution UV cross-linking and immunoprecipitation (iCLIP). This protocol was designed to acquire information of cross-linking sites at a single nucleotide resolution [18]. The amino acid residues at the cross-linking sites are barriers to the RT reaction, which often results in mutation and deletion, and may even stop the RT reaction [17,19]. These mutation and stop sites can provide information regarding the cross-linking sites [5]. However, because amplification of cDNAs depends on 5′ and 3′ adaptor sequences, cDNAs lacking 5′ adaptors, which are caused by incomplete RT, cannot be amplified by classical cloning strategies. The iCLIP protocol introduces a 5′ linker by circularization of cDNA and restriction enzymes, which bypasses the limitation of a lack of a 5′ linker due to incomplete RT (Figure 2).

The CRAC and CLASH methods employ affinity-purification of tagged proteins rather than immunoprecipitation of endogenous proteins [20–22]. CLASH is an improved version of CRAC. The most important improvement is ligation of base-paired RNA molecules and the generation of chimeric RNAs. RNA guiding is a commonly used mechanism for many types of RNAs, such as miRNA, piRNA and snoRNA [23], and the search for targets of these RNAs is a challenge [24]. The chimeric RNAs generated by CLASH can detect direct, high-throughput mapping of RNA-RNA interactions and thus provide an excellent solution for target searching.

The major differences, advantages and disadvantages of these different versions of CLIP are summarized in Table 1.

2 Key steps of the CLIP protocol

Because the detailed protocols of each version of CLIP are illustrated clearly in the original studies, methodological special issues and reviews (Table 1). Here, we do not include protocol details but instead focus on the different strategies adopted by different variants of CLIP in each step. The bioinformatic analysis of CLIP data is also described here, and the available tools are summarized in Table 2.

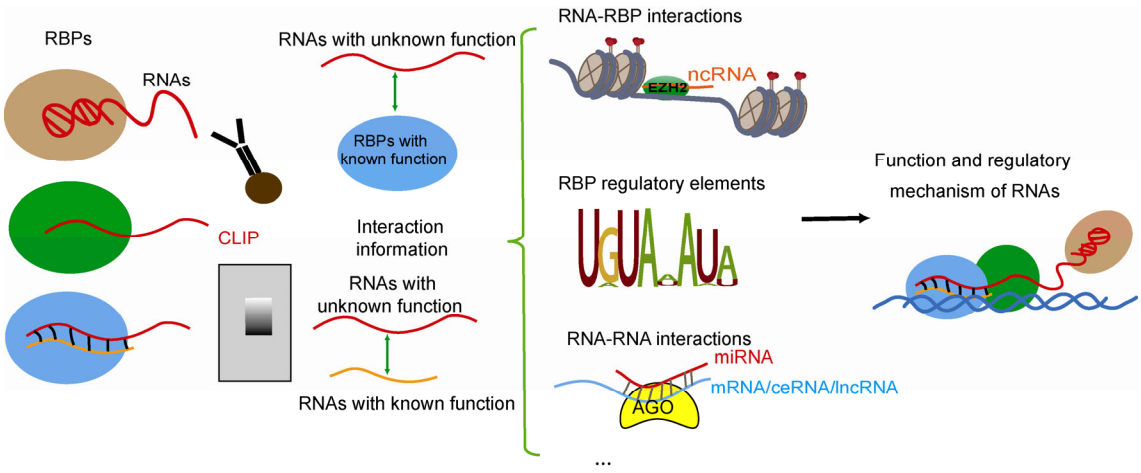

Figure 1 CLIP as an excellent tool for functional RNomics study. CLIP provides useful information of RNA-RBP interactions, RPB regulatory elements and RNA-RNA interactions. Integrated analysis of this information can help researchers to understand the function and regulatory mechanism of RNAs, especially novel ncRNAs.

Table 1 Differences among CLIP variants

Protocol name	Major modification	Advantages and disadvantages	References
HITS-CLIP	Original version of CLIP	Can apply to both tissue and cell; no potential toxicity to cell compared to PAR-CLIP Low cross-linking efficiency; lack of indicator of binding sites	[6,9]
PAR-CLIP	Use 4SU to enhance cross-link efficiency	Higher cross-linking efficiency; T to C mutation as a indicator of binding sites Can only apply to cell line; potential toxicity to cell	[12,13]
iCLIP	Use a new strategy to clone incomplete RT product due to cross-linking	More efficient RT-PCR; RT blocked sites can be indicator of binding sites	[17,18,25]
CRAC	Use affinity-purification to isolate RNA-binding protein	Have no restriction of antibody; can be applied for non-model organism The tag and overexpression of protein may have potential influence on the binding pattern between protein and RNA	[20]
CLASH	Use affinity-purification to isolate RNA-binding protein; conduct intermolecular RNA-RNA ligation	The formation of chimeric RNAs is a very useful imformation for RNA that function as a guider such as miRNA and snoRNA The tag and overexpression of protein may have potential influence on the binding pattern between protein and RNA	[21,22]

2.1 Sample preparation and UV cross-linking

The samples used for UV cross-linking must be fresh, and cryopreservation can only be performed after UV cross-linking. For PAR-CLIP, a final concentration of 100 μmol L^{-1} of 4SU should be added into the medium 14 h before cross-linking, and then 365 nm UV was selected to cross-link RNA-protein complex. Except for PAR-CLIP, 254 nm UV is used to cross-link the RNA-protein complexes. Tissues should be homogenized or triturated before the UV cross-linking. The strength of the UV cross-link should be optimized; tissues usually require higher energy levels than cells. The UV cross-linked sample can be either processed to the next step or shock-frozen in liquid nitrogen and stored at −80°C for up to one year [9,26].

2.2 Cell lysis and RNase digestion

Newly cross-linked or frozen samples are lysed in cooled lysis buffer for approximately 30 min. The composition of the lysis buffer may differ among the methods. The key components are Tris-HCl or PBS buffer, detergent and protease inhibitors. After cell lysis, RNase is added to the lysis buffer. For HITS-CLIP and iCLIP, DNase is also added before RNase digestion. However, for PAR-CLIP, CRAC and CLASH, the DNase digestion is omitted. different RNase such as RNase A [9], RNase T1 [12], and RNase I [27] are used in different laboratories. RNase A specifically cleaves single-stranded RNA at C and U residues and RNase T1 at G residues; RNase I has no biases. To minimize biases caused by RNase, a mixture of RNase A and T1 is used in CRAC and CLASH, and the RNA digestion is performed on beads rather than in lysis buffer [20–22]. The concentration of RNase and the time of digestion should be determined to obtain RNA fragments of the appropriate length.

2.3 Immunoprecipitation

For HITS-CLIP, PAR-CLIP and iCLIP, immunoprecipita-

Table 2 Available software or database for CLIP

Name	Implementation	Application	Features	References
PARalyzer	C/C++	PAR-CLIP	Utilizes kernel density estimations to distinguish T to C mutations signal and background and thus identify binding sites	[14]
wavClusteR	R package	PAR-CLIP	Use nonparametric mixture models to identify binding sites	[45]
rcpphmmparclip	R package	PAR-CLIP	Adopts non-homogeneous HMM and considers spatial dependence of genomic locations	[46]
CIMS	Perl/shell	all CLIP variants	Deletion and substitution are tracked to identify cross-linking sites	[4,17]
dCLIP	Perl	all CLIP variants	Quantitative CLIP-seq comparative analysis	[47]
Piranha	C/C++	all CLIP variants	Peak-calling tool that can allows consideration of transcript abundance, direct comparison of site usage across cell-types or conditions	[32]
ASPeak	Perl	all CLIP variants	Peak-calling tool that can used in CLIP and RIP	[48]
RIPSeeker	R package	all CLIP variants	*De novo* peak predictions based on HMM	[34]
CLIPper	Python	all CLIP variants	https://github.com/YeoLab/clipper/wiki/CLIPper-Home A tool for defining peaks of CLIP data	
CapR	C/C++	all CLIP variants	Prediction secondary structure of RBPs binding motif	[39]
GraphProt	Perl	all CLIP variants	Provides motif search, structure analysis and diverse funtions like predicting RBP affinities and scanning for RBP target sites	[40]
PIPE-CLIP	Galaxy server	all CLIP variants	An online pipeline to process and analyze CLIP-seq data which allows user-specified parameters	[49]
CLIPZ	Web server	all CLIP variants	Provides database and analysis environment for analyzing and visualizing CLIP-seq data	[41]
doRiNA	Web server	all CLIP variants	Database stores and integrates CLIP data for RBPs and miRNAs, visualized by local copy of the UCSC genome browser	[42]
starBase	Web server	all CLIP variants	Database provides RBPs binding sites, integrated ncRNAs analysis and Pan-Cancer analysis based on comprehensive CLIP data, which allows visualizing , downloading and plotting of these data	[43,44]

tion is used to purify the RNA-protein complex, so the choice of antibody is vital. Because many antibodies used for Western blotting perform poorly in immunoprecipitation, and the specificity of antibodies varies under different conditions, pre-testing the antibodies under the same conditions of CLIP is strongly recommended. Immunoprecipitation of endogenous or tagged protein and silver staining can be used to detect the antibodies. The antibodies are pre-incubated with Dynabead-coupled protein A or G to immobilize them. For CRAC and CLASH, the protein is fused to a His6-TEV-Protein A tag; the protein is first purified by IgG coupled on beads and then by Ni-NTA affinity chromatography.

