

REVIEW

Mapping transcriptome-wide protein-RNA interactions to elucidate RNA regulatory programs

Molly M. Hannigan[†], Leah L. Zagore[†] and Donny D. Licatalosi*

Center for RNA Science and Therapeutics, Case Western Reserve University, Cleveland, OH 44106, USA

* Correspondence: ddl33@case.edu

Received January 20, 2018; Revised March 27, 2018; Accepted April 3, 2018

Background: Our understanding of post-transcriptional gene regulation has increased exponentially with the development of robust methods to define protein-RNA interactions across the transcriptome. In this review, we highlight the evolution and successful applications of crosslinking and immunoprecipitation (CLIP) methods to interrogate protein-RNA interactions in a transcriptome-wide manner.

Results: Here, we survey the vast array of *in vitro* and *in vivo* approaches used to identify protein-RNA interactions, including but not limited to electrophoretic mobility shift assays, systematic evolution of ligands by exponential enrichment (SELEX), and RIP-seq. We particularly emphasize the advancement of CLIP technologies, and detail protocol improvements and computational tools used to analyze the output data. Importantly, we discuss how profiling protein-RNA interactions can delineate biological functions including splicing regulation, alternative polyadenylation, cytoplasmic decay substrates, and miRNA targets.

Conclusions: In summary, this review summarizes the benefits of characterizing RNA-protein networks to further understand the regulation of gene expression and disease pathogenesis. Our review comments on how future CLIP technologies can be adapted to address outstanding questions related to many aspects of RNA metabolism and further advance our understanding of RNA biology.

Keywords: RNA binding proteins; CLIP; post-transcriptional regulation; RNA networks

Author summary: Rapidly advancing sequencing methodologies continue to improve our understanding of gene regulation and disease pathogenesis. An emerging area of interest is the contribution of RNA regulation in development and disease. In this review, we highlight techniques widely used to interrogate RNA regulatory networks that fine tune gene expression and how they impact cell biology.

INTRODUCTION

The diversity in morphology and function of the numerous cell types in the human body is driven primarily through differential gene expression. Perturbations in normal gene expression, at the level of both transcription and post-transcriptional regulatory mechanisms, lead to a number of human diseases, including cancer and neurodegenerative disorders [1–3]. Elucidating regulatory

mechanisms that control gene expression networks is therefore critical to understand the complexity of development, cell differentiation, and disease.

RNA binding proteins (RBPs) orchestrate many essential processes in the expression of protein-coding genes. This includes co-transcriptional pre-mRNA processing events such as exon splicing and 3' end formation, as well as mRNA export to the cytoplasm, localization, translation, and decay [4]. It is estimated that up to 1,500 proteins are able to bind RNA in humans [5,6], illustrating

[†] These authors contributed equally to this work.

the complexity of post-transcriptional gene regulatory networks. Furthermore, protein-RNA interactions are dynamically assembled and remodeled at different stages of the mRNA lifecycle [7], with the fate of a particular mRNA determined by the composition of attendant proteins. Accordingly, the diverse spectrum of RBPs expressed in different cell types and developmental stages generates distinct programs of combinatorial RNA regulation to fine tune gene output in different biological contexts.

Here, we review high-throughput, *in vivo*, protein-centric techniques used to explore protein-RNA interactions and highlight the power of integrative approaches to advance our understanding of how complex protein-RNA interaction networks regulate the transcriptome to specialize cell properties.

REDUCTIONIST APPROACHES TO INTERROGATE PROTEIN-RNA INTERACTIONS

Historically, the study of protein-RNA interactions was limited to reconstituted *in vivo* and *in vitro* approaches, including quantitative methods such as electrophoretic mobility shift assays (EMSA) [8], surface plasmon resonance [9], and yeast three-hybrid systems [10]. Additionally, *in vitro* selection of RNA [11] or systematic evolution of ligands by exponential enrichment (SELEX) [12] can identify subpopulations of RNA molecules that bind to a particular ligand with high affinity. While these approaches provide detailed kinetic information regarding the specificity and affinity of protein-RNA interactions, they do not recapitulate the native cellular environments that occur *in vivo*. In contrast, RNA immunoprecipitation (RIP) uses a protein-specific antibody to co-precipitate associated RNAs followed by sequencing (RIP-seq) or microarray analysis (RIP-chip) [13] to identify global RNA substrates. While RIP has been used successfully in many studies to identify relevant RNA targets [13,14], the technique does not identify direct sites of protein interaction with RNA. Additionally, in some conditions, protein-RNA complexes can remodel after cell lysis, leading to the capture of non-physiological interactions [15].

The use of UV crosslinking to lock protein-RNA interactions *in vivo* overcomes the limitations inherent to traditional RIP procedures. When exposed to UV light, photoreactive molecules of RNA form a covalent association with directly bound proteins, a principle first described by Dreyfuss and colleagues in the 1980s [16]. The combination of UV crosslinking with immunoprecipitation, or crosslinking and immunoprecipitation (CLIP), is a powerful approach to identify direct protein-RNA

interactions [17]. Because of the covalent crosslinks formed, stringent purification techniques can be applied during the immunoprecipitation, improving the signal to noise ratio. Additionally, by partially digesting the RNA prior to immunoprecipitation, the interactions captured by CLIP provide positional binding information, as well as a list of direct RNA substrates.

The first method for isolating and purifying small RNA fragments crosslinked to a particular RBP was developed in the Darnell lab by Ule *et al.* in 2003 [17]. Briefly, triturated tissue or cells is UV irradiated to crosslink protein to RNA *in vivo*, followed by cell lysis and RNA fragmentation with RNase. The protein of interest and associated RNA fragments are purified by immunoprecipitation with stringent washes. This is followed by ligation of a RNA linker to the 3' end of the RNA fragments, and radiolabeling of the 5' end. The radiolabeled protein-RNA complexes are then subjected to SDS-PAGE and transferred to nitrocellulose. These steps allow the separation of protein-RNA complexes based on RNA fragment length, removal of free RNA, and visualization of protein-RNA complexes by autoradiography. After extraction of protein-RNA complexes from a nitrocellulose membrane, the target protein is degraded by proteinase K and a second RNA linker is ligated to the 5' end of RNA. Reverse-transcription PCR is then performed to generate cDNA that either can be cloned into plasmids for Sanger sequencing (as in the original study) or sequenced using massively parallel sequencing (Figure 1). In the first study [17], CLIP was used to identify RNAs in mouse brain that are directly bound by NOVA, a neuron-specific splicing factor. Importantly, the functional importance of specific NOVA-RNA interactions identified could be interrogated using a gene knockout mouse lacking NOVA expression. These efforts showed that many RNA targets dependent on NOVA for proper alternative splicing had known functions in synaptic biology. The data therefore provided new insights into the aberrant neurological phenotype of NOVA KO mice while also validating CLIP as a powerful tool for RNA molecular biology research.

