The Future of CLIP

Jernej Ule[1,2], Hun-Way Hwang[3] and Robert B. Darnell[4,5,*]

[1] The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK
[2] Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, WC1N 3BG, UK
[3] Department of Pathology, University of Pittsburgh, School of Medicine, 3550 Terrace Street, Pittsburgh, PA 15261, USA
[4] Laboratory of Molecular Neuro-Oncology
[5] Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, NY 10065 USA

*Corresponding author: Robert B. Darnell, M.D., Ph.D.
Laboratory of Molecular Neuro-Oncology
The Rockefeller University
Howard Hughes Medical Institute
Box 226
1230 York Avenue
New York, NY 10021
Phone: (212) 327-7474
e-mail: darnelr@rockefeller.edu

Table of Contents

Abstract

To understand the assembly and functional outcomes of protein-RNA regulation, it is crucial to precisely identify the positions of such interactions. Crosslinking and immunoprecipitation (CLIP) serves this purpose by exploiting covalent protein-RNA crosslinking and RNA fragmentation, along with a series of stringent purification and quality control steps to prepare cDNA libraries for sequencing. Here we describe the core steps of CLIP, its primary variations and the approaches to data analysis. We present the application of CLIP to studies of specific cell types in genetically engineered mice, and discuss the mechanistic and physiologic insights that have already been gained from studies employing CLIP. We conclude by discussing the future opportunities for CLIP, including studies of human postmortem tissues from disease patients and controls, RNA epigenetic modifications, and RNA structure. These and other applications of CLIP will continue to unravel fundamental gene regulatory mechanisms, while providing important biologic and clinically relevant insights.

Introduction

A crucial step in understanding RNA-related gene regulation and its relationship to disease is identifying how RNAs are bound and hence regulated by RNA binding proteins (RBPs) in specific cell types and subcellular compartments. Driving our interest in this topic has been the growing list of human neurologic diseases that have RNA dysregulation at their core (Conlon and Manley 2017). The first method developed for this purpose used antibodies against the spliceosomal Sm proteins (lupus autoimmune sera) to identify the small nuclear RNAs, which interact with Sm proteins within the abundant snRNPs (Lerner and Steitz 1979). This method, later referred to as RIP (for RNP/RNA immunoprecipitation), relies on immunoprecipitation (IP) of an RBP under conditions that preserve ribonucleoprotein complexes (RNPs) (Niranjanakumari et al. 2002). However, RIP can suffer from low specificity, partly because it preserves protein-protein interactions and can therefore purify multiple RBPs in complex with their bound RNAs, and partly because RNA-protein complexes can reassociate in vitro (Mili and Steitz 2004).

To understand the assembly of RNPs, a method to identify direct protein-RNA interactions with high specificity was required. Moreover, it was important to also identify the positions on RNAs that are recognized by a specific RBP in order to identify the RNA sequences and structures where regulation takes place. We have developed crosslinking and immunoprecipitation (CLIP) for this purpose by exploiting zero-length covalent protein-RNA crosslinking and RNA fragmentation, along with a series of stringent purification and quality control steps, followed by the preparation of a cDNA library for sequencing (Ule et al. 2003; R. B. Darnell 2010). With CLIP, only RNA fragments that are crosslinked to the specific RBP are isolated under well-optimized conditions, which avoid co-purification of additional RBPs. CLIP data can be validated with the use of appropriate negative controls, such as the use of cells or tissue that lack the RBP of interest or have not been crosslinked (Donny D. Licatalosi and Darnell 2010). CLIP is now generally accepted as a highly reliable method for identifying the positions of endogenous RNA-protein interactions (Conlon and Manley 2017).

Initially, CLIP relied on Sanger sequencing to identify 340 sequences corresponding to RNA interactions of splicing factors Nova1 and Nova2 in mouse brain (Ule et al. 2003). In spite of the limited number of these initial sequences, many were located next to alternative exons that turned out to be regulated by Nova proteins, thus demonstrating the capacity of CLIP to identify functionally important binding sites. Subsequent development of HITS-CLIP (high-throughput sequencing of RNA isolated by CLIP) was then applied to further studies of the

functional interactions of Nova and several other RBPs with neuronal transcripts in the mouse brain (D. D. Licatalosi et al. 2008; R. B. Darnell 2013).

This work established a general strategy to link functional studies with a transcriptome-wide map of protein-RNA interactions. Such maps validated regulatory rules that had been bioinformatically predicted, in which the position of binding of RBPs within primary transcripts largely determines the outcome of RNA regulation (Ule et al. 2006; D. D. Licatalosi et al. 2008; Donny D. Licatalosi and Darnell 2010). These rules are now recognized as a general feature of pre-mRNA processing, applicable to dozens of different RBPs (Witten and Ule 2011). Following these approaches, CLIP has been applied to a large number of different RBPs in a variety of cell types, organs and species.