2.4 Radiolabeling of RNA, SDS-PAGE and transfer to nitrocellulose

To visualize the band of RNA-protein complex in gel or nitrocellulose membrane, RNAs or linkers are radiolabeled. For HITS-CLIP and iCLIP, 3′ linkers are radiolabeled and then ligated to RNAs on beads. As for PAR-CLIP, CRAC and CLASH, radiolabeling is carried on RNAs in the RNA-protein complex. After radiolabeling, the samples are denatured and separated by SDS-PAGE. The RNA-protein complexes are then transferred to nitrocellulose membranes, exposed on film to visualize the complexes and cut the corresponding band. The transfer nitrocellulose membrane permeates free RNAs and detains RNA-protein complexes, which can reduce background. However, for PAR-CLIP, the transfer nitrocellulose is omitted; the gel is exposed on film, cut, and then electro-eluted.

2.5 Isolation and cloning of RNA

The nitrocellulose membranes or electroelution products are incubated with proteinase K to digest protein and release the RNAs. The RNAs are then recovered by phenol/chloroform extraction and precipitated by ethanol and glycogen. Re-dissolved RNAs are linked to adaptors and then reverse-transcribed. For this process, different CLIP methods take different strategies. For HITS-CLIP, the 3′ linker is added prior to SDS-PAGE on beads and the 5′ linker is added to the RNA. For iCLIP, the 3′ linker is already ligated to the beads and the RT primer contains two cleavable linker regions. After RT and size selection by urea-PAGE, circularization and restriction enzyme digestion are carried out. The two cleavable linker regions separate and became the 5′ linker and the 3′ linker. This ingenious cloning strategy bypasses the blockages during the RT reaction. Moreover, the blocking site of RT can be a good indicator for cross-linking. For PAR-CLIP, CRAC and CLASH, the isolated RNA is added to the 3′ linker and the 5′ linker and is reverse-transcribed. Many measures, such as blockage of the 3′ linker, sequential dephosphorylation and phosphorylation, employment of truncated ligases, were adopted to avoid self-ligation and other undesired ligations in a different CLIP protocol. The cDNA is then amplified by PCR of proper cycles and then sent to undergo high-throughput se-

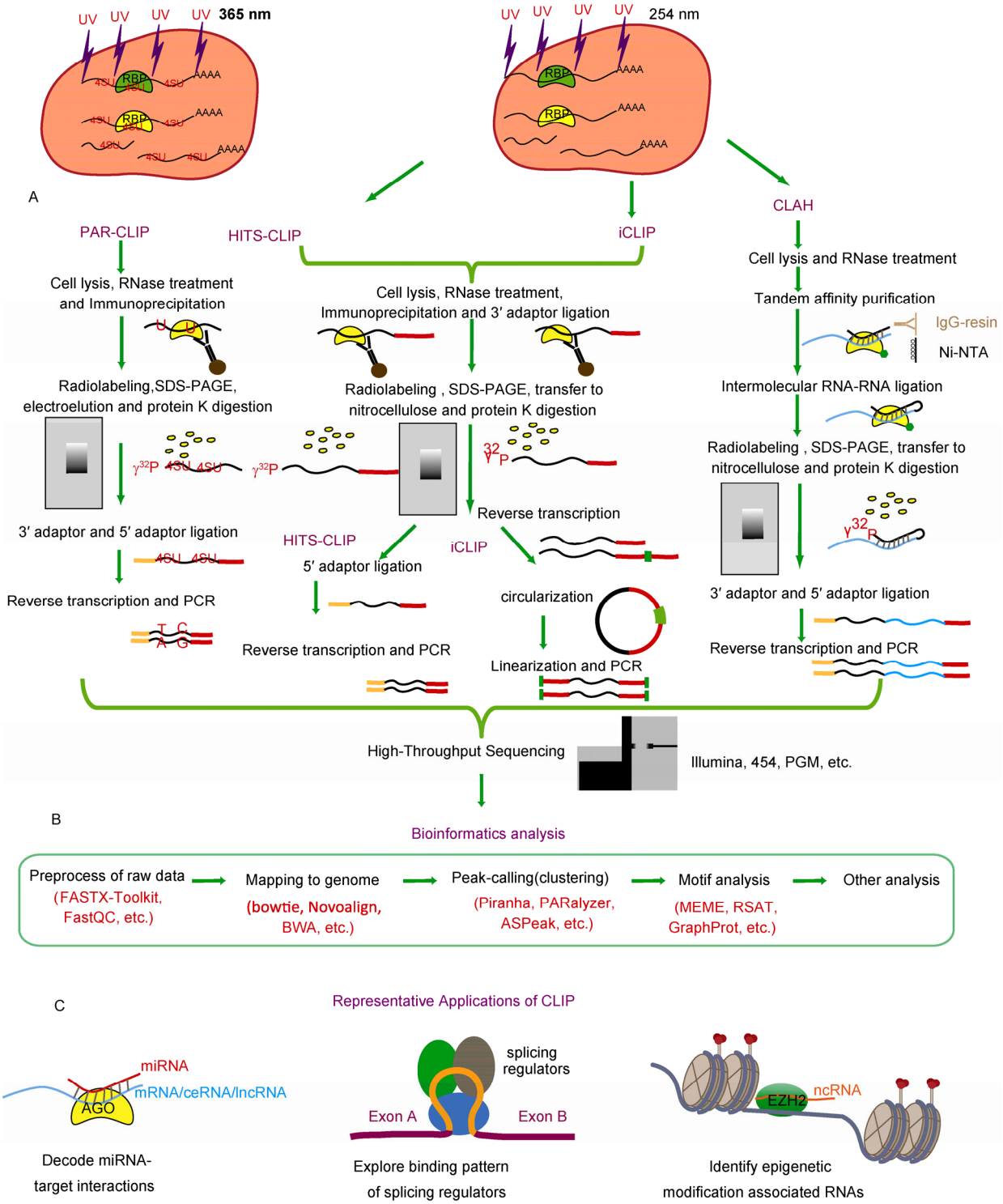

Figure 2 The CLIP methods and their applications. A, The schematic illustrates the workflow of CLIP and differences of each method. B, Key steps of bioinformatic analysis of CLIP and the available tools. C, Examples of successful applications of CLIP.

quencing.

2.6 Bioinformatic analysis of CLIP data

High-throughput sequencing using the CLIP procedure generates raw data that should be processed before further analysis. Preprocessing of reads includes barcode splitting, quality filtering, and adaptor removal. Several tools are available to clean up raw sequencing data. FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)

provides the overall quality status of raw data, and the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) and PRINSEQ [28] are collections of command line tools for multiple preprocessing.

Preprocessed data in FASTQ or FASTA format is then mapped to the genome or transcriptome. Because of insertions and deletions caused by cross-linking, the software used for mapping should allow gaps. Algorithms such as BLAST+ [29], Novoalign (http://www.novocraft.com/main/index.php), BWA [30] and Bowtie 2 [31] are capable of aligning gapped CLIP tags to the genome or transcriptome and report the position of mutations. For PAR-CLIP, more mismatches should be tolerated due to inherent T to C mutations caused by 4SU.

Overlapped tags are fused and the read coverage of each position is calculated. To distinguish between true binding sites and background caused by non-specific sequences, peak calling is required. There are some computer programs designed for peak-calling of ChIP-seq, but they are not suitable for CLIP due to intrinsic differences between ChIP-seq and CLIP. Programs such as Piranha [32], Pyicos [33], and RIPSeeker [34] were developed to solve this problem. PARalyzer [14] takes advantage of T to C mutations to identify cross-linking sites of PAR-CLIP data. To find features of binding sites, motif searching is usually conducted. Algorithms such as MEME [35], GMS [36], AlignACE [37] and RSAT [38] can be used to identify motifs within CLIP data. Recently, tools specifically designed for motif and structure analysis of CLIP have been developed [39,40].

There are also some databases that allow for collecting and visualizing CLIP data, such as CLIPZ [41], doRiNA [42] and starBase [43,44]. StarBase v 2.0 is the most comprehensive CLIP database, which integrates 108 CLIP-Seq (HITS-CLIP, PAR-CLIP, iCLIP, CLASH) datasets of different RBPs. In addition to RNA-protein interactions, miRNA-lncRNA, miRNA-mRNA, miRNA-circRNA, miRNA-pseudogene, miRNA-sncRNA interactions, competing endogenous RNA (ceRNA) networks, miRNA function and ceRNA function based on CLIP data are also available. In the recently updated version, it also provides Pan-Cancer Networks of lncRNAs, miRNAs, ceRNAs and RNA-binding RBPs by mining clinical and expression profiles of 14 cancer types (>6000 samples) from The Cancer Genome Atlas (TCGA) Data Portal.