HIGH-RESOLUTION TRANSCRIPTOME-WIDE MAPPING OF PROTEIN-RNA INTERACTIONS

Applications of CLIP methods have vast potential to facilitate new mechanistic insights into RBP functions and post-transcriptional control of gene expression. Over the last decade, a plethora of CLIP based techniques have been developed, including hiCLIP [18], CLASH [19], irCLIP [20], RipIT-seq [21], TRIBE [22], sCLIP [23], dCLIP [24], and Fr-iCLIP [25]. Each technique improves

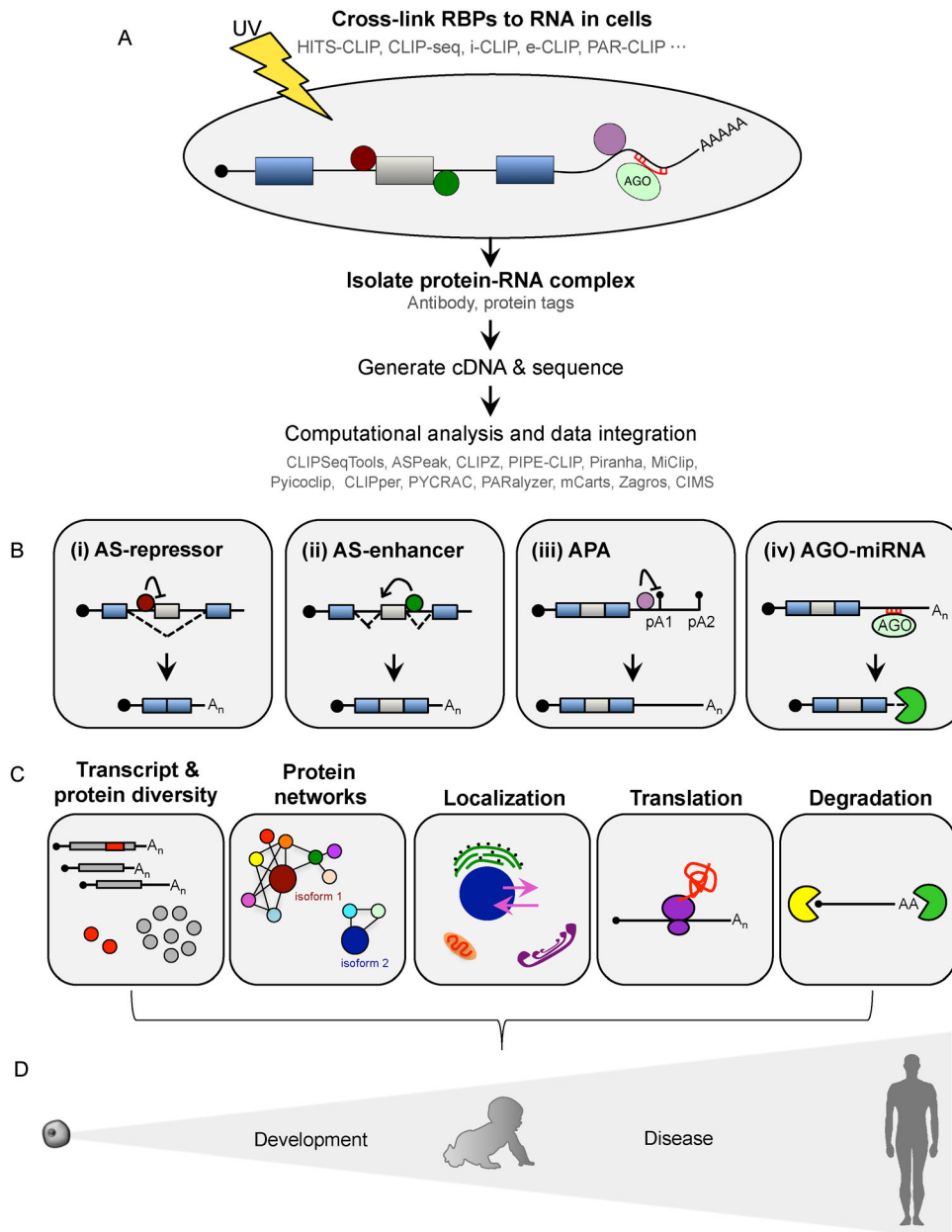


Figure 1. Integration of CLIP methods with computational analysis to explore protein-RNA interactions and their biological impact.

(A) Simplified schematic illustrating the CLIP workflow to generate transcriptome-side maps of RBP-RNA interactions occurring *in vivo*. (B) Integration of next generation sequencing with CLIP data can be used to identify RBPs that function as (i) repressors of alternative splicing (AS), (ii) enhancers of alternative splicing, or (iii) regulators of alternative polyadenylation (APA). CLIP can also be performed on AGO proteins to map miRNA binding sites that downregulate target mRNAs (iv). (C) Application of CLIP methods to define RBP regulatory networks from position-dependent and combinatorial functions. CLIP profiles, in conjunction with high-throughput sequencing, can characterize transcript/protein diversity in specific cellular contexts, protein networks, localization patterns, translational fates of mRNAs, and substrates of mRNA decay. (D) Identification of complex protein-RNA interaction networks that regulate gene expression is critical to understanding human development and disease.

or refines a particular aspect of the methodology. Below, we review the most widely used adaptations.