The CLIP method

A. Covalent protein-RNA crosslinking

B. Cell lysis and RNA fragmentation

C. Purification of protein-RNA complexes

D. Ligation of adapter to the 3' end of RNA

E. Purification of crossinked RNA fragments

+ proteinase K
+SDS/urea
+heat

F1. Preparation of cDNA library (HITS-CLIP)

Ligation of adapter to the 5' end of RNA

RNA: 5'   3'

Reverse transcription,
cDNA amplification and sequencing

readthrough cDNA:

sequencing read

3'   5'

F2. Preparation of cDNA library (iCLIP)

Reverse transcription, and ligation
of adapter to the 3' end of cDNA

truncated cDNA:

3'   5'
UMI
barcode

cDNA amplification and sequencing
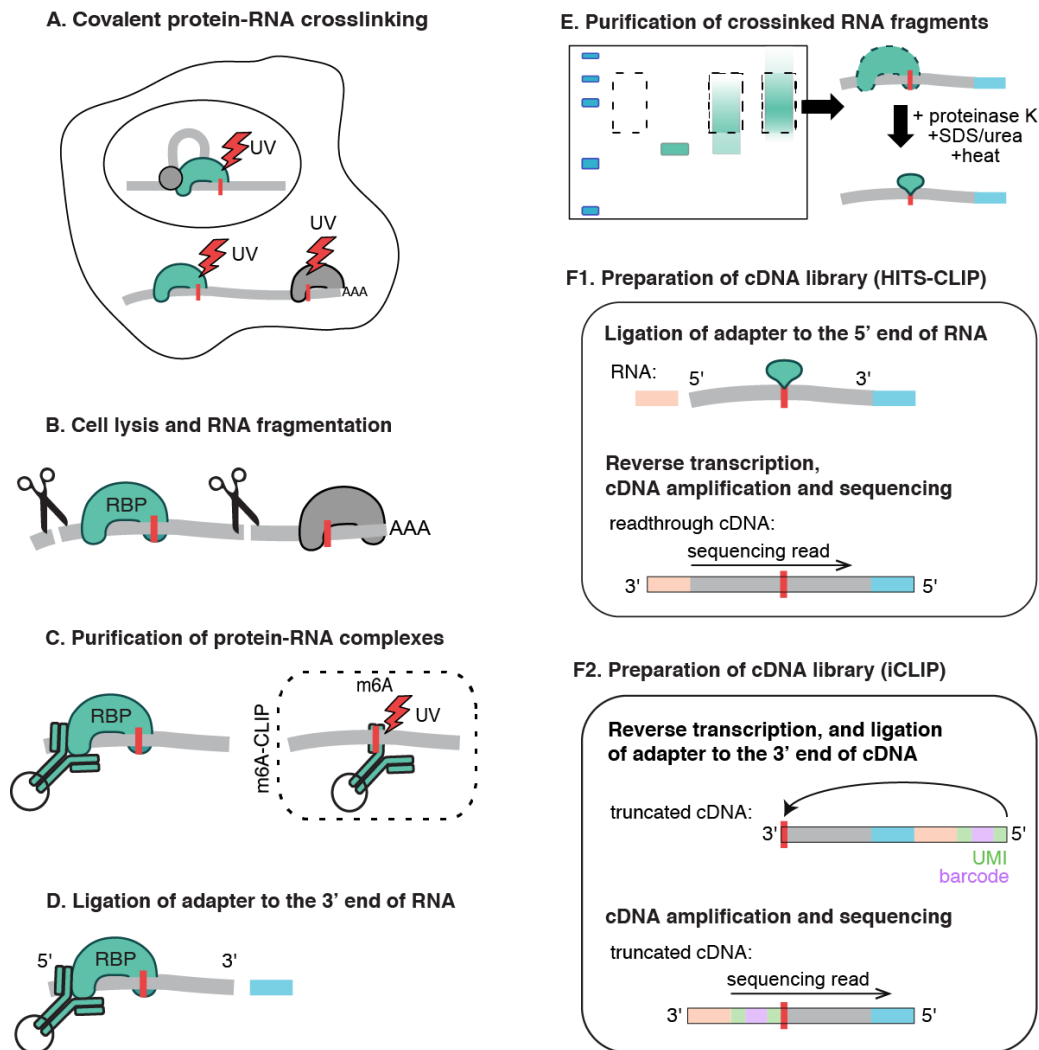
truncated cDNA:

sequencing read

3'   5'

Figure 1: The core steps of CLIP

The steps A-E are common across most variants of CLIP. Methods that identify the sites of RNA methylation crosslink purified RNA to an antibody specific to a methylated base, as in m6A-CLIP, and then proceed from step C onwards. The first option to prepare the cDNA library is as in HITS-CLIP (F1), where the cDNA library is prepared by ligating an adapter to the 5' end of RNA to allow amplification and sequencing of readthrough cDNAs, but not truncated cDNAs. The second option is as in iCLIP (F2), where adapter is ligated to the 3' end of cDNAs, which allows amplification of both readthrough and truncated cDNAs (only a truncated cDNA is shown here, since this type most often represents >90% of the iCLIP cDNA library (Haberman et al. 2017; Sugimoto et al. 2012). The unique molecular identifiers (UMI) and experimental barcodes, which are introduced in iCLIP during reverse transcription, are color coded in green and violet. The RNA fragments and their

corresponding cDNA inserts are in grey. The figure is adapted based on previous publication by Lee and Ule (Lee and Ule 2018).

The fundamentals

Many variant protocols of CLIP have been derived during the last decade. Their detailed variations have been reviewed elsewhere (Lee and Ule 2018; R. B. Darnell 2010), and therefore we will focus here only on the primary steps that are common to most protocols (Figure 1). A feature common to all protocols is that RBPs are crosslinked by ultraviolet (UV) light to their bound RNAs within the intact endogenous environment. Acutely harvested cells, triturated or pulverized tissues, or even whole organisms such as *C. elegans* are immediately exposed to UV light, which crosslinks RBPs to their bound RNAs in their native (intracellular) state. We have adopted use of the term "CLIP" as a general way to refer to all protocols that use UV crosslinking and protein purification of directly bound native RNA-protein complexes.