For convenient analysis of CLIP data, we summarize the available tools and databases especially developed for CLIP, as well as their implementation, applications and main features in Table 2.

3 Successful applications of CLIP technology

Since its development, over 300 studies using CLIP have been published [4]. Collecting CLIP data for each RBP could provide a genome-wide RNA-protein interaction map and thus largely promote studies in the field.

3.1 Decoding miRNA-target interactions

The most successful and extensive applications of CLIP are identifying miRNA targets. miRNAs are approximately 22 nt non-coding RNAs that play important roles in regulating gene expression, primarily at the post-transcriptional level [50,51]. Although many studies have explored their functions in proliferation, differentiation, and development, as well as pathologies such as tumorigenesis, the identification of miRNA targets remains a challenge [24,50]. Many algorithms, mainly based on complementarity of "seed regions" and conservation of 3′ UTRs, have been developed to predict miRNA targets [52–54]. High false-positive rates and different performances of each program [51,55,56] are major problems for bioinformatic prediction of miRNA-target interactions. High-throughput and experimental identification of physiological targets and their binding sites is urgently needed.

The first report of this type of work is from 2009 [9], describing the analysis of an Ago2-miRNA-mRNA ternary complex in mouse brain using HITS-CLIP. Combined with bioinformatic analysis, thousands of miRNA-targets were identified. The specificity, false-positive and false-negative rates were approximately 93%, 13%–27% and 15%–25%, respectively, estimated by comparing previously identified targets of miR-124.

Markus et al. [12] studied all four human Ago proteins by PAR-CLIP in HEK293 cells that stably express tagged Ago proteins. They obtained approximately 4000 overlapped clusters of each Ago protein, and found that human Ago1–4 have similar binding patterns. The mRNA abundance of identified targets was statistically significant, decreasing after overexpression of miRNAs, which agreed with the observation that miRNAs cause destabilization of targets.

The existing algorithms for miRNA target prediction are mostly based on a perfect match at a "seed region". However, a previous structural study of a ternary complex of *T. thermophilus* Ago indicated that the mismatches and bulges in the seed region could be tolerated [57]. Another study suggested that these non-canonical sites were quite rare (less than 2% of all preferentially conserved sites) [58]. Ago CLIP makes it possible for the first time to systematically identify these non-canonical targets of miRNA. By analyzing Ago HITS-CLIP "orphan clusters," Sung et al. [59] identified a type of non-canonical seed match: G-bulge sites, which comprise ⩾15% of all Ago-miRNA interactions in mouse brain. By carrying out HITS-CLIP for miR-155 in wild type and knockout T cells, Gabriel et al. [60] systemically identified non-canonical targets of miR-155. For CLASH, the chimeric RNAs provide direct evidence of RNA-RNA interaction. This information is extremely valuable for studying miRNA targets, especially non-canonical

targets. Using the CLASH method, Aleksandra et al. reported over 18000 miRNA-target interactions, and found that approximately 60% of seed interactions are non-canonical [22]. Today, there are several tools identifying non-canonical targets that take advantage of CLIP data [61–63]. Decoding these non-canonical targets will expand our knowledge of miRNA, especially its target recognition mechanisms.

The properties of miRNA make it an ideal tool for viruses to regulate the gene expression of host cells. More than 200 viral miRNAs have been identified. These viral miRNAs can target their own mRNAs, as well as host cell mRNAs. Cellular miRNAs can also target viral mRNAs [64,65]. Moreover, some viral miRNAs are orthologs of cellular miRNAs. For example, KSHV miR-K12-11 is an ortholog of miR-155, and they share the same seed sequence [66]. The miRNA-mRNA relationships between viruses and host cells, and their role in the regulation of infection and latent/lytic switching are quite intriguing, and can serve as an excellent model to investigate host-virus coevolution. To uncover the roles of miRNAs in host-virus interactions, the identification of bona fide targets is vital. Several reports employing CLIP to comprehensively understand these questions have been published, including KSHV [67,68] and EBV [69,70]. In each study, 500–2000 targets of viral miRNAs were identified. These reports increase our knowledge of the function of viral miRNA.

Embryonic stem cell (ESCs)-specific miRNAs play critical roles in establishing the stemness of ESCs [71]. The miRNAs that are specifically or highly expressed in ESCs can also reprogram mouse and human somatic cells to iPSCs alone or with the help of transcription factors [72–74]. Various reports indicated that miRNAs are a key regulator of cell fate decision and conversion [75,76]. Although some important targets of these ESCs-enriched miRNAs have been revealed, many others remain to be identified. Leung et al. [77] conducted Ago2 CLIP in wild type mouse ESCs (mESCs) and *Dicer*$^{-/-}$ mESCs. They reported that the GCACUU motif, corresponding to the most highly expressed miRNA family, was significantly enriched in wild-type mESCs 3′ UTRs, but not in *Dicer*$^{-/-}$ mESCs. Many validated and novel targets of mESC-enriched miRNAs were identified. Human ESCs (hESCs) were also investigated by PAR-CLIP [78]; hundreds of miR-302/367 targets were identified. Combining the dynamic change in miRNA profile at different stages of differentiation in hESCs, the roles of miR-302/367 in modulating TGFβ and the BMP signaling pathway were revealed.

Inquiries into disease-related miRNA-mRNA interactions can help us to understand pathogenesis and provide potential therapeutic targets for the disease. The miR-17–92 family is the first miRNA gene implicated in cancer [79]. By combining PAR-CLIP analysis of human B cells and experimental validation in B cell-specific miR-17–92 transgenic mice that developed lymphomas, Hyun et al. [80] reported that miR-17–92 acts as a powerful cancer driver via activation of multiple oncogenic pathways, such as PI3K and NFκB. HITS-CLIP performed on human brain samples revealed that many disease-associated SNPs are within the target sites of brain-specific miRNAs, which indicated that these SNPs could result in disease by destroying miRNA-target interactions [81]. The imprinted DLK1-MEG3 gene region is associated with type 2 diabetes (T2DM). Using HITS-CLIP, Vasumathi et al. [82] identified disease-relevant targets of miRNAs in the DLK1-MEG3 gene region, implying their roles in T2DM.

3.2 Identifying RNA-binding sites of splicing regulators

RNA splicing is one of the most important post-transcriptional regulation mechanisms, and plays a critical role in diverse life activities. Alternative splicing makes a large contribution to the polymorphism of mature mRNAs and proteins. Dozens of RNA binding proteins are involved in the regulation of RNA splicing [83,84]. Apart from Ago proteins, splicing regulators are the most intensively studied proteins by CLIP. RNA binding patterns of Nova [6,8], PTBs [85,86], FOX2 [87], SFRS1 [88], TIA [89], MBNL1 [90], Cugbp1 [91], SRSF1/2 [92] and hnRNP proteins [18,93,94] have been investigated.

In addition to exploring the interplay of RNA-protein and decoding their roles in splicing, there are also some unexpected discoveries made or inspired by CLIP analysis of these splicing regulators. For example, hnRNP A1 is a nucleocytoplasmic shuttling RBP that functions in alternative splicing and other mRNA metabolism process. Although the regular roles of hnRNP A1 in mRNA processing are well studied [95–97], only a few targets of hnRNP A1 have been identified. Using CLIP, Sonia et al. [98] obtained a genome-wide map of hnRNP A1 binding sites. Unexpectedly, they found that hnRNP A1 specifically binds the primary transcript of miR-18a at the stem loop region, rather than other miRNAs. Additional study revealed that hnRNP A1 facilitates processing of miR-18a in as well as repression of targets. This unexpected finding indicated that some RBPs could function as co-regulators for a specific miRNA at the post-transcriptional level, which is a more flexible strategy.

Another interesting finding revealed by CLIP data was reported by Fu and colleagues, who revealed that knockdown of PTB can induce trans-differentiation of fibroblasts into functional neurons [99]. PTB (also known as hnRNP I), a well-characterized splicing suppressor, is downregulated during neural development, which coincides with an increased expression of its NS-enriched homolog (hPTB). Based on CLIP data and previous results [85,100], they identified a PTB-regulated miR124-REST loop, in which the downregulation of PTB leads to relief of PTB-mediated blockage of miR-124 repression of the REST complex, and

thus promote many neuronal genes, finally inducing trans-differentiation into neuron cells.