HITS-CLIP/CLIP-seq

The emergence of next generation sequencing platforms in the mid 2000s, such as 454 Life Sciences and Solexa ushered in a new era of RNA molecular biology research. For the first time, the transcriptome could be examined in a global, unbiased, and high-resolution manner, thus revolutionizing our understanding of transcriptome complexity and its regulation. Integrating high-throughput sequencing with CLIP (HITS-CLIP or CLIP-seq) led to a robust expansion of the number of protein-RNA interactions that could be identified from a single experiment. To emphasize the magnitude of this advance, the first CLIP experiment generated 340 NOVA CLIP reads from mouse brain using Sanger sequencing [17]. In comparison, a 2008 study coupling CLIP with high-throughput sequencing [26] generated ~168,000 CLIP reads and revealed thousands of biologically reproducible NOVA-RNA interactions across the brain transcriptome.

PAR-CLIP

UV crosslinking efficiency *in vivo* is estimated to be quite low — on the order of 1%–5% [27]. Photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) improves the crosslinking efficiency, reportedly improving RNA recovery 100 to 1,000 fold as compared to traditional UV 254 nm crosslinking [28]. Such improvement is achieved by the incorporation of UV-reactive ribonucleosides into nascent RNA molecules by adding a modified nucleoside, such as 4-thiouridine (4sU), to the cell media. At low concentrations 4sU is not reported to affect mRNA processing or protein synthesis; however, detection of protein-RNA interactions is limited to newly synthesized RNAs with recognition sites that contain the supplied nucleoside. Additionally, this technique is restricted to cell culture applications and cannot be used to investigate RBPs in whole tissue. Direct comparisons between HITS-CLIP and PAR-CLIP show highly reproducible binding sites and similar transcriptomic landscapes [29].

iCLIP

Individual nucleotide resolution CLIP (iCLIP) [30] is similar to conventional HITS-CLIP with notable improvements. This includes circularization ligation of single strand cDNA instead of a 5' RNA linker ligation step. Not only is this reaction more efficient than linker ligation, it also allows for the capture of truncated reverse transcription products. After protein degradation, several

amino acids remain covalently attached to the RNA molecule at the crosslink site [31] which cause frequent premature termination of reverse transcription at this position. The vast majority of these sequences are lost in traditional HITS-CLIP as PCR amplification requires reverse transcription read-through into the 5' adaptor sequence [32]. In iCLIP, the circularization step is performed after reverse transcription, and captures both full-length and truncated products. The site of cross-linking is predicted to correspond to the nucleotide immediately upstream of the 5' end of the sequenced CLIP tag, providing a means to identify the direct site of protein-RNA interaction with single nucleotide resolution. While iCLIP improves upon the traditional HITS-CLIP protocol with greater library complexity and superior binding site resolution, the technique remains time consuming and challenging.

eCLIP

The complexity of traditional CLIP protocols has proven to be a significant deterrent to large scale profiling efforts [33]. Enhanced CLIP (eCLIP) improves iCLIP with a reported ~1,000 fold increase in preamplification products [34] and reduced library preparation time, while maintaining single nucleotide resolution. These results are achieved by replacing the circularization step after reverse transcription with a more efficient ssDNA linker ligation. The ssDNA adapter contains a unique bar code to identify PCR duplicates generated during PCR amplification. Notably, eCLIP reduces the percentage of CLIP reads discarded due to PCR duplicates by ~60%, producing libraries with greater complexity even in cases with limiting samples or low abundance RBPs. Additional optimization and modifications shorten the protocol from almost two weeks to approximately 4 days. These advances have led to a significant expansion in the number of RBPs that have been interrogated by CLIP, with the datasets publicly available at www.encodeproject.org. Additional resources available to researchers include databases of experimentally validated antibodies [35]. As the RBP of interest is immunoprecipitated using mono- or poly-clonal antibodies, the quality and specificity of the antibody is critical to the success of the experiment. Alternatively, a fusion protein with an exogenous tag can be used [36,37], however validation must be performed to verify the fusion protein remains functional and expression is comparable.

Considerations for interpreting protein-RNA interactions from global analyses

The staggering amount of descriptive information generated in the modern CLIP era has led to significant

advances in our understanding of RNA regulation. However, with comprehensive, global interaction maps comes the need to distinguish biologically functional from non-functional opportunistic interactions. Thus, integration of CLIP maps with ancillary data sets and experimental validation must be performed. Knockdown or knockout of the RBP of interest, followed by RNA analyses can determine which RNA and regulatory events are sensitive to the deletion of the RBP in question. A caveat of these approaches is the potential for global remodeling of protein-RNA interactions, and functional compensation by other RBPs. Multiple strategies can be used to further interrogate mechanism and confirm functionality of specific RBP binding sites defined by CLIP, such as the mutation of protein binding sites. Alternatively, artificial tethering assays using MS2 hairpins or RNA guided Cas9 [38] can bring an RBP to an atypical target. However, artificial manipulation of any system may have unanticipated outcomes. An additional consideration is the inefficiency of 254 nm UV cross-linking and intrinsic crosslinking biases [39], whereby pyrimidines crosslink more efficiently than purines, thus potentially influencing capture of specific transcripts based on their nucleotide composition.

COMPUTATIONAL TOOLS

The emergence of CLIP has prompted a new generation of bioinformatic tools and platforms developed to investigate protein-RNA regulation. At a fundamental level, analysis of deep sequencing data involves steps to decipher between signal and noise. General strategies include analysis of uniquely mapped tags, removal of potential PCR duplicates, and identification of biologically reproducible interactions. With the widespread adoption and use of CLIP methods, a large number of computational approaches (discussed below) have been developed to facilitate data analysis to gain insights into the molecular function(s) of the RBP of interest.

Peak calling tools

Since the advent of CLIP methods, an array of software tools has been developed to distinguish RBP binding sites from the raw data. Such tools include CLIPSeqTools [40], ASPeak [41], CLIPZ [42], PIPE-CLIP [43], Piranha [44], MiClip [45], Pyclip [46], CLIPper [47], PYCRAC [48] and PARalyzer [49]. Some of these algorithms can be applied to data from different CLIP platforms and include commands that pre-process input files to remove adapters, mask repeats, align reads to the genome, remove PCR duplicates, identify chemical modifications, and evoke binomial models to identify enriched peaks. The majority

of these packages provide outputs that are compatible with any genome browser for easy visualization. Fortunately, given the many bioinformatic tools that have spawned from CLIP technologies, users should be able to find a tool to interrogate data generated from HITS-, PAR-, i-, or e-CLIP, and is also suitable for different machine platforms (linux, mac, PC).