Most CLIP variants exploit the capacity of UV light to promote formation of covalent bonds between RBPs and their direct RNA binding sites (Table S1). Unlike formaldehyde crosslinking that is used in chromatin immunoprecipitation (ChIP) and some variants of RIP, UV does not crosslink proteins to each other. UV crosslinking requires direct (Angstrom) contact between an amino acid and a nucleobase, and therefore ensures that only direct protein-RNA interactions are preserved. The high strength of the covalent bond allows a stringent buffer containing ionic detergents to be used for cell lysis, which disrupts most non-crosslinked protein-protein and protein-RNA interactions. The stringent buffer also increases the accessibility of RNA for RNase fragmentation and minimizes the chance of co-purifying multiple associated RBPs during later immunoprecipitation, thus increasing data specificity. RNA fragmentation is crucial to avoid co-purifying multiple RBPs that crosslink to the same RNAs, and to provide insight into the position of RNA binding sites, since the RNA fragment contains the crosslink site. A well-optimized RNase concentration is important to minimize the constraints at the sites of RNase cleavage, which can lead to biased binding site assignment (Haberman et al. 2017).

Purification of endogenous (i.e., untagged) RBPs normally requires that antibodies are available for efficient immunoprecipitation. The need for stringent purification and quality control varies to some extent on the type and expression of the RBP being studied, but in general this improves signal to noise of detection of directly bound RNAs. Some RBPs participate in stable RNPs that may not efficiently dissociate under standard CLIP

immunoprecipitation conditions, thus increasing their risk of co-purifying multiple RBPs, and careful protocols have articulated how to optimize the stringency of CLIP IPs (see (R. Darnell 2012). Even a minor co-purification of an abundant RBP that crosslinks with high efficiency can lead to a major loss of specificity if the RBP-of-interest is of low abundance or crosslinks poorly. In such cases, RBPs can be epitope-tagged to allow purification under denaturing conditions (Granneman et al. 2009; Huppertz et al. 2014). With epitope-tagging, it is important to avoid overexpression, which can give artificial binding to unnatural RNA partners, and to test if the function, stability and localization of the tagged RBP remain unperturbed.

To prepare cDNA libraries from the CLIP RNA fragments, an adapter is ligated to the 3' end of the RNA. The sequence of the adapter is complementary to the primers that are later used in reverse transcription (RT) and PCR. This is followed by purification of the protein-RNA complexes with the use of SDS-PAGE and membrane transfer. Both steps reduce the contamination of non-crosslinked RNAs and excess adapters, which normally run at a lower molecular weight in the gel than fragments crosslinked to the RBP; moreover, free (contaminating) RNA doesn't bind well to the membrane. The labelled protein-RNA complex can be visualized on the membrane, serving to optimize RNA fragmentation and to control for the specificity of purified complexes. Labelling can be achieved radioactively with ^{32}P on the 5' end of the RNA (Ule et al. 2003) or the adapter (Chi et al. 2009), or by covalently coupling an infrared dye to the adapter, which allows visualization with infrared imaging (irCLIP) (Zarnegar et al. 2016). Under well-optimized RNase conditions, the protein-RNA complexes migrate on the gel diffusely above the apparent MW of the immunoprecipitated RBP, since the diverse sizes of RNA fragments variably affect the migration of the RBP-RNA complex. These complexes are then excised from the appropriate region of the membrane according to recommended guidelines (Huppertz et al., 2014; Ule et al., 2005), and the RBP is digested with proteinase K, which leaves only a short peptide at the crosslink site and releases the RNA fragments into solution. To prepare a cDNA library from these RNA fragments, the original CLIP protocol ligated a second adapter to the 5' end of RNA fragments. This approach has been used when CLIP was first combined with high-throughput sequencing (HITS-CLIP, Figure 1). With the HITS-CLIP protocol, it is important that the cDNAs read through the crosslink site during the reverse transcription in order to ensure that they contain both adapters, which is required for the amplification and sequencing of cDNAs.

Primary variations in CLIP protocols

The first aspect that has been varied in the CLIP methods is the crosslinking of protein-RNA complexes. Most variants of CLIP expose cells or tissues to the UV-C wavelength (254 nm), a procedure that usually takes less than a minute with cells placed on ice (Ule et al. 2003) or frozen in liquid nitrogen. PAR-CLIP introduces a variation in the crosslinking strategy (Hafner et al. 2010), where live cells are pre-incubated for hours with photoactivatable ribonucleosides 4-thiouridine (4SU) or 6-thioguanosine (6SG), which label the RNA in vivo and enable protein-RNA crosslinking to be performed with UV-A wavelength (365 nm). A mass spectrometry study has compared this protocol to UV-C crosslinking to conclude that the efficiency of the two approaches is quite similar across all RBPs (Castello et al. 2012). The advantage of UV-A crosslinking is its decreased capacity to induce DNA damage as compared to UV-C, although in most CLIP experiments samples are chilled or frozen during ~30 second long UV crosslinking and then immediately solubilized, so that biologic consequences of DNA damage are minimal. The disadvantages of PAR-CLIP are its limitation to biological systems where the photoactivatable nucleosides can be efficiently incorporated, and its recommended use of >100 million cells (Danan, Manickavel, and Hafner 2016), while 1 million or less are used by standard variants of CLIP. Moreover, prolonged preincubation with 6SG or 4SU can cause cellular toxicity, including stress responses and inhibition of rRNA synthesis (Burger et al. 2013; Huppertz et al. 2014) and therefore care needs to be taken to monitor the cellular response to these ribonucleosides (Burger et al. 2013; Huppertz et al. 2014).