3.3 Studying epigenetic modification-associated RNAs

Polycomb Repressive Complex 2 (PRC2) can induce the methylation of histone H3 on lysine 27 (H3K27 me3) and result in transcriptional gene silencing, which controls cell fate, development and tumorigenesis [101]. How PRC2s are recruited to the target region of the genome is unknown. Some studies showed that non-coding RNAs guide PRC2 to specific regions via interacting with SUZ12 or EZH2 [102,103]. To systemically study the associated RNAs of PRC2, Guil et al. [104] and Syuzo et al. [105] conducted CLIP experiments on EZH2 in human colorectal cancer cells and mouse ES cells. In human colorectal cancer cells, EZH2 binds many intronic RNAs. However, in mouse ES cells, EZH2 preferentially binds the 5′ region of nascent RNAs. These results indicated that EZH2 might incorporate different types of RNAs in different cell types.

Increasing evidence suggests that non-coding RNAs, including small RNAs and long non-coding RNA, play important roles in epigenetic regulation [106,107]. Proteins that can bind both DNA and RNA are ideal candidates for transcriptional and/or epigenetic regulators, in an RNA-guiding manner. Like the multifunction protein EZH2, there are some well-studied DNA-binding proteins, such as DNA methylase [108], Smad [109], p53 [110,111], and others [112], that bind RNAs. Similarly, some RBPs can also bind DNA [113–115]. Further investigation of their DNA or RNA binding characteristics, meta-analysis ChIP and CLIP data for these proteins will become research hotspots in the future.

CTCF is a well-known DNA binding protein which plays extremely important roles in higher-order chromatin organization and has great impacts on gene expression at genome-wide [116]. Recently, by using PAR-CLIP, Ricardo et al. [117] reported CTCF can bind a variety of RNAs *in vivo*. Moreover, they found CTCF can regulate p53 expression via binding Wrap53 RNA, which is an antisense transcript originating from the p53 locus. This discovery not only revealed new regulatory mechanism of CTCF, but also provided a vivid example that DNA binding proteins also function as RBPs.

3.4 Exploring roles of ceRNA

The ceRNA hypothesis [118,119] highlights interactions among mRNAs, transcribed pseudogenes and long non-coding RNAs by competitive binding to microRNA responding elements (MREs). More and more evidence supports this hypothesis, which suggests that these ceRNAs play essential roles in development and tumorigenesis [120–126]. The identification of ceRNAs by traditional miRNA-target prediction algorithms is challenging due to an incomplete understanding of targeting rules [127].

CLIP technology provides a good opportunity to study ceRNAs. A representative example was a type of circular RNA reported by Memczak et al. [128], which contains 63 conserved sites of miR-7 that act as a "super sponge" for miRNA. This exciting discovery revealed "an important RNA that had flown under the radar" [129]. Another exemplary study was conducted by Libri et al. [130] using CRAC. They demonstrated that an OFR of murine cytomegalovirus (MCMV), m169, acted as a miRNA sponge and mediated the degradation of miR-27 in host cells.

The integration of large-scale data from CLIP will give some clues for discovering new ceRNAs. StarBase, developed by our laboratory [43,44], provides the most comprehensive prediction of ceRNA networks, miRNA-lncRNA interactions and functions of lncRNAs based on CLIP data.

4 Limitations and further improvements of CLIP

Although CLIP performs quite well and is increasing in popularity, it has some limitations. For better applications of CLIP, we summarize the major defects and present some solutions.

4.1 Reducing noise

By rigorously washing and purifying, CLIP has much less background noise compared with RIP [5], but due to the extreme sensitivity of high-throughput sequencing, there are still background signals that cannot be ignored.

Matthew et al. [131] reported that covalently cross-linked backgrounds were reproducible and most likely universal to PAR-CLIP procedures. This cross-linked background contains T>C conversions and thus cannot be removed by bioinformatic analysis alone. Because background reads are usually identical among libraries from different RBPs and different laboratories, they may be inherently introduced by some PAR-CLIP processes and amplified by high-throughput sequencing. Moreover, it is reasonable that this cross-linked background also exists in other variants of CLIP. Negative controls, such as normal IgG or irrelevant proteins, can be used to efficiently subtract the background and are strongly recommended.

Some background may result from contamination in several steps. Commercial enzymes, such as ligase and proteinase K, are recombinant and contain bacterial RNAs, which can be a source of contamination. These contaminants, most of which are rRNAs, cannot be completely excluded by alignment because some short bacterial rRNAs can also be aligned to the mammalian genome. In improved HITS-CLIP and iCLIP protocols, 3′ linker ligation is performed on beads. Then, unbound bacterial RNA contaminants can be removed by washing, SDS-PAGE and nitro-

cellulose transfer; the bacterial RNAs introduced after nitrocellulose transfer cannot be cloned due to lack of 3′ linker [7]. However, in the PAR-CLIP protocol, a 3′ adaptor is added to free RNA rather than the on-bead RNA; the bacterial RNAs from proteinase K and ligase can also be cloned. Sixty percent to 80% of reads in PAR-CLIP cannot be mapped to the genome [132], which is much more than HITS-CLIP and iCLIP. Adopting the on-bead ligation strategy may improve the mapping rate and thus reduce the background for PAR-CLIP and CLASH.

Biological replicates can be a good method for reducing false positives, and only clusters with good reproducibility should be considered potential RBP binding sites.

4.2 Increasing RNA output efficiency

A large bottleneck for CLIP is low RNA output efficiency, making it difficult to use CLIP to study low abundance RBPs. Low cross-linking efficiency and significant loss of RNAs during the experimental process are the causes.

The cross-linking efficiency of HITS-CLIP is approximately 1%–5% [5]. Although PAR-CLIP improves cross-linking efficiency by introducing 4SU or 6SG, its performance varies among different RBPs. Some BPS, such as HuR, have an efficiency similar to HITS-CLIP [13]. Searching high efficiency, low side effects cross-linking aids is in need for improving performance of CLIP. Methylene blue, which can increase dsRNA-protein cross-linking efficiency to 10%–15% in visible light [133] and may provide an alternative photochemistry regent. Methods based on psoralen photochemistry to decode RNA-protein interactions have been reported [134]. To utilize these regents in CLIP, pre-tests of toxicity, cross-linking efficiency and RNA recovery need to be conducted.

CLIP achieves a high signal-to-noise ratio at the cost of RNA abundance, and only very small parts of the RNAs are reserved after rigorous purification. Thus, the purified RNAs are extremely precious and the productivity of the clone procedure is vital. RNA ligation is the limiting process of library preparation for CLIP. As mentioned above, HITS-CLIP and iCLIP add a 3′ linker to the beads. This strategy can reduce bacterial RNA contamination, but the ligation reaction may be suppressed on the beads. Wang et al. [27] reported that performing a 5′ adaptor ligation rather than a 3′ adaptor on beads can increase overall ligation efficiency, primarily because 5′ adaptor ligation is more efficient. Other improvements, such as using truncated T4 RNA ligase and adding PEG, have been introduced, and good performance of these strategies has been reported [135,136]. CLASH takes great advantage of intermolecular RNA-RNA ligation to explore RNA-RNA interactions directly. However, the intermolecular ligation efficiency is extremely low, approximately 2%. This low ligation efficiency, most likely caused by steric hindrance, hampers widespread use of CLASH. Attempts to overcome or improve this limitation should be made in the future. Recently, Grosswendt et al. [137] reported that intermolecular RNA-RNA ligation also occurred in the absence of the exogenous ligase and the chimeric RNA can also be found in HITS-CLIP and PAR-CLIP.

4.3 Quantitative analysis

Another limitation is the poor performance of quantitative analysis. Theoretically, the normalized CLIP read count is positively correlated to the amount of RNA bound to RBP, which reflects the affinity of the RNA-protein interaction. There are several factors that may bias read count among RNA ligation and PCR are the most determinant. Ligation efficiency varies dramatically, sometimes more than 1000-fold, among RNA with different sequences and secondary structures. Some methods, such as high-efficiency RNA ligation and a pooled adapter strategy, have been employed to reduce the deviation introduced by ligation [136,138]. The barcode sequence in the RNA adaptor also contributes to the inequality of ligation efficiency. Introducing barcodes by PCR rather than in the RNA adaptor can efficiently reduce the bias [139].

The non-linear amplification of PCR becomes another major source of quantitative inaccuracy in high-throughput sequencing. This problem becomes more serious when the low input RNA and requirement of more PCR cycles for CLIP is taken into consideration. Random sequence barcodes provide an ingenious way to remove PCR bias. Reads mapped to the same genome region with identical barcodes are regarded as the duplicates from a single RNA molecule and thus are counted as one. This strategy has been reported to greatly improve the quantification performance of CLIP [18,60,89,140]. Ion Torrent and Ion Proton sequencers require much less cDNA compared to Illumina sequencers. Fewer PCR cycles are accordingly needed, which may reduce non-linear amplification. Recently developed third generation sequencing requires no PCR and allows for direct RNA sequencing [141], which gives a good prospect for quantitative analysis for CLIP. However, due to high read error rates, its performance in CLIP remains to be tested. Moreover, software such as dCLIP [47] has been developed to facilitate the quantitative analysis of CLIP.