Binding site predictions

High-resolution RBP-RNA interactome maps generated by CLIP methodologies present an opportunity to identify RBP binding signatures. Multiple motif discovery tools, such as MEME [50], compseq (EMBOSS utilities), HOMER [51], mCarts [52], and ChIPMunk [53] can be applied to determine RBP binding motifs. Additionally, software specifically designed to interrogate CLIP data has been developed. In 2011, a computational framework, known as CIMS, was developed to identify protein-RNA interactions at single nucleotide resolution [54]. The foundation of this platform takes advantage of the residual amino-acid-RNA adduct remaining after UV crosslinking and protein degradation, which can obstruct the reverse transcriptase or result in a mutation in the cDNA at the crosslink site. Another motif discovery algorithm, called Zagros, was modeled to exploit sequence, secondary structure, and technology-specific crosslinking events from CLIP data [55]. These techniques can be readily utilized to characterize sequences within CLIP peaks.

Overcoming CLIP biases

As described above, there are several resources used to annotate RBP binding sites based on CLIP data, however many of these methods cannot distinguish binding of a single RBP from the binding of a protein complex; nor can they account for preferential detection of uridine-rich sequences [32]. To circumvent these biases, CLIP data can be integrated with other protein-RNA binding methods. One such method is RNACompete, which systematically analyzes RNA binding specificities of RBPs in a rapid and low-cost fashion [56]. In brief, RNACompete uses *in vitro* RNA-protein binding, followed by high-throughput analyses to identify RNA binding motifs. Notably, this approach provides binding preferences for a large number of sequences with only one binding reaction, and is absent of any possible cross-linking modifications. On the other hand, *in vitro* approaches do not account for the competitive and cooperative binding of other proteins in the cell. Furthermore, the influence of secondary structure on binding is complicated by the number of probes required for the assay.

Following RNACompete, RNA Bind-n-Seq (RBNS) [57] was developed to quantitatively measure RBP binding affinities to a range of bound RNAs. The high-throughput nature of RBNS is more robust than previous quantitative assays, such as electrophoretic mobility shift assays and surface plasmon resonance, which only provide low-throughput K_d values. Evolved from previous protein-DNA binding protocols, RBNS optimized RBP concentration relative to binding affinities, and advanced an analytical framework to estimate the effects of secondary structure on protein binding [57]. Importantly, Lambert *et al.* demonstrated that motifs enriched by CLIP, but not by RBNS, are not associated with regulatory activity, suggesting that RBNS is a valuable supplement in creating RBP profiles when using CLIP [57].

DELINEATING FUNCTIONS OF PROTEIN-RNA INTERACTIONS

RBP binding profiles and splicing maps

Our understanding of splicing regulatory networks advanced in the late 1990s, when SELEX-based experiments were adopted to investigate the position of SR proteins relative to spliced exons [58]. Liu *et al.* showed that individual SR proteins have a high affinity to bind exonic splicing enhancers, thereby promoting the inclusion of alternative exons [58]. Since then, the development of splicing microarrays [59–62], RIP-chip [14], and CLIP [26,63–72] technologies have allowed metagene analyses and the construction of “splicing maps” to reveal position dependent effects of RBPs on splicing function. Generally, RBPs bound upstream of an alternative exon repress splicing of the exon, while RBPs bound downstream promote exon inclusion. More specifically, NOVA, hnRNP C, hnRNP L, hnRNP H, PTBP1, PTBP2, and MBNL1 have been shown to silence exon inclusion by binding at positions close to the branch site or the 3' splice site, thus interfering with spliceosome machinery. In contrast, NOVA, RBFOX, hnRNP L, and TIA proteins bind downstream to promote alternative exon inclusion, potentially by RNA looping. Intron versus exon binding also dictates splicing behavior. In general, hnRNPs inhibit splicing when bound to an exon, yet can have either positive or negative regulation when bound to intronic sites [73,74].

While many splicing maps have been generated for individual RBPs, global splicing networks of combinatorial RBPs remain poorly understood. In 2010, the binding patterns of NOVA and its subsequent contribution on splicing were characterized using RBP-RNA transcriptome-wide maps [75]. Not only did Zhang *et al.*

show conserved NOVA binding sites, they also discovered an enrichment of the RBFOX (UGCAUG) binding element, suggesting combinatorial regulation of splicing by NOVA and RBFOX. Shortly after, additional evidence supporting widespread combinatorial regulation of alternative splicing was demonstrated between ESRPs and RBFOX2/RBM47 during the epithelial-mesenchymal transition [76,77]. More recently, Damianov *et al.* demonstrated that nuclear RBFOX proteins are bound within a large assembly of splicing regulators (LASR), which include hnRNP M, hnRNP H, hnNP C, Matrin3, NF110/NFAR-2, NF45, and DDX5 [70]. This study implicates that protein complexes may function in controlling splicing networks, and identification of such complexes may be required to fully decipher regulatory codes.

Position-dependent regulation of alternative polyadenylation

In addition to elucidating general regulatory rules associated with known functions of RBPs, high-throughput sequencing after CLIP can uncover new functions for RBPs. The unanticipated identification of NOVA CLIP tags overlapping or near polyadenylation (pA) sites led to a hypothesis that this protein could have a second, splicing-independent function in mRNA processing in the brain [26]. Indeed, subsequent experiments indicated that NOVA impacts alternative polyadenylation choices to promote expression of mRNAs with brain-specific 3' UTRs. In 2014, Batra *et al.* also utilized HITS-CLIP data alongside minigene reporter assays to show that direct binding of MBNL to target RNAs regulates pA site selection [78]. Importantly, polyadenylation regulatory maps revealed that pA sites are repressed when MBNL binds in close proximity to the core 3' end processing region, while more distal binding (upstream) activates pA selection. Similarly, CLIP-seq technology was used to investigate the functional impact of FUS binding clusters near alternative pA sites. Extensive transcriptome-wide mapping demonstrated a positional dependence of FUS binding to RNA for activation of alternative polyadenylation and suggested the involvement of RNA polymerase accumulation in this regulation [79]. Two new computational platforms, expressRNA and RNAmotifs2, were developed to advance the position-dependent principles of pre-mRNA processing [80]. These platforms demonstrated that TDP-43 most often binds the proximal same-exon pA site to repress site usage, but also has enriched binding downstream of activated pA sites (consistent with TDP-43 splicing regulation) [80]. While RNA binding maps are fundamental to understanding the role of RBPs in alternative polyadenylation, a need remains for an

integrative platform, which accounts for structure and combinatorial regulation with other *trans*-acting factors, including the cleavage and polyadenylation machinery.