For methods aimed at studying RNA methylation, it is also possible to use a mutant RNA methylase enzyme, which forms a covalent bond with RNA spontaneously without the need for UV-induced crosslinking, as has been done by m5C-miCLIP for transcriptome-wide mapping of 5-methylcytosine (m5C) modification sites (Sugimoto et al. 2012; Hussain et al. 2013). Alternatively, proteins can be crosslinked to purified RNA with UV light *in vitro*. This has been exploited by protocols that crosslink an antibody that binds to a methylated base in order to study RNA methylation, such as m6A-CLIP (Chen et al. 2015; Ke et al. 2015; Linder et al. 2015) (Figure 1). The antibody-RNA complex is then captured on protein A/G magnetic beads, and the sample continues to the rest of the CLIP protocol.

The second aspect of CLIP methods in which major variations have been explored involves ligating adaptors to the fragmented RNA. By modifying the ligation conditions, it is possible to enable proximity ligation of the two strands of RNA duplexes that are bound by double-strand RNA binding proteins (dsRBPs). When the two strands get ligated, the resulting sequencing reads are referred to as 'hybrid reads', since they need to be split into two sequences that map to separate loci in the genome. This approach has been exploited to

study intermolecular duplexes, such as snoRNA-rRNA, mRNA-miRNA and lncRNA-mRNA, as well as intramolecular RNA-RNA duplexes bound by dsRBPs (Chi et al. 2009; Imig et al. 2014; Kudla et al. 2011; Sugimoto et al. 2015; Moore et al. 2015). Further technical details of these protocols, their data analysis, and relationship to other methods to study RNA secondary structure are reviewed in more depth elsewhere (Sugimoto et al. 2017).

A major advance in analysis of RNA structure was made with the development ligation conditions to allow hybrid CLIP (hiCLIP), which ligates the two RNA strands with an additional adaptor to gain control over the intermolecular ligation. This is achieved with the use of two types of RNA adaptors, which are ligated to RNA duplexes: RNA adaptor A with an irreversible 3' block (dideoxynucleotide), and RNA adaptor B with an enzymatically removable 3' block (phosphate). While the RNA duplexes remain bound to dsRBPs during IP, the two adaptors are first ligated in equimolar concentration, thus ~50% of the duplexes are expected to contain a different adaptor at each strand. The 3' block of adaptor B is then removed, and the next ligation reaction is performed to ligate the two arms of RNA duplexes via the adaptor B. This results in hybrid reads that contain the two strands of RNA duplexes separated by the adaptor B, which allows unambiguous identification of the two arms of the hybrid reads. The method has been used to identify RNA duplexes bound by Staufen 1 in HEK293 cells, which identified intermolecular mRNA-mRNA and mRNA-lncRNA duplexes, in addition to RNA duplexes that connect complementary parts of the same RNA. In the coding sequence, these complementary parts were generally proximally located, while in UTRs they were more often located over 100 nts apart. Initial studies with a reporter assay indicated a role for such 'long-range' duplexes in the regulation of mRNA stability (Sugimoto et al. 2015), and their broader functional importance remains to be examined.

The final major variations in CLIP involve the preparation of the cDNA library. Individual nucleotide CLIP (iCLIP) ligates the second adapter to the 3' end of the cDNA via circularization (König et al. 2010). This enables amplification of cDNAs that prematurely truncate at the crosslinked nucleotide, in addition to the readthrough cDNAs, thus increasing the sensitivity of the method. 17 other published protocols similarly amplify truncated cDNAs, including BrdU CLIP, eCLIP and irCLIP (Van Nostrand et al. 2016; Zarnegar et al. 2016; Weyn-Vanhentenryck et al. 2014). In iCLIP and most other CLIP methods, additional sequences are part of the adapter oligonucleotide, including an experimental barcode that enables multiplexing of different experiments, and a unique molecular identifier (UMI) that enables quantification of unique cDNAs to remove the artefacts of variable PCR amplification (König et al., 2010) (Figure 1). Following adapter ligation, cDNAs are then

purified and amplified by PCR, which is achieved in multiple ways by variant protocols, as described in detail elsewhere (Lee and Ule 2018).

CLIP data analysis

The primary advantage of CLIP is that it allows investigators to narrow down the very large number of computationally possible binding sites for a given RBP to a smaller subset of sites with evidence of direct RBP contact in the relevant cell type. This was clearly illustrated in the ability of Ago-miRNA-mRNA CLIP to identify true miRNA binding sites from among a larger set of computationally predicted sites, improving accuracy (Chi et al. 2009) and leading to unexpected discoveries (Chi, Hannon, and Darnell 2012; Moore et al. 2015; Luna et al. 2015).