5 Conclusion and prospects

The diverse types of RNAs and their spatial and temporal-specific expression contribute greatly to the complexity of life. In addition, the functions of non-coding RNAs are crucial for filling the gap between genotype and phenotype. Thus, the identification and functional study of new species of RNAs will be a momentous task for RNA research. The booming growth of next generation sequencing and knowledge of the human genome offers numerous op-

portunities to systematically identify novel non-coding RNAs. Querying the associated RBPs of these novel non-coding RNAs provides an optimal method to study their functions. CLIP and its variants are precise, and serve as a high-throughput method to decode RNA-protein interactions, with excellent application prospect in studies of RNAs.

Decoding RNA-protein interactions by CLIP provides a good perspective to explore the mysterious RNA world. However, only a few RBPs have been studied by CLIP, which are well characterized and functionally important. To further understand the RNA world, systematic identification and characterization of RBPs in the genome should be conducted. Some preliminary work has been conducted [142−144], and several hundred RBPs were identified. Further characterization and CLIP study of these RBPs will yield valuable information. Recently, adenomatous polyposis coli (APC), a microtubule plus-end scaffolding protein, was identified as an RBP by HITS-CLIP. Further analysis indicated that its associated RNAs were highly enriched for APC-related functions [145].

For functional study of RNAs by their associated RBPs, it is important to distinguish whether the protein is processing-related or function-related. This task is challenging because our understanding of RBPs is limited, and processing and function are sometimes coupled. Subcellular localization can also give some key clues to understanding the function of RBP and its associated RNAs. For example, many studies indicated that Ago proteins could locate in the nucleus and miRNAs could function in the nucleus [146−149], but the targets of miRNA in the nucleus remain unknown. CLIP analysis of nuclear Ago may provide a global understanding of miRNA-target interactions.

Some different RBPs can competitively or cooperatively bind the same RNA motif, and the binding of one RBP may change the RNA-protein interaction of others and thus influence the stability and function of RNAs. The interplay between the RISC complex and other RBPs and the impact on miRNA function have been studied [3,150]. Interestingly, a single RBP, HuR, can either facilitate [151] or attenuate [152] the combining of miRNA-RISC in the 3′ UTR of mRNA. The crosstalk between different RBPs adds a layer of complexity and flexibility to post-transcriptional regulation. Integrating CLIP data from different types of RBPs will provide information on their interplay.

To uncover the hidden RNA world, comprehensive usage of different aspects of information is vital. To this end, meta-analysis of CLIP data, ChIP data, RNA expression profiles and other types of data from different cell types and physiological conditions will broaden our horizons in the RNA world.

We thank Xu ShiJun, Li Bin and Yang JianHua (State Key Laboratory for Biocontrol, Sun Yat-Sen University) for useful suggestion. This work was supported by the National Natural Science Foundation of China (31200593, 31230042), the Guangdong Natural Science Foundation (S2011040001760), the Fundamental Research Funds for the Central Universities (13lgpy40), and the National Basic Research Program of China (2011CB811300).

1 Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Roder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Bar NS, Batut P, Bell K, Bell I, Chakrabortty S, Chen X, Chrast J, Curado J, Derrien T, Drenkow J, Dumais E, Dumais J, Duttagupta R, Falconnet E, Fastuca M, Fejes-Toth K, Ferreira P, Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Gunawardena H, Howald C, Jha S, Johnson R, Kapranov P, King B, Kingswood C, Luo OJ, Park E, Persaud K, Preall JB, Ribeca P, Risk B, Robyr D, Sammeth M, Schaffer L, See LH, Shahab A, Skancke J, Suzuki AM, Takahashi H, Tilgner H, Trout D, Walters N, Wang H, Wrobel J, Yu Y, Ruan X, Hayashizaki Y, Harrow J, Gerstein M, Hubbard T, Reymond A, Antonarakis SE, Hannon G, Giddings MC, Ruan Y, Wold B, Carninci P, Guigo R, Gingeras TR. Landscape of transcription in human cells. Nature, 2012, 489: 101–108

2 Consortium EP, Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F, Parker SCJ, Sabo PJ, Sandstrom R, Shafer A, Vetrie D, Weaver M, Wilcox S, Yu M, Collins FS, Dekker J, Lieb JD, Tullius TD, Crawford GE, Sunyaev S, Noble WS, Dunham I, Denoeud F, Reymond A, Kapranov P, Rozowsky J, Zheng D, Castelo R, Frankish A, Harrow J, Ghosh S, Sandelin A, Hofacker IL, Baertsch R, Keefe D, Dike S, Cheng J, Hirsch HA, Sekinger EA, Lagarde J, Abril JF, Shahab A, Flamm C, Fried C, Hackermüller J, Hertel J, Lindemeyer M, Missal K, Tanzer A, Washietl S, Korbel J, Emanuelsson O, Pedersen JS, Holroyd N, Taylor R, Swarbreck D, Matthews N, Dickson MC, Thomas DJ, Weirauch MT, Gilbert J, Drenkow J, Bell I, Zhao X, Srinivasan KG, Sung W-K, Ooi HS, Chiu KP, Foissac S, Alioto T, Brent M, Pachter L, Tress ML, Valencia A, Choo SW, Choo CY, Ucla C, Manzano C, Wyss C, Cheung E, Clark TG, Brown JB, Ganesh M, Patel S, Tammana H, Chrast J, Henrichsen CN, Kai C, Kawai J, Nagalakshmi U, Wu J, Lian Z, Lian J, Newburger P, Zhang X, Bickel P, Mattick JS, Carninci P, Hayashizaki Y, Weissman S, Hubbard T, Myers RM, Rogers J, Stadler PF, Lowe TM, Wei C-L, Ruan Y, Struhl K, Gerstein M, Antonarakis SE, Fu Y, Green ED, Karaöz U, Siepel A, Taylor J, Liefer LA, Wetterstrand KA, Good PJ, Feingold EA, Guyer MS, Cooper GM, Asimenos G, Dewey CN, Hou M, Nikolaev S, Montoya-Burgos JI, Löytynoja A, Whelan S, Pardi F, Massingham T, Huang H, Zhang NR, Holmes I, Mullikin JC, Ureta-Vidal A, Paten B, Seringhaus M, Church D, Rosenbloom K, Kent WJ, Stone EA, Program NCS, Center BCoMHGS, Center WUGS, Institute B, Institute CsHOR, Batzoglou S, Goldman N, Hardison RC, Haussler D, Miller W, Sidow A, Trinklein ND, Zhang ZD, Barrera L, Stuart R, King DC, Ameur A, Enroth S, Bieda MC, Kim J, Bhinge AA, Jiang N, Liu J, Yao F, Vega VB, Lee CWH, Ng P, Shahab A, Yang A, Moqtaderi Z, Zhu Z, Xu X, Squazzo S, Oberley MJ, Inman D, Singer MA, Richmond TA, Munn KJ, Rada-Iglesias A, Wallerman O, Komorowski J, Fowler JC, Couttet P, Bruce AW, Dovey OM, Ellis PD, Langford CF, Nix DA, Euskirchen G, Hartman S, Urban AE, Kraus P, Van Calcar S, Heintzman N, Kim TH, Wang K, Qu C, Hon G, Luna R, Glass CK, Rosenfeld MG, Aldred SF, Cooper SJ, Halees A, Lin JM, Shulha HP, Zhang X, Xu M, Haidar JNS, Yu Y, Ruan Y, Iyer VR, Green RD, Wadelius C, Farnham PJ, Ren B, Harte RA, Hinrichs AS, Trumbower H, Clawson H, Hillman-Jackson J, Zweig AS, Smith K, Thakkapallayil A, Barber G, Kuhn RM, Karolchik D, Armengol L, Bird CP, de Bakker PIW, Kern AD,

Lopez-Bigas N, Martin JD, Stranger BE, Woodroffe A, Davydov E, Dimas A, Eyras E, Hallgrímsdóttir IB, Huppert J, Zody MC, Abecasis GR, Estivill X, Bouffard GG, Guan X, Hansen NF, Idol JR, Maduro VVB, Maskeri B, McDowell JC, Park M, Thomas PJ, Young AC, Blakesley RW, Muzny DM, Sodergren E, Wheeler DA, Worley KC, Jiang H, Weinstock GM, Gibbs RA, Graves T, Fulton R, Mardis ER, Wilson RK, Clamp M, Cuff J, Gnerre S, Jaffe DB, Chang JL, Lindblad-Toh K, Lander ES, Koriabine M, Nefedov M, Osoegawa K, Yoshinaga Y, Zhu B, de Jong PJ. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature, 2007, 447: 799–816

3 Ho JJD, Marsden PA. Competition and collaboration between RNA-binding proteins and microRNAs. Wiley Interdiscip Rev RNA, 2014, 5: 69–86

4 Moore MJ, Zhang C, Gantman EC, Mele A, Darnell JC, Darnell RB. Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. Nat Protoc, 2014, 9: 263–293

5 Darnell RB. HITS-CLIP: panoramic views of protein-RNA regulation in living cells. Wiley Interdiscip Rev RNA, 2010, 1: 266–286