Insights into cytoplasmic mRNA regulation from protein-RNA interaction maps

CLIP methods have also been applied to the study of cytoplasmic mRNA metabolism. In conjunction with RNA-seq analysis and ribosome profiling, CLIP analysis of mammalian UPF1 has been used to identify direct targets of the nonsense-mediated decay (NMD) [81]. Opposite of what was expected, Hurt *et al.* identified reproducible UPF1 binding to 3' UTRs with a density 10 times greater than that seen in the coding region [81]. While CLIP analysis did not suggest a clear binding motif, bound sequences were enriched for guanosine residues and secondary structure. Further characterization found UPF1 targets to have longer 3' UTRs and increased translational efficiency. Another study identified that MBNL1 and CUGBP1 preferentially bind 3' UTRs to facilitate mRNA decay in C2C12 cells [82]. Although NMD has been studied for nearly forty years, mechanisms of NMD substrate recognition remain unclear.

CLIP methods have also been used to study gene regulation by microRNAs (miRNAs), which post-transcriptionally control gene expression by binding to 3' UTRs. Functional miRNAs are loaded on AGO complexes to bind target RNAs, leading to silencing by translation repression or nucleolytic turnover. While bioinformatic predictions can identify miRNA binding sites, even the most stringent analyses yield high rates of false positives. Application of HITS-CLIP to AGO *in vivo* allows functional miRNA target sites to be mapped [83]. Since miRNA-mRNA base pairing occurs within short seed regions, size selection by SDS-PAGE after immunoprecipitation is especially crucial. Moore *et al.* have published a detailed protocol for AGO HITS-CLIP library construction, as well as downstream computational analyses [83] that can be applied to cells or tissues of interest. Nearly five years after AGO HITS-CLIP method was published, CLASH (crosslinking, ligation, and sequencing of hybrids) [84] was developed as a technique to capture and map miRNA-RNA duplexes associated with human AGO1. Using this approach, Helwak *et al.* annotated several miRNA “seed” binding sites, a substantial number of interactions that do not involve contacts within the seed region, and enriched motifs within these sites. CLASH technology and the nearly forty AGO HITS-CLIP datasets published inspired the development of new algorithms to resolve highly expressed miRNAs associated with AGO CLIP peaks, including microMUMMIE [85], DIANA-microT-CDS [86], STarMir [87], MIRZA [88], CLIPZ [42], and

chimiRic [89]. These computational models use cross-linking mutations; energy based duplex predictions; protocol specific sequence signals; and/or common AGO sequence preferences to assign the most likely canonical and noncanonical seed. Together, these experimental and bioinformatic technologies improve miRNA target mapping transcriptome-wide, and demonstrate that miRNA-mRNA targeting is much more widespread than anticipated.

FUTURE OF TRANSCRIPTOME-WIDE PROTEIN-RNA MAPPING TO INVESTIGATE REGULATION OF GENE EXPRESSION

The ability to map protein-RNA contacts in a transcriptome-wide manner has revolutionized our understanding of RNA regulatory mechanisms and functions. Future studies of RBP regulation will significantly benefit from the advancement of tools to map RNA secondary structure. Interestingly, it has been postulated that not all regulatory factors recognize consensus sequences [90], but instead may be recruited to the RNA via secondary structure or other co-factors. Moving forward, integration of RNA structure predictions with protein-protein and protein-RNA interaction tools will give rise to a more complete characterization of RNA regulation by RBPs.

The experimental methods described in this review are fundamentally based on the purification of RNA bound to a specific protein of interest. These protein-focused techniques provide detailed positional information and may illuminate networks of RNAs under coordinate regulation. However, they do not yield data regarding the composition of the unique complex of proteins bound to each mRNA molecule. While not discussed in depth in this review, a variety of methods make use of the RNA molecule as bait to capture and identify bound proteins using mass spectrometry analysis. Such methodologies include CHART [91], ChIRP [92], and RAP-MS [93]. Additionally, global interactome capture of the entire mRNA-associated proteome has identified many previously unknown RBPs [94,95]. Coupling CLIP methods with increasingly powerful biochemical and genetic tools [96] has the potential to reveal RNA regulatory networks that underlie cell type-specific functions and developmental programs.

In summary, unraveling the complexity of RNA regulatory events and gene regulation is key to understanding gene expression and disease pathogenesis. In addition to elucidating general regulatory rules associated with known functions of RBPs, high-throughput sequencing after CLIP can uncover new functions for RBPs and shed light on global RNA networks. In the future, CLIP

technologies can be adapted to address questions relating to countless aspects of RNA metabolism.

ACKNOWLEDGEMENTS

This work was supported by funds from the National Institutes of Health to LLZ (T32 GM08056) and DDL (R01 GM107331), USA.

COMPLIANCE WITH ETHICS GUIDELINES

The authors Molly M. Hannigan, Leah L. Zagore and Donny D. Licatalosi declare that they have no conflict of interests.

This article is a review article and does not contain any studies with human or animal subjects performed by any of the authors.