The first steps in analyzing CLIP sequencing data involve using the experimental barcodes to demultiplex the cDNA libraries, mapping the data to the genome, and identifying unique cDNAs with the use of UMIs, if available. Moreover, multiple methods can be used normalize the binding patterns relative to RNA abundance (Lee and Ule 2018). Depending on whether the protocol amplified truncated cDNAs or not, the peaks of RNA binding sites can be assigned in two primary ways. For HITS-CLIP and other protocols that only amplify readthrough cDNAs, the full position of mapped reads is examined. The most likely position of the crosslink site can be identified within a proportion of the reads by identifying crosslinking-induced mutation sites, such as C to T transitions when 4SU is used in PAR-CLIP (Hafner et al. 2010), or deletions and other types of mutations in HITS-CLIP (C. Zhang and Darnell 2011). In iCLIP and 17 other protocols that amplify truncated cDNAs, crosslink sites are located at the start of these cDNAs. The truncated cDNAs strongly dominate the iCLIP libraries, and thus analysis of cDNA starts enables the nucleotide-resolution mapping of the crosslink sites, especially when combined with high-resolution peak-calling algorithms (Haberman et al. 2017; Sugimoto et al. 2012). The simplest peak-calling approach is to identify clusters of crosslinking events with cDNA counts that are significantly higher compared to randomized data in same genomic regions (König et al. 2010).

While peak analysis of RBP-RNA binding sites has been of great value in understanding binding specificity and demarcating discrete points of regulation (Chakrabarti et al. 2017), not all regulatory factors have been shown to bind to individual discrete peaks. For example, hnRNP C binds widely across the transcriptome, with a periodicity suggesting the possibility of nucleosomal-like organization on RNA transcripts, which also reflects its binding to the dual U-rich sequences in antisense Alu elements (König et al. 2010; Zarnack et al. 2013). FMRP, the protein lost in the commonest inherited form of intellectual disability, the Fragile-X

Syndrome, was found by CLIP to bind to a subset of transcripts across their entire coding length (J. C. Darnell et al. 2011). This was then shown to result from dual interactions of FMRP with the ribosome and the mRNA to cause ribosomal elongation arrest and translational pausing on its target transcripts. In such instances, more complex computational modelling may be needed in addition to peak analysis to understand how the RBPs recognize specific target RNAs.

Paralleling the genomic mapping and peak assignment, an essential concomitant is to examine the quality of data. Analysis of negative controls is particularly important for interpreting the data specificity. For example, purification can be performed from cells where the RBP is absent, such as knockout cells, or when using tag-based purification, cells that do not express a tagged protein. As an alternative, non-specific serum or IgG can be used for IP, or the RBP can be purified from non-crosslinked cells. Ideally, these controls should not produce any clearly detectable signal during SDS-PAGE analysis, and the number of unique cDNAs in their resulting libraries should be at least 100-fold fewer than in the specific experiments (König et al. 2010).

The sensitivity and specificity of data can be further examined by assessing the total number of unique cDNAs in the sequencing library and the peaks, respectively. To monitor the specificity of CLIP, a suitable peak-calling program needs to be chosen according to the CLIP protocol used to produce the data (Chakrabarti et al. 2017). A low number of peaks indicates either that the cDNAs are dispersed along the transcripts, or that they are concentrated in a small number of abundant RNAs (such as rRNA) – this could either reflect low preference of the RBPs for specific sites on longer RNAs, or low quality of data. Thus, some understanding of the expected binding properties of the RBP can help when interpreting CLIP data.

Finally, an extremely informative way to examine CLIP data sensitivity and specificity is to assess binding peaks with orthogonal data in the format of RNA maps (Donny D. Licatalosi and Darnell 2010; D. D. Licatalosi et al. 2008), which allow the investigator to examine if binding is enriched at exons or mRNAs that are differentially regulated by the RBP. Analysis of crosslink peaks in the format of RNA maps is particularly valuable when comparing multiple datasets for the same RBP. For example, data produced by iCLIP of PTBP1 led to a larger number of peaks around regulated exons than data produced by other protocols, even though the number of unique cDNAs in iCLIP is equal or smaller (Chakrabarti et al. 2017).

cTag-CLIP

A major forefront in biological research is to understand the functional roles of individual cell types in a multicellular organism. Cell-type specific regulation of RNA expression and processing emerges as a common phenomenon from recent RNA profiling studies in selected cell populations (X. Zhang et al. 2016; Yap et al. 2016; Bhate et al. 2015). Importantly, it is clear from these studies that many cell-type specific RNA regulatory events are not identifiable in whole tissue analysis. Therefore, developing new technologies with cell-type resolutions is critical to advance our understanding of cell-type specific RNA regulation.