6 Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB. CLIP identifies Nova-regulated RNA networks in the brain. Science, 2003, 302: 1212–1215

7 Ule J, Jensen K, Mele A, Darnell RB. CLIP: a method for identifying protein-RNA interaction sites in living cells. Methods, 2005, 37: 376–386

8 Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, Darnell JC, Darnell RB. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature, 2008, 456: 464–469

9 Chi SW, Zang JB, Mele A, Darnell RB. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature, 2009, 460: 479–486

10 Licatalosi DD, Darnell RB. RNA processing and its regulation: global insights into biological networks. Nat Rev Genet, 2010, 11: 75–87

11 Favre A, Saintomé C, Fourrey JL, Clivio P, Laugâa P. Thionucleobases as intrinsic photoaffinity probes of nucleic acid structure and nucleic acid-protein interactions. J Photochem Photobiol B, 1998, 42: 109–124

12 Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jr., Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell, 2010, 141: 129–141

13 Ascano M, Hafner M, Cekan P, Gerstberger S, Tuschl T. Identification of RNA-protein interaction networks using PAR-CLIP. Wiley Interdiscip Rev RNA, 2012, 3: 159–177

14 Corcoran DL, Georgiev S, Mukherjee N, Gottwein E, Skalsky RL, Keene JD, Ohler U. PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data. Genome Biol, 2011, 12: R79

15 Kemény-Beke A, Berényi E, Facskó A, Damjanovich J, Horváth A, Bodnár A, Berta A, Aradi J. Antiproliferative effect of 4-thiouridylate on OCM-1 uveal melanoma cells. Eur J Ophthalmol, 2006, 16: 680–685

16 Burger K, Mühl B, Kellner M, Rohrmoser M, Gruber-Eber A, Windhager L, Friedel CC, Dölken L, Eick D. 4-thiouridine inhibits rRNA synthesis and causes a nucleolar stress response. RNA Biol, 2013, 10: 1623–1630

17 Zhang C, Darnell RB. Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. Nat Biotechnol, 2011, 29: 607–614

18 König J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol, 2010, 17: 909–915

19 Urlaub H, Hartmuth K, Lührmann R. A two-tracked approach to analyze RNA-protein crosslinking sites in native, nonlabeled small nuclear ribonucleoprotein particles. Methods, 2002, 26: 170–181

20 Granneman S, Kudla G, Petfalski E, Tollervey D. 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, 2009, 106: 9613–9618

21 Kudla G, Granneman S, Hahn D, Beggs JD, Tollervey D. Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. Proc Natl Acad Sci USA, 2011, 108: 10010–10015

22 Helwak A, Kudla G, Dudnakova T, Tollervey D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell, 2013, 153: 654–665

23 Hüttenhofer A, Schattner P. The principles of guiding by RNA: chimeric RNA-protein enzymes. Nat Rev Genet, 2006, 7: 475–482

24 Akbari Moqadam F, Pieters R, den Boer ML. The hunting of targets: challenge in miRNA research. Leukemia, 2013, 27: 16–23

25 Broughton JP, Pasquinelli AE. Identifying Argonaute binding sites in *Caenorhabditis elegans* using iCLIP. Methods, 2013, 63: 119–125

26 Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jungkamp A-C, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. PAR-CliP—a method to identify transcriptome-wide the binding sites of RNA binding proteins. J Vis Exp, 2010, doi: 10.3791/2034

27 Wang Z, Tollervey J, Briese M, Turner D, Ule J. CLIP: construction of cDNA libraries for high-throughput sequencing from RNAs cross-linked to proteins *in vivo*. Methods, 2009, 48: 287–293

28 Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics, 2011, 27: 863–864

29 Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. BMC Bioinformatics, 2009, 10: 421

30 Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 2009, 25: 1754–1760

31 Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods, 2012, 9: 357–359

32 Uren PJ, Bahrami-Samani E, Burns SC, Qiao M, Karginov FV, Hodges E, Hannon GJ, Sanford JR, Penalva LOF, Smith AD. Site identification in high-throughput RNA-protein interaction data. Bioinformatics, 2012, 28: 3013–3020

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

34 Li Y, Zhao DY, Greenblatt JF, Zhang Z. RIPSeeker: a statistical package for identifying protein-associated transcripts from RIP-seq experiments. Nucleic Acids Res, 2013, 41: e94

35 Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol, 1994, 2: 28–36

36 Thompson WA, Newberg LA, Conlan S, McCue LA, Lawrence CE. The Gibbs Centroid Sampler. Nucleic Acids Res, 2007, 35: W232–W237

37 Roth FP, Hughes JD, Estep PW, Church GM. Finding DNA regulatory motifs within unaligned noncoding sequences clustered by whole-genome mRNA quantitation. Nat Biotechnol, 1998, 16: 939–945

38 van Helden J. Regulatory sequence analysis tools. Nucleic Acids Res, 2003, 31: 3593–3596

39 Fukunaga T, Ozaki H, Terai G, Asai K, Iwasaki W, Kiryu H. CapR: revealing structural specificities of RNA-binding protein target recognition using CLIP-seq data. Genome Biol, 2014, 15: R16

40 Maticzka D, Lange SJ, Costa F, Backofen R. GraphProt: modeling binding preferences of RNA-binding proteins. Genome Biol, 2014, 15: R17

41 Khorshid M, Rodak C, Zavolan M. CLIPZ: a database and analysis environment for experimentally determined binding sites of RNA-binding proteins. Nucleic Acids Res, 2011, 39: D245–252

42 Anders G, Mackowiak SD, Jens M, Maaskola J, Kuntzagk A, Rajewsky N, Landthaler M, Dieterich C. doRiNA: a database of RNA interactions in post-transcriptional regulation. Nucleic Acids Res, 2012, 40: D180–186