REFERENCES

- Linder, B., Fischer, U. and Gehring, N. H. (2015) mRNA metabolism and neuronal disease. *FEBS Lett.*, 589, 1598–1606
- Brais, B., Bouchard, J. P., Xie, Y. G., Rochefort, D. L., Chrétien, N., Tomé, F. M., Lafrentère, R. G., Rommens, J. M., Uyama, E., Nohira, O., *et al.* (1998) Short GCG expansions in the *PABP2* gene cause oculopharyngeal muscular dystrophy. *Nat. Genet.*, 18, 164–167
- Sreedharan, J., Blair, I. P., Tripathi, V. B., Hu, X., Vance, C., Rogelj, B., Ackerley, S., Durnall, J. C., Williams, K. L., Buratti, E., *et al.* (2008) TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science*, 319, 1668–1672
- Bentley, D. L. (2014) Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.*, 15, 163–175
- Baltz, A. G., Munschauer, M., Schwanhäusser, B., Vasilic, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., *et al.* (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell*, 46, 674–690
- Gerstberger, S., Hafner, M. and Tuschl, T. (2014) A census of human RNA-binding proteins. *Nat. Rev. Genet.*, 15, 829–845
- Dreyfuss, G., Kim, V. N. and Kataoka, N. (2002) Messenger-RNA-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.*, 3, 195–205
- Setzer, D. R. (1999) Measuring equilibrium and kinetic constants using gel retardation assays. *Methods Mol. Biol.*, 118, 115–128
- Katsamba, P. S., Park, S. and Laird-Offringa, I. A. (2002) Kinetic studies of RNA-protein interactions using surface plasmon resonance. *Methods*, 26, 95–104
- Hook, B., Bernstein, D., Zhang, B. and Wickens, M. (2005) RNA-protein interactions in the yeast three-hybrid system: affinity, sensitivity, and enhanced library screening. *RNA*, 11, 227–233
- Ellington, A. D. and Szostak, J. W. (1990) *In vitro* selection of RNA molecules that bind specific ligands. *Nature*, 346, 818–822
- Tuerk, C. and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, 249, 505–510
- Tenenbaum, S. A., Carson, C. C., Lager, P. J. and Keene, J. D. (2000) Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. *Proc. Natl. Acad. Sci. USA*, 97, 14085–14090
- Keene, J. D., Komisarow, J. M. and Friedersdorf, M. B. (2006) RIP-chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat. Protoc.*, 1, 302–307
- Mili, S. and Steitz, J. A. (2004) Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA*, 10, 1692–1694
- Dreyfuss, G., Adam, S. A. and Choi, Y. D. (1984) Physical change in cytoplasmic messenger ribonucleoproteins in cells treated with inhibitors of mRNA transcription. *Mol. Cell. Biol.*, 4, 415–423
- 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
- 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
- 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
- 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
- Singh, G., Ricci, E. P. and Moore, M. J. (2014) RIPiT-Seq: a high-throughput approach for footprinting RNA:protein complexes. *Methods*, 65, 320–332
- 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
- Kargapolova, Y., Levin, M., Lackner, K. and Danckwardt, S. (2017) sCLIP—an integrated platform to study RNA-protein interactomes in biomedical research: identification of CSTF2tau in alternative processing of small nuclear RNAs. *Nucleic Acids Res.*, 45, 6074–6086
- Rosenberg, M., Blum, R., Kesner, B., Maier, V. K., Szanto, A. and Lee, J. T. (2017) Denaturing CLIP, dCLIP, pipeline identifies discrete RNA footprints on chromatin-associated proteins and reveals that CBX7 targets 3' UTRs to regulate mRNA expression. *Cell Syst.*, 5, 368–385
- Brugiolo, M., Botti, V., Liu, N., Müller-McNicoll, M. and Neugebauer, K. M. (2017) Fractionation iCLIP detects persistent SR protein binding to conserved, retained introns in chromatin, nucleoplasm and cytoplasm. *Nucleic Acids Res.*, 45, 10452–10465
- Licatalosi, D. D., Mele, A., Fak, J. J., Ule, J., Kayikci, M., Chi, S. W., Clark, T. A., Schweitzer, A. C., Blume, J. E., Wang, X., *et al.* (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*, 456, 464–469
- Darnell, R. B. (2010) HITS-CLIP: panoramic views of protein-RNA regulation in living cells. *Wiley Interdiscip. Rev. RNA*, 1, 266–286

28. Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M. Jr, 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
29. Kishore, S., Jaskiewicz, L., Burger, L., Hausser, J., Khorshid, M. and Zavolan, M. (2011) A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nat. Methods*, 8, 559–564
30. 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
31. Urlaub, H., Hartmuth, K. and Lührmann, R. (2002) A two-tracked approach to analyze RNA-protein crosslinking sites in native, nonlabeled small nuclear ribonucleoprotein particles. *Methods*, 26, 170–181
32. Sugimoto, Y., König, J., Hussain, S., Zupan, B., Curk, T., Frye, M. and Ule, J. (2012) Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. *Genome Biol.*, 13, R67
33. Marchese, D., de Groot, N. S., Lorenzo Gotor, N., Livi, C. M. and Tartaglia, G. G. (2016) Advances in the characterization of RNA-binding proteins. *Wiley Interdiscip. Rev. RNA*, 7, 793–810
34. 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
35. Björling, E. and Uhlén, M. (2008) Antibodypedia, a portal for sharing antibody and antigen validation data. *Mol. Cell. Proteomics*, 7, 2028–2037
36. de Boer, E., Rodriguez, P., Bonte, E., Krijgsveld, J., Katsantoni, E., Heck, A., Grosveld, F. and Strouboulis, J. (2003) Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci. USA*, 100, 7480–7485
37. Gerace, E. and Moazed, D. (2015) Affinity purification of protein complexes using TAP tags. *Methods Enzymol.*, 559, 37–52
38. Nelles, D. A., Fang, M. Y., O’Connell, M. R., Xu, J. L., Markmiller, S. J., Doudna, J. A. and Yeo, G. W. (2016) Programmable RNA tracking in live cells with CRISPR/Cas9. *Cell*, 165, 488–496
39. Wheeler, E. C., Van Nostrand, E. L. and Yeo, G. W. (2018) Advances and challenges in the detection of transcriptome-wide protein-RNA interactions. *Wiley Interdiscip. Rev. RNA*, 9, e1436
40. Maragkakis, M., Alexiou, P., Nakaya, T. and Mourelatos, Z. (2016) CLIPSeqTools — a novel bioinformatics CLIP-seq analysis suite. *RNA*, 22, 1–9
41. Kucukural, A., Özadam, H., Singh, G., Moore, M. J. and Cenik, C. (2013) ASPeak: an abundance sensitive peak detection algorithm for RIP-seq. *Bioinformatics*, 29, 2485–2486
42. Khorshid, M., Rodak, C. and Zavolan, M. (2011) CLIPZ: a database and analysis environment for experimentally determined binding sites of RNA-binding proteins. *Nucleic Acids Res.*, 39, D245–D252
43. Chen, B., Yun, J., Kim, M. S., Mendell, J. T. and Xie, Y. (2014) PIPE-CLIP: a comprehensive online tool for CLIP-seq data analysis. *Genome Biol.*, 15, R18
44. 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. and Smith, A. D. (2012) Site identification in high-throughput RNA-protein interaction data. *Bioinformatics*, 28, 3013–3020
45. Wang, T., Chen, B., Kim, M., Xie, Y. and Xiao, G. (2014) A model-based approach to identify binding sites in CLIP-seq data. *PLoS One*, 9, e93248
46. Althammer, S., González-Vallinas, J., Ballaré, C., Beato, M. and Eyra, E. (2011) Pyicos: a versatile toolkit for the analysis of high-throughput sequencing data. *Bioinformatics*, 27, 3333–3340
47. 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
48. Webb, S., Hector, R. D., Kudla, G. and Granneman, S. (2014) PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. *Genome Biol.*, 15, R8
49. 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
50. Bailey, T. L., Johnson, J., Grant, C. E. and Noble, W. S. (2015) The MEME Suite. *Nucleic Acids Res.*, 43, W39–W49
51. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H. and Glass, C. K. (2010) Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, 38, 576–589
52. Weyn-Vanhenenryck, S. M. and Zhang, C. (2016) mCarts: genome-wide prediction of clustered sequence motifs as binding sites for RNA-binding proteins. *Methods Mol. Biol.*, 1421, 215–226
53. Kulakovskiy, I. V., Boeva, V. A., Favorov, A. V. and Makeev, V. J. (2010) Deep and wide digging for binding motifs in ChIP-seq data. *Bioinformatics*, 26, 2622–2623
54. Zhang, C. and Darnell, R. B. (2011) Mapping *in vivo* protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat. Biotechnol.*, 29, 607–614
55. Bahrami-Samani, E., Penalva, L. O., Smith, A. D. and Uren, P. J. (2015) Leveraging cross-link modification events in CLIP-seq for motif discovery. *Nucleic Acids Res.*, 43, 95–103
56. Ray, D., Kazan, H., Chan, E. T., Castillo, L. P., Chaudhry, S., Talukder, S., Blencowe, B. J., Morris, Q. and Hughes, T. R. (2009) Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. *Nat. Biotechnol.*, 27, 667–670
57. Lambert, N., Robertson, A., Jangi, M., McGeary, S., Sharp, P. A. and Burge, C. B. (2014) RNA Bind-n-Seq: quantitative assessment

- of the sequence and structural binding specificity of RNA binding proteins. *Mol. Cell*, 54, 887–900
58. Liu, H. X., Zhang, M. and Krainer, A. R. (1998) Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes Dev.*, 12, 1998–2012
 59. Johnson, J. M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P. M., Armour, C. D., Santos, R., Schadt, E. E., Stoughton, R. and Shoemaker, D. D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, 302, 2141–2144
 60. Ule, J., Ule, A., Spencer, J., Williams, A., Hu, J. S., Cline, M., Wang, H., Clark, T., Fraser, C., Ruggiu, M., *et al.* (2005) Nova regulates brain-specific splicing to shape the synapse. *Nat. Genet.*, 37, 844–852
 61. Sugnet, C. W., Srinivasan, K., Clark, T. A., O'Brien, G., Cline, M. S., Wang, H., Williams, A., Kulp, D., Blume, J. E., Haussler, D., *et al.* (2006) Unusual intron conservation near tissue-regulated exons found by splicing microarrays. *PLoS Comput. Biol.*, 2, e4
 62. Boutz, P. L., Stoilov, P., Li, Q., Lin, C. H., Chawla, G., Ostrow, K., Shiue, L., Ares, M. Jr and Black, D. L. (2007) A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev.*, 21, 1636–1652
 63. Licatalosi, D. D., Yano, M., Fak, J. J., Mele, A., Grabinski, S. E., Zhang, C. and Darnell, R. B. (2012) Ptp2 represses adult-specific splicing to regulate the generation of neuronal precursors in the embryonic brain. *Genes Dev.*, 26, 1626–1642
 64. Hannigan, M. M., Zagore, L. L. and Licatalosi, D. D. (2017) Ptp2 controls an alternative splicing network required for cell communication during spermatogenesis. *Cell Reports*, 19, 2598–2612
 65. Wang, E. T., Cody, N. A., Jog, S., Biancolella, M., Wang, T. T., Treacy, D. J., Luo, S., Schroth, G. P., Housman, D. E., Reddy, S., *et al.* (2012) Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell*, 150, 710–724
 66. Charizanis, K., Lee, K. Y., Batra, R., Goodwin, M., Zhang, C., Yuan, Y., Shiue, L., Cline, M., Scotti, M. M., Xia, G., *et al.* (2012) Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. *Neuron*, 75, 437–450
 67. Änkö, M. L., Müller-McNicoll, M., Brandl, H., Curk, T., Gorup, C., Henry, I., Ule, J. and Neugebauer, K. M. (2012) The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes. *Genome Biol.*, 13, R17
 68. Pandit, S., Zhou, Y., Shiue, L., Coutinho-Mansfield, G., Li, H., Qiu, J., Huang, J., Yeo, G. W., Ares, M. Jr and Fu, X. D. (2013) Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. *Mol. Cell*, 50, 223–235
 69. 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 Reports*, 6, 1139–1152
 70. Damianov, A., Ying, Y., Lin, C. H., Lee, J. A., Tran, D., Vashisht, A. A., Bahrami-Samani, E., Xing, Y., Martin, K. C., Wohlschlegel, J. A., *et al.* (2016) Rbfox proteins regulate splicing as part of a large multiprotein complex LASR. *Cell*, 165, 606–619
 71. Wang, E. T., Ward, A. J., Cherone, J. M., Giudice, J., Wang, T. T., Treacy, D. J., Lambert, N. J., Freese, P., Saxena, T., Cooper, T. A., *et al.* (2015) Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res.*, 25, 858–871
 72. Kapeli, K., Pratt, G. A., Vu, A. Q., Hutt, K. R., Martinez, F. J., Sundararaman, B., Batra, R., Freese, P., Lambert, N. J., Huelga, S. C., *et al.* (2016) Distinct and shared functions of ALS-associated proteins TDP-43, FUS and TAF15 revealed by multisystem analyses. *Nat. Commun.*, 7, 12143
 73. Rothrock, C. R., House, A. E. and Lynch, K. W. (2005) HnRNP L represses exon splicing via a regulated exonic splicing silencer. *EMBO J.*, 24, 2792–2802
 74. Motta-Mena, L. B., Heyd, F. and Lynch, K. W. (2010) Context-dependent regulatory mechanism of the splicing factor hnRNP L. *Mol. Cell*, 37, 223–234
 75. Zhang, C., Frias, M. A., Mele, A., Ruggiu, M., Eom, T., Marney, C. B., Wang, H., Licatalosi, D. D., Fak, J. J. and Darnell, R. B. (2010) Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. *Science*, 329, 439–443
 76. Yang, Y., Park, J. W., Bebee, T. W., Warzecha, C. C., Guo, Y., Shang, X., Xing, Y. and Carstens, R. P. (2016) Determination of a comprehensive alternative splicing regulatory network and combinatorial regulation by key factors during the epithelial-to-mesenchymal transition. *Mol. Cell Biol.*, 36, 1704–1719
 77. Dittmar, K. A., Jiang, P., Park, J. W., Amirkian, K., Wan, J., Shen, S., Xing, Y. and Carstens, R. P. (2012) Genome-wide determination of a broad ESRP-regulated posttranscriptional network by high-throughput sequencing. *Mol. Cell Biol.*, 32, 1468–1482
 78. Batra, R., Charizanis, K., Manchanda, M., Mohan, A., Li, M., Finn, D. J., Goodwin, M., Zhang, C., Sobczak, K., Thornton, C. A., *et al.* (2014) Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol. Cell*, 56, 311–322
 79. Masuda, A., Takeda, J., Okuno, T., Okamoto, T., Ohkawara, B., Ito, M., Ishigaki, S., Sobue, G. and Ohno, K. (2015) Position-specific binding of FUS to nascent RNA regulates mRNA length. *Genes Dev.*, 29, 1045–1057
 80. Rot, G., Wang, Z., Huppertz, I., Modic, M., Lenče, T., Hallegger, M., Haberman, N., Curk, T., von Mering, C. and Ule, J. (2017) High-resolution RNA maps suggest common principles of splicing and polyadenylation regulation by TDP-43. *Cell Reports*, 19, 1056–1067
 81. Hurt, J. A., Robertson, A. D. and Burge, C. B. (2013) Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. *Genome Res.*, 23, 1636–1650
 82. Masuda, A., Andersen, H. S., Doktor, T. K., Okamoto, T., Ito, M., Andresen, B. S. and Ohno, K. (2012) CUGBP1 and MBNL1 preferentially bind to 3' UTRs and facilitate mRNA decay. *Sci. Rep.*, 2, 209
 83. Moore, M. J., Zhang, C., Gantman, E. C., Mele, A., Darnell, J. C. and Darnell, R. B. (2014) Mapping Argonaute and conventional

- RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. *Nat. Protoc.*, 9, 263–293
84. 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
85. Majoros, W. H., Lekprasert, P., Mukherjee, N., Skalsky, R. L., Corcoran, D. L., Cullen, B. R. and Ohler, U. (2013) MicroRNA target site identification by integrating sequence and binding information. *Nat. Methods*, 10, 630–633
86. Reczko, M., Maragkakis, M., Alexiou, P., Grosse, I. and Hatzigeorgiou, A. G. (2012) Functional microRNA targets in protein coding sequences. *Bioinformatics*, 28, 771–776
87. Liu, C., Mallick, B., Long, D., Rennie, W. A., Wolenc, A., Carmack, C. S. and Ding, Y. (2013) CLIP-based prediction of mammalian microRNA binding sites. *Nucleic Acids Res.*, 41, e138
88. Khorshid, M., Hausser, J., Zavolan, M. and van Nimwegen, E. (2013) A biophysical miRNA-mRNA interaction model infers canonical and noncanonical targets. *Nat. Methods*, 10, 253–255
89. Lu, Y. and Leslie, C. S. (2016) Learning to predict miRNA-mRNA interactions from AGO CLIP sequencing and CLASH data. *PLoS Comput. Biol.*, 12, e1005026
90. Zhu, Y., Wang, X., Forouzmand, E., Jeong, J., Qiao, F., Sowd, G. A., Engelman, A. N., Xie, X., Hertel, K. J. and Shi, Y. (2018) Molecular mechanisms for CFIm-mediated regulation of mRNA alternative polyadenylation. *Mol. Cell*, 69, 62–74
91. Simon, M. D. (2013) Capture hybridization analysis of RNA targets (CHART). *Curr. Protoc. Mol. Biol.* 101, 21.25.1–21.25.16
92. Chu, C., Quinn, J. and Chang, H. Y. (2012) Chromatin isolation by RNA purification (ChIRP). *J. Vis. Exp.*, 61, 3912
93. McHugh, C. A., Chen, C. K., Chow, A., Surka, C. F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., *et al.* (2015) The *Xist* lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, 521, 232–236
94. Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L. M., Krijgsveld, J. and Hentze, M. W. (2016) Comprehensive identification of RNA-binding proteins by RNA interactome capture. *Methods Mol. Biol.*, 1358, 131–139
95. Bao, X., Guo, X., Yin, M., Tariq, M., Lai, Y., Kanwal, S., Zhou, J., Li, N., Lv, Y., Pulido-Quetglas, C., *et al.* (2018) Capturing the interactome of newly transcribed RNA. *Nat. Methods*, 15, 213–220
96. Hwang, H. W., Park, C. Y., Goodarzi, H., Fak, J. J., Mele, A., Moore, M. J., Saito, Y. and Darnell, R. B. (2016) PAPERCLIP identifies microRNA targets and a role of CstF64/64tau in promoting non-canonical poly(A) site usage. *Cell Reports*, 15, 423–435