To meet the challenge of applying CLIP to study cell-type specific biology, we developed the conditionally-tagged (cTag) CLIP technology, which utilizes a genetically engineered mouse that carries a latent GFP-tagged allele of the RNA-binding protein (RBP) of interest. The latent GFP-tagged allele is "off" by default, leaving the native allele unaffected, and only turns "on" in the presence of Cre recombinase (Figure 2A) (Hwang et al. 2017). Importantly, the GFP-tagged allele is expressed by homologous recombination into the endogenous locus, such that the quantitative, qualitative (alternate isoforms), spatial and temporal expression of GFP-tagged RBP is minimally disruptive to the biology of the system, and can be precisely controlled through Cre expression to achieve cell-type specificity. Cell-type specific CLIP analysis for the RBP of interest is then be performed using a mix of high affinity GFP antibodies to immunoprecipitate the GFP-tagged RBPs under stringent conditions (Figure 2B) only from cells expressing Cre recombinase.
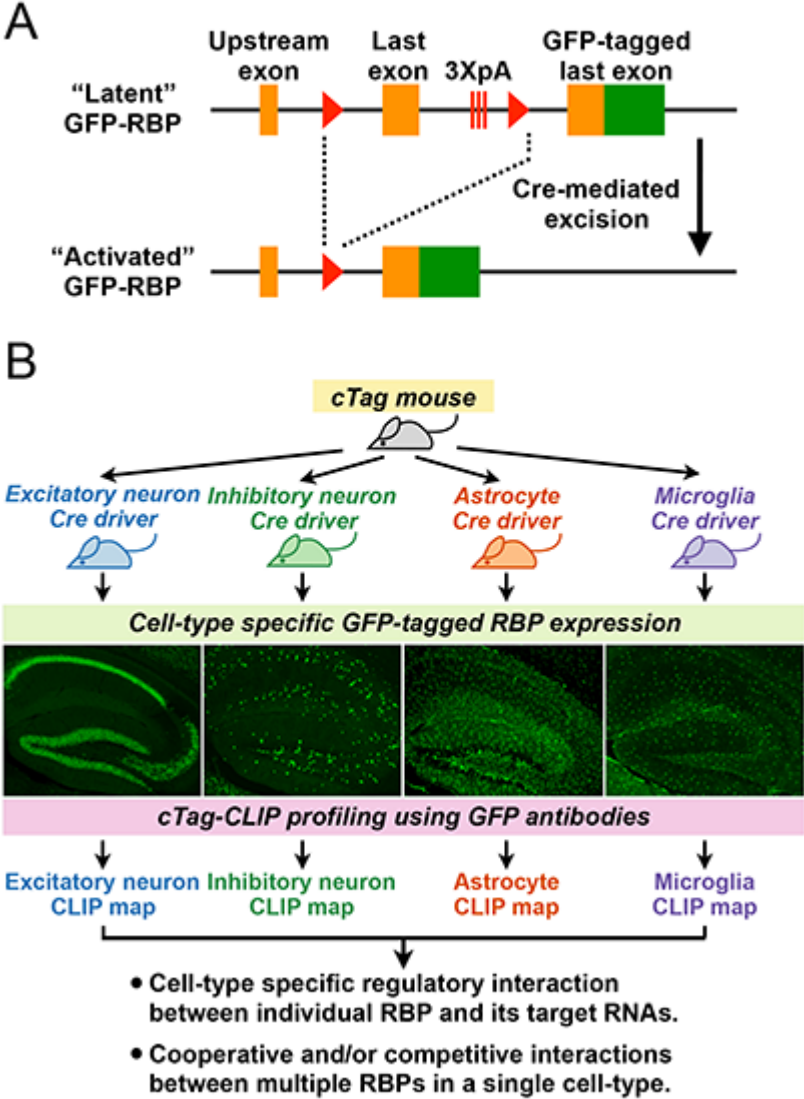
Figure 2: cTag-CLIP

A) The basic design of a cTag mouse. An extra copy of the last coding exon fused with GFP sequence (green) is inserted downstream of sequences encoding the RNA-binding protein of interest. To prevent read-through, triple poly(A) sites (3XpA) are inserted before the GFP-tagged exon. Cre-mediated excision of the endogenous last exon results in expression of GFP-tagged RBP from its native locus. Arrowheads: loxP sites. B) Illustrative examples of cTag-CLIP with results from the cTag-PABP mouse (Hwang et al. 2017).

The cTag-CLIP technology has several important features to ensure a proper representation of endogenous RNA-protein interactions but at the same time to maintain a broad applicability. First, since cTag mice are created using a "knock-in" approach, the proper cellular expression of RBPs and the RNA-RBP stoichiometry is likely to be maintained, both of which are difficult to achieve with either traditional or even bacterial artificial chromosome (BAC) transgenes (He et al. 2012; Gerfen, Paletzki, and Heintz 2013) (Figure 2A). Second, the utilization of a mix of high affinity GFP antibodies (Heiman et al. 2014) allows cTag-CLIP analysis to be performed with a standardized and optimized high stringency protocol for different RBPs (Hwang et al. 2017). The standardization eliminates the need for RBP-specific antibodies and the subsequent experimental optimization to achieve high stringency immunoprecipitation for individual RBPs. Furthermore, it also provides an equal footing for cross-comparison of cTag-CLIP data from different RBPs to facilitate data integration. Third, its compatibility with the Cre/Lox system allows direct application of existing or newly developed Cre mouse lines, which makes it a versatile tool for wide-ranging biological investigations.

We have since successfully established a series of cTag mice to perform cTag-CLIP analysis for different RBPs in multiple cell types. We validated this approach using conditionally tagged poly(A)-binding protein (PABP; gene symbol: *Pabpc1*) to generate the cTag-PABP mouse and to establish the cTag-PAPERCLIP technique (Hwang et al. 2017). cTag-PAPERCLIP utilizes PABP as a biological filter to specifically clone distal mRNA 3′ ends to generate comprehensive mRNA alternative polyadenylation (APA) maps. With cTag-PAPERCLIP, we performed an in-depth APA analysis in individual cell types *in vivo* in adult mouse brain to demonstrate the dynamics of APA regulation and discovered a previously unknown role of APA in diversifying protein expression in the brain.

Several observations from this cTag-PAPERCLIP study exemplify the technical advances and, more importantly, biological implications of the cTag-CLIP technology, which are detailed below.