43 Yang JH, Li JH, Shao P, Zhou H, Chen YQ, Qu LH. starBase: a da-

tabase for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. Nucleic Acids Res, 2011, 39: D202–209
44 Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res, 2014, 42: D92–97
45 Sievers C, Schlumpf T, Sawarkar R, Comoglio F, Paro R. Mixture models and wavelet transforms reveal high confidence RNA-protein interaction sites in MOV10 PAR-CLIP data. Nucleic Acids Res, 2012, 40: e160
46 Yun J, Wang T, Xiao G. Bayesian hidden Markov models to identify RNA-protein interaction sites in PAR-CLIP. Biometrics, 2014, doi: 10.1111/biom.12147
47 Wang T, Xie Y, Xiao G. dCLIP: a computational approach for comparative CLIP-seq analyses. Genome Biol, 2014, 15: R11
48 Kucukural A, Özadam H, Singh G, Moore MJ, Cenik C. ASPeak: an abundance sensitive peak detection algorithm for RIP-Seq. Bioinformatics, 2013, 29: 2485–2486
49 Chen B, Yun J, Kim MS, Mendell JT, Xie Y. PIPE-CLIP: a comprehensive online tool for CLIP-seq data analysis. Genome Biol, 2014, 15: R18
50 Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell, 2009, 136: 215–233
51 Ambros V. The functions of animal microRNAs. Nature, 2004, 431: 350–355
52 Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell, 2005, 120: 15–20
53 Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N. Combinatorial microRNA target predictions. Nat Genet, 2005, 37: 495–500
54 Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R. Fast and effective prediction of microRNA/target duplexes. RNA, 2004, 10: 1507–1517
55 Rajewsky N. microRNA target predictions in animals. Nat Genet, 2006, 38(Suppl): S8–13
56 Bentwich I. Prediction and validation of microRNAs and their targets. FEBS Lett, 2005, 579: 5904–5910
57 Wang Y, Juranek S, Li H, Sheng G, Tuschl T, Patel DJ. Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. Nature, 2008, 456: 921–926
58 Friedman RC, Farh KK-H, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res, 2009, 19: 92–105
59 Chi SW, Hannon GJ, Darnell RB. An alternative mode of microRNA target recognition. Nat Struct Mol Biol, 2012, 19: 321–327
60 Loeb GB, Khan AA, Canner D, Hiatt JB, Shendure J, Darnell RB, Leslie CS, Rudensky AY. Transcriptome-wide miR-155 binding map reveals widespread noncanonical microRNA targeting. Mol Cell, 2012, 48: 760–770
61 Kim KK, Ham J, Chi SW. miRTCat: a comprehensive map of human and mouse microRNA target sites including non-canonical nucleation bulges. Bioinformatics, 2013, 29: 1898–1899
62 Betel D, Koppal A, Agius P, Sander C, Leslie C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. Genome Biol, 2010, 11: R90
63 Khorshid M, Hausser J, Zavolan M, van Nimwegen E. A biophysical miRNA-mRNA interaction model infers canonical and noncanonical targets. Nat Methods, 2013, 10: 253–255
64 Skalsky RL, Cullen BR. Viruses, microRNAs, and host interactions. Annu Rev Microbiol, 2010, 64: 123–141
65 Cullen BR. Viruses and microRNAs: RISCy interactions with serious consequences. Genes Dev, 2011, 25: 1881–1894
66 Gottwein E, Mukherjee N, Sachse C, Frenzel C, Majoros WH, Chi JTA, Braich R, Manoharan M, Soutschek J, Ohler U, Cullen BR. A viral microRNA functions as an orthologue of cellular miR-155. Nature, 2007, 450: 1096–1099
67 Gottwein E, Corcoran DL, Mukherjee N, Skalsky RL, Hafner M, Nusbaum JD, Shamulailatpam P, Love CL, Dave SS, Tuschl T, Ohler U, Cullen BR. Viral microRNA targetome of KSHV-infected primary effusion lymphoma cell lines. Cell Host Microbe, 2011, 10: 515–526
68 Haecker I, Gay LA, Yang Y, Hu J, Morse AM, McIntyre LM, Renne R. Ago HITS-CLIP expands understanding of Kaposi's sarcoma-associated herpesvirus miRNA function in primary effusion lymphomas. PLoS Pathog, 2012, 8: e1002884
69 Skalsky RL, Corcoran DL, Gottwein E, Frank CL, Kang D, Hafner M, Nusbaum JD, Feederle R, Delecluse H-J, Luftig MA, Tuschl T, Ohler U, Cullen BR. The viral and cellular microRNA targetome in lymphoblastoid cell lines. PLoS Pathog, 2012, 8: e1002484
70 Riley KJ, Rabinowitz GS, Yario TA, Luna JM, Darnell RB, Steitz JA. EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. EMBO J, 2012, 31: 2207–2221
71 Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet, 2007, 39: 380–385
72 Anokye-Danso F, Trivedi CM, Juhr D, Gupta M, Cui Z, Tian Y, Zhang Y, Yang W, Gruber PJ, Epstein JA, Morrisey EE. Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. Cell Stem Cell, 2011, 8: 376–388
73 Judson RL, Babiarz JE, Venere M, Blelloch R. Embryonic stem cell-specific microRNAs promote induced pluripotency. Nat Biotechnol, 2009, 27: 459–461
74 Miyoshi N, Ishii H, Nagano H, Haraguchi N, Dewi DL, Kano Y, Nishikawa S, Tanemura M, Mimori K, Tanaka F, Saito T, Nishimura J, Takemasa I, Mizushima T, Ikeda M, Yamamoto H, Sekimoto M, Doki Y, Mori M. Reprogramming of mouse and human cells to pluripotency using mature microRNAs. Cell stem cell, 2011, 8: 633–638
75 Xie S, Zhang Y, Qu L, Xu H. A Helm model for microRNA regulation in cell fate decision and conversion. Sci China Life Sci, 2013, 56: 897–906
76 Leonardo TR, Schultheisz HL, Loring JF, Laurent LC. The functions of microRNAs in pluripotency and reprogramming. Nat Cell Biol, 2012, 14: 1114–1121
77 Leung AK, Young AG, Bhutkar A, Zheng GX, Bosson AD, Nielsen CB, Sharp PA. Genome-wide identification of Ago2 binding sites from mouse embryonic stem cells with and without mature microRNAs. Nat Struct Mol Biol, 2011, 18: 237–244
78 Lipchina I, Elkabetz Y, Hafner M, Sheridan R, Mihailovic A, Tuschl T, Sander C, Studer L, Betel D. Genome-wide identification of microRNA targets in human ES cells reveals a role for miR-302 in modulating BMP response. Genes Dev, 2011, 25: 2173–2186
79 Mendell JT. miRiad roles for the miR-17-92 cluster in development and disease. Cell, 2008, 133: 217–222
80 Jin HY, Oda H, Lai M, Skalsky RL, Bethel K, Shepherd J, Kang SG, Liu WH, Sabouri-Ghomi M, Cullen BR, Rajewsky K, Xiao C. MicroRNA-17~92 plays a causative role in lymphomagenesis by coordinating multiple oncogenic pathways. EMBO J, 2013, 32: 2377–2391
81 Boudreau RL, Jiang P, Gilmore BL, Spengler RM, Tirabassi R, Nelson JA, Ross CA, Xing Y, Davidson BL. Transcriptome-wide Discovery of microRNA Binding Sites in Human Brain. Neuron, 2013, doi: 10.1016/j.neuron.2013.10.062
82 Kameswaran V, Bramswig NC, McKenna LB, Penn M, Schug J, Hand NJ, Chen Y, Choi I, Vourekas A, Won K-J, Liu C, Vivek K, Naji A, Friedman JR, Kaestner KH. Epigenetic regulation of the DLK1-MEG3 microRNA cluster in human type 2 diabetic islets. Cell Metab, 2014, 19: 135–145
83 Black DL. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem, 2003, 72: 291–336
84 Kornblihtt AR, Schor IE, Alló M, Dujardin G, Petrillo E, Muñoz MJ. Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat Rev Mol Cell Biol, 2013, 14: 153–165
85 Xue Y, Zhou Y, Wu T, Zhu T, Ji X, Kwon YS, Zhang C, Yeo G, Black DL, Sun H, Fu XD, Zhang Y. Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. Mol Cell, 2009, 36:

996–1006
86 Licatalosi DD, Yano M, Fak JJ, Mele A, Grabinski SE, Zhang C, Darnell RB. Ptbp2 represses adult-specific splicing to regulate the generation of neuronal precursors in the embryonic brain. Genes Dev, 2012, 26: 1626–1642
87 Yeo GW, Coufal NG, Liang TY, Peng GE, Fu X-D, Gage FH. An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nat Struct Mol Biol, 2009, 16: 130–137
88 Sanford JR, Wang X, Mort M, Vanduyn N, Cooper DN, Mooney SD, Edenberg HJ, Liu Y. Splicing factor SFRS1 recognizes a functionally diverse landscape of RNA transcripts. Genome Res, 2009, 19: 381–394
89 Wang Z, Kayikci M, Briese M, Zarnack K, Luscombe NM, Rot G, Zupan B, Curk T, Ule J. iCLIP predicts the dual splicing effects of TIA-RNA interactions. PLoS Biol, 2010, 8: e1000530
90 Wang ET, Cody NAL, Jog S, Biancolella M, Wang TT, Treacy DJ, Luo S, Schroth GP, Housman DE, Reddy S, Lécuyer E, Burge CB. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. Cell, 2012, 150: 710–724
91 Daughters RS, Tuttle DL, Gao W, Ikeda Y, Moseley ML, Ebner TJ, Swanson MS, Ranum LPW. RNA gain-of-function in spinocerebellar ataxia type 8. PLoS Genet, 2009, 5: e1000600
92 Pandit S, Zhou Y, Shiue L, Coutinho-Mansfield G, Li H, Qiu J, Huang J, Yeo GW, Ares M Jr., Fu XD. Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. Mol Cell, 2013, 50: 223–235
93 Zarnack K, König J, Tajnik M, Martincorena I, Eustermann S, Stévant I, Reyes A, Anders S, Luscombe NM, Ule J. Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. Cell, 2013, 152: 453–466
94 Huelga SC, Vu AQ, Arnold JD, Liang TY, Liu PP, Yan BY, Donohue JP, Shiue L, Hoon S, Brenner S, Ares M, Yeo GW. Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. Cell Rep, 2012, 1: 167–178
95 Mayeda A, Krainer AR. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell, 1992, 68: 365–375
96 Mayeda A, Munroe SH, Cáceres JF, Krainer AR. Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. EMBO J, 1994, 13: 5483–5495
97 Cáceres JF, Stamm S, Helfman DM, Krainer AR. Regulation of alternative splicing *in vivo* by overexpression of antagonistic splicing factors. Science, 1994, 265: 1706–1709
98 Guil S, Cáceres JF. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. Nat Struct Mol Biol, 2007, 14: 591–596
99 Xue Y, Ouyang K, Huang J, Zhou Y, Ouyang H, Li H, Wang G, Wu Q, Wei C, Bi Y, Jiang L, Cai Z, Sun H, Zhang K, Zhang Y, Chen J, Fu XD. Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. Cell, 2013, 152: 82–96
100 Makeyev EV, Zhang J, Carrasco MA, Maniatis T. The microRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. Mol Cell, 2007, 27: 435–448
101 Sparmann A, van Lohuizen M. Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer, 2006, 6: 846–856
102 Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science, 2008, 322: 750–756
103 Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell, 2007, 129: 1311–1323
104 Guil S, Soler M, Portela A, Carrère J, Fonalleras E, Gómez A, Villanueva A, Esteller M. Intronic RNAs mediate EZH2 regulation of epigenetic targets. Nat Struct Mol Biol, 2012, 19: 664–670
105 Kaneko S, Son J, Shen SS, Reinberg D, Bonasio R. PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. Nat Struct Mol Biol, 2013, 20: 1258–1264
106 Castel SE, Martienssen RA. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat Rev Genet, 2013, 14: 100–112
107 Lee JT. Epigenetic regulation by long noncoding RNAs. Science, 2012, 338: 1435–1439
108 Jeffery L, Nakielny S. Components of the DNA methylation system of chromatin control are RNA-binding proteins. J Biol Chem, 2004, 279: 49479–49487
109 Davis BN, Hilyard AC, Nguyen PH, Lagna G, Hata A. Smad proteins bind a conserved RNA sequence to promote microRNA maturation by Drosha. Mol Cell, 2010, 39: 373–384
110 Riley KJ-L, James Maher L. Analysis of p53-RNA interactions in cultured human cells. Biochem Biophys Res Commun, 2007, 363: 381–387
111 Riley KJL, Maher LJ. p53 RNA interactions: new clues in an old mystery. RNA, 2007, 13: 1825–1833
112 Cassiday LA, Maher LJ. Having it both ways: transcription factors that bind DNA and RNA. Nucleic Acids Res, 2002, 30: 4118–4126
113 Suswam EA, Li YY, Mahtani H, King PH. Novel DNA-binding properties of the RNA-binding protein TIAR. Nucleic Acids Res, 2005, 33: 4507–4518
114 Law WJ, Cann KL, Hicks GG. TLS, EWS and TAF15: a model for transcriptional integration of gene expression. Brief Funct Genomic Proteomic, 2006, 5: 8–14
115 Huang V, Zheng J, Qi Z, Wang J, Place RF, Yu J, Li H, Li LC. Ago1 Interacts with RNA polymerase II and binds to the promoters of actively transcribed genes in human cancer cells. PLoS Genet, 2013, 9: e1003821
116 Phillips JE, Corces VG. CTCF: master weaver of the genome. Cell, 2009, 137: 1194–1211
117 Saldana-Meyer R, Gonzalez-Buendia E, Guerrero G, Narendra V, Bonasio R, Recillas-Targa F, Reinberg D. CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. Genes Dev, 2014, 28: 723–734
118 Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell, 2011, 146: 353–358
119 Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. Nature, 2014, 505: 344–352
120 Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature, 2010, 465: 1033–1038
121 Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, Tramontano A, Bozzoni I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell, 2011, 147: 358–369
122 Kallen AN, Zhou XB, Xu J, Qiao C, Ma J, Yan L, Lu L, Liu C, Yi JS, Zhang H, Min W, Bennett AM, Gregory RI, Ding Y, Huang Y. The imprinted H19 lncRNA antagonizes let-7 microRNAs. Mol Cell, 2013, 52: 101–112
123 Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, Leyva A, Weigel D, García JA, Paz-Ares J. Target mimicry provides a new mechanism for regulation of microRNA activity. Nat Genet, 2007, 39: 1033–1037
124 Karreth FA, Tay Y, Perna D, Ala U, Tan SM, Rust AG, DeNicola G, Webster KA, Weiss D, Perez-Mancera PA, Krauthammer M, Halaban R, Provero P, Adams DJ, Tuveson DA, Pandolfi PP. In vivo identification of tumor-suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. Cell, 2011, 147: 382–395
125 Sumazin P, Yang X, Chiu HS, Chung WJ, Iyer A, Llobet-Navas D, Rajbhandari P, Bansal M, Guarnieri P, Silva J, Califano A. An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. Cell, 2011, 147: 370–381
126 Kumar MS, Armenteros-Monterroso E, East P, Chakravorty P, Matthews N, Winslow MM, Downward J. HMGA2 functions as a competing endogenous RNA to promote lung cancer progression. Nature, 2014, 505: 212–217
127 Thomas M, Lieberman J, Lal A. Desperately seeking microRNA tar-