1. ***cTag-CLIP achieves desired cell-type specificity to provide a powerful means to dissect cell-type specific biology:*** The ubiquitous expression of PABPC1 presents an ideal test case for the capabilities of cTag-CLIP technology in achieving high cell-type specificity. To evaluate the cell-type specificity at the anatomical level, we examined the expression patterns of PABP-GFP in brains from four mouse models generated by crossing the cTag-PABP mouse to four Cre drivers with different cell-type specificity (excitatory neurons, inhibitory neurons, astrocytes and microglia (Figure 2B)). The analyses indeed

confirmed the distinct expression patterns of PABP-GFP from the four mouse models and their consistency with the corresponding cell-type markers. Next, we evaluated the cell-type specificity at the molecular level by analyzing the four resulting RNA profiles from each mouse model generated by cTag-PAPERCLIP. These cTag-PAPERCLIP profiles correlated with independently generated cell-type specific RNA-seq profiles and showed expected segregation of known cell-type specific marker genes by the Cre activities. Overall, these data demonstrate that cTag-CLIP indeed accomplishes its goals: 1) specific expression of the tagged RBP in the desired cell populations and 2) generation of cell-type specific CLIP maps. Importantly, cTag-CLIP data obtained from cTag mouse models are very consistent with data generated in wild-type mouse through a variety of methods in multiple occasions (Saito et al. 2018; Hwang et al. 2017; Jereb et al. 2018), showing that cTag-CLIP approach in general does not interfere with normal RBP biology and hence provides a powerful means for dissecting cell-type specific biology.

2.   ***cTag-CLIP has excellent power to "zoom in" on less abundant cell-types to discover cell-type specific regulatory events:*** Our in-depth cTag-PAPERCLIP analysis of APA patterns across four major brain cell types and between different cellular states in microglia (quiescent vs. activated) revealed novel cell-type specific regulatory events. For example, we demonstrated that cell-type specific RBPs directly shape the distinct APA patterns of *Itsn1*, *Map4* and *Cdc42* between neurons and glia/non-neural tissues. Moreover, in activated microglia, we identified that *Araf* switches from a truncated APA isoform (which generates a kinase-deficient protein) to a full-length APA isoform (which generates a kinase-intact protein). Importantly, these findings nicely illustrate how gaining resolution at the cell-type level allows cTag-CLIP to generate new biological insights as these four genes are expressed in multiple cell types in the brain and it would be difficult for whole tissue analysis to pick up these cell-type specific regulatory events.

3.   ***cTag-CLIP well-preserves native transcriptome states and has an advantage in profiling sensitive cell types:*** One distinctive feature of cTag-CLIP is its "built-in" cell-type specificity imparted through genetic expression of cell-type restricted Cre in cTag mice (Figure 2B). As a result, cTag-CLIP can utilize the whole tissue as the input material without going through the process of tissue dissociation and cell-type purification, which is necessary for common RNA profiling methods such as RNA-seq to achieve cell-type specificity. Tissue dissociation and cell-type purification disrupts the tissue microenvironment, causes cellular stress and may alter gene expression (Cardona et al. 2006; Okaty, Sugino, and Nelson 2011). Most importantly, cTag-CLIP kept the expression of microglia activation markers and immediate early genes (which respond to cellular stress)

low compared with RNA-seq profiling from *ex vivo* purified microglia. These results demonstrate that cTag-CLIP better preserves native transcriptome states of individual cell types and therefore is suitable for accurate *in vivo* RNA profiling.

The success of our initial cTag-CLIP studies starts to reveal a diversity of RNA regulatory events in individual cell-types. Generation and integration of cTag-CLIP maps for different RBPs in multiple cells types will provide essential information to refine the existing rules of RNA-protein interaction for individual cell-types. Importantly, with the cTag-CLIP maps, it is now feasible to deduce the cooperative and/or competitive interactions among different RBPs for the same RNA binding sites in a single cell-type to construct high-order RNA-protein interaction maps. Recent improvement in generating knock-in mouse models from advances in synthetic biology and gene-targeting technology will speed up the process of expanding the cTag mice repertoire (Shalem, Sanjana, and Zhang 2015; Wright, Nuñez, and Doudna 2016). Moreover, continued identification of cell-type specific markers and technical advancement enhancing the sensitivity of CLIP technique will broaden the application of cTag-CLIP to new and rare cell types (Tasic et al. 2016; Lee and Ule 2018).

Insights from CLIP and Future Directions

The development of CLIP has been driven largely by the need to understand the regulation of pre-mRNA processing, although it can be applied to all forms of post-transcriptional RNA regulation, including translation as demonstrated with CLIP studies of FMRP. For studies of pre-mRNA, it is crucial to map the binding of nuclear RBPs to introns, which are generally of low abundance and of transient nature. To allow studies of RBP interactions on introns, specificity of the method had to be maximized, which led to the inclusion of covalent crosslinking followed by several stringent purification and quality control steps. While the success of this approach was demonstrated by the very first studies of Nova, CLIP also led to an unexpected finding that nuclear RBPs have many thousands of high-affinity binding sites, most of which are far from exons (Ule et al. 2003; D. D. Licatalosi et al. 2008). In recent years, it has become clear that many abundant nuclear RBPs, such as hnRNPC, often bind to specific transposable elements or other repetitive sequences that are present in introns, where they repress processing of cryptic exons and polyadenylation sites (Zarnack et al. 2013; Sibley, Blazquez, and Ule 2016). While this repressive binding safeguards the transcriptome from newly-emerging intronic sites of RNA processing, it may also create an opportunity for a gradual evolutionary emergence of new alternative exons upon mutations

that weaken the repressive binding (Attig et al. 2016). Thus, even though the large number of intronic binding sites of nuclear RBPs initially came as a surprise, it is becoming clear that this plays an important role in the control of rarely included cryptic exons, including those mediating nonsense mediated decay (Eom et al. 2013), and in the evolution of gene expression (Sibley, Blazquez, and Ule 2016).

Molecular and Physiologic Insights from CLIP

Following identification of in vivo binding sites, further integration with orthogonal functional data can then allow interpretation of regulatory mechanisms that operate in cells and organs. For example, the integrative approach to examine binding profiles around regulated exons and polyadenylation sites has been visualized in the form of RNA maps (Ule et al. 2006; Witten and Ule 2011; D. D. Licatalosi et al. 2008). This union of biochemical CLIP data with functional evidence via the bioinformatic generation of global "RNA maps" led to the recognition that the position of protein binding sites in transcripts provided information relevant to their mechanisms and regulatory outcomes (Ule et al. 2003; D. D. Licatalosi et al. 2008). For example, it revealed that nuclear RBPs often regulate splicing and 3' processing via similar positional principles, with repressive binding located close to the sites of RNA processing, and enhancing binding located further downstream (D. D. Licatalosi et al. 2008; Rot et al. 2017). The insights from such analysis can complement predictive programs such as TargetScan, which predicts possible binding sites for miRNAs by incorporating information on evolutionary conservation and general binding rules; for example, CLIP could identify important miRNA binding sites that are unique to human (Chi et al. 2009) or are located at non-canonical sequences that don't conform to general rules (Chi, Hannon, and Darnell 2012; Moore et al. 2015).

Many studies have shown that RNP assembly on the nascent and mature mRNA, and the resulting regulatory mechanisms are coordinated at many levels (Singh et al. 2015), and CLIP has solidified these observations. Many RBPs shuttle between nucleus and cytoplasm, which allows mRNP assembly to start in the nucleus on pre-mRNAs and then continue to influence cytoplasmic regulation. For example, while CLIP identified roles for Nova proteins in nuclear pre-mRNA processing (Ule et al. 2003; D. D. Licatalosi et al. 2008), it also revealed 3' UTR binding sites in the vicinity of factors involved in alternative polyadenylation; again using genetics, such binding was found to be functional in determining the use of APA sites (Ule et al. 2003; D. D. Licatalosi et al. 2008). Moreover, CLIP, together with immuno-EM localization of Nova in the dendrite, identified a new link between nuclear regulation and synaptic RNA localization within neurons (Racca et al. 2010). This demonstrated that an RNA regulatory factor was governed by its binding affinity

to motifs that can be located at various positions in the transcriptome, and that a mammalian splicing factor could also bind and regulate mature mRNA biology.

Biologically, some of the more powerful conclusions from CLIP biochemistry have come from pairing it with genetic manipulations. CLIP studies have been most definitive in vivo in mice where individual (R. B. Darnell 2013) and more recently multiple (Saito et al. 2018) RBPs have been genetically deleted. Similarly, CLIP studies have been successful in tissue culture cells where RBPs have been knocked down. Such genetic studies have been complemented by genetic rescue experiments. This was illustrated in early studies demonstrating the requirement of Nova for splicing of the Z+ isoform of the agrin gene in motor neurons (Ule et al. 2005), and as a result the Nova1 KO mice have abnormal neuromuscular junctions and are paralyzed (Ruggiu et al. 2009). Moreover, Nova proteins regulate an alternative exon in Dab1 (Ule et al. 2006), an important regulator of reelin signaling in newly born migrating neurons (Yano et al. 2010). Importantly, specific phenotypes of Nova KO mice can be rescued by expressing the appropriately spliced isoform; interestingly, when this was done by mating Nova1 KO mice to transgenic mice constitutively expressing Z+ agrin, the normal neuromuscular junction was re-established, but the mice remained paralyzed, illustrating the multi-layered aspect of RBP-RNA regulation in vivo (Ruggiu et al. 2009). Beyond studies of Nova, CLIP has been successfully applied to characterize the function and assembly of many additional disease-relevant RBPs in the brain and other tissues of model organisms (Modic, Ule, and Sibley 2013; R. B. Darnell 2013).

CLIP: Future Directions

One of the most recent applications of CLIP has been to undertake studies of epigenetic modifications of RNA. These were pioneered with studies of m6A RNA modifications, using a variant of CLIP in which the modified nucleotide itself was directly crosslinked, *in vitro*, to m6A specific antibodies (Ke et al. 2015; Linder et al. 2015; Patil, Pickering, and Jaffrey 2018). This provides a potentially general strategy for analyzing what is an expanding number of post-transcriptional RNA modifications important in cell function and human disease (Jonkhout et al. 2017).

The same general principles of CLIP experiments apply in analyses of RNA modifications as in traditional CLIP: unbiased, genome-wide genetic studies allow demarcation of specific sites, here using direct modification-specific antibodies to crosslink the modified RNA nucleotide. For example, m6A CLIP was used to identify terminal exons as the most common positions for m6A modification on RNA transcripts, rather than exons harboring

stop codons (Ke et al. 2015); to demonstrate primary functions of m6A in RNA turnover rather than regulation of alternative splicing (Ke et al. 2017); and to clarify our current understanding of the biologic role of m6A RNA modifications (Ke et al. 2017; R. B. Darnell, Ke, and Darnell 2018).

Importantly, CLIP has been shown to be of value also for studies of human postmortem tissues (Tollervey et al. 2011). This is important as such tissues are often frozen in liquid nitrogen. RNA-protein complexes appear to survive such conditions well, and allow CLIP to be applied to tissues relevant for a range of human disease (