gets. Nat Struct Mol Biol, 2010, 17: 1169–1174
128 Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, Loewer A, Ziebold U, Landthaler M, Kocks C, le Noble F, Rajewsky N. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature, 2013, 495: 333–338
129 Ledford H. Circular RNAs throw genetics for a loop. Nature, 2013, 494: 415
130 Libri V, Helwak A, Miesen P, Santhakumar D, Borger JG, Kudla G, Grey F, Tollervey D, Buck AH. Murine cytomegalovirus encodes a miR-27 inhibitor disguised as a target. Proc Natl Acad Sci USA, 2012, 109: 279–284
131 Friedersdorf MB, Keene JD. Advancing the functional utility of PAR-CLIP by quantifying background binding to mRNAs and lncRNAs. Genome Biol, 2014, 15: R2
132 Hafner M, Lianoglou S, Tuschl T, Betel D. Genome-wide identification of miRNA targets by PAR-CLIP. Methods, 2012, 58: 94–105
133 Liu ZR, Wilkie AM, Clemens MJ, Smith CW. Detection of double-stranded RNA-protein interactions by methylene blue-mediated photo-crosslinking. RNA, 1996, 2: 611–621
134 Wang Z, Rana TM. Probing RNA-protein interactions by psoralen photocrosslinking. Methods Mol Biol, 1999, 118: 49–62
135 Hafner M, Renwick N, Farazi TA, Mihailovic A, Pena JT, Tuschl T. Barcoded cDNA library preparation for small RNA profiling by next-generation sequencing. Methods, 2012, 58: 164–170
136 Zhang Z, Lee JE, Riemondy K, Anderson EM, Yi R. High-efficiency RNA cloning enables accurate quantification of miRNA expression by deep sequencing. Genome Biol, 2013, 14: R109
137 Grosswendt S, Filipchyk A, Manzano M, Klironomos F, Schilling M, Herzog M, Gottwein E, Rajewsky N. Unambiguous identification of miRNA: target site interactions by different types of ligation reactions. Mol Cell, 2014, 54: 1042–1054
138 Jayaprakash AD, Jabado O, Brown BD, Sachidanandam R. Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. Nucleic Acids Res, 2011, 39: e141
139 Alon S, Vigneault F, Eminaga S, Christodoulou DC, Seidman JG, Church GM, Eisenberg E. Barcoding bias in high-throughput multiplex sequencing of miRNA. Genome Res, 2011, 21: 1506–1511
140 Kishore S, Jaskiewicz L, Burger L, Hausser J, Khorshid M, Zavolan M. A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. Nat Methods, 2011, 8: 559–564
141 Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. Hum Mol Genet, 2010, 19: R227–240
142 Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM, Krijgsveld J, Hentze MW. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell, 2012, 149: 1393–1406
143 Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS, Li X, Gueroussov S, Albu M, Zheng H, Yang A, Na H, Irimia M, Matzat LH, Dale RK, Smith SA, Yarosh CA, Kelly SM, Nabet B, Mecenas D, Li W, Laishram RS, Qiao M, Lipshitz HD, Piano F, Corbett AH, Carstens RP, Frey BJ, Anderson RA, Lynch KW, Penalva LOF, Lei EP, Fraser AG, Blencowe BJ, Morris QD, Hughes TR. A compendium of RNA-binding motifs for decoding gene regulation. Nature, 2013, 499: 172–177
144 Kwon SC, Yi H, Eichelbaum K, Fohr S, Fischer B, You KT, Castello A, Krijgsveld J, Hentze MW, Kim VN. The RNA-binding protein repertoire of embryonic stem cells. Nat Struct Mol Biol, 2013, 20: 1122–1130
145 Preitner N, Quan J, Nowakowski DW, Hancock ML, Shi J, Tcherkezian J, Young-Pearse TL, Flanagan JG. APC is an RNA-binding protein, and its interactome provides a link to neural development and microtubule assembly. Cell, 2014, 158: 368–382
146 Gagnon KT, Li L, Chu Y, Janowski BA, Corey DR. RNAi factors are present and active in human cell nuclei. Cell Rep, 2014, 6: 211–221
147 Truesdell SS, Mortensen RD, Seo M, Schroeder JC, Lee JH, LeTonqueze O, Vasudevan S. MicroRNA-mediated mRNA translation activation in quiescent cells and oocytes involves recruitment of a nuclear microRNP. Sci Rep, 2012, 2: 842
148 Gagnon KT, Corey DR. Argonaute and the nuclear RNAs: new pathways for RNA-mediated control of gene expression. Nucleic Acid Ther, 2012, 22: 3–16
149 Liao JY, Ma LM, Guo YH, Zhang YC, Zhou H, Shao P, Chen YQ, Qu LH. Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of miRNAs and tRNA 3′ trailers. PLoS One, 2010, 5: e10563
150 van Kouwenhove M, Kedde M, Agami R. MicroRNA regulation by RNA-binding proteins and its implications for cancer. Nat Rev Cancer, 2011, 11: 644–656
151 Kim HH, Kuwano Y, Srikantan S, Lee EK, Martindale JL, Gorospe M. HuR recruits let-7/RISC to repress c-Myc expression. Genes Dev, 2009, 23: 1743–1748
152 Kundu P, Fabian MR, Sonenberg N, Bhattacharyya SN, Filipowicz W. HuR protein attenuates miRNA-mediated repression by promoting miRISC dissociation from the target RNA. Nucleic Acids Res, 2012, 40: 5088–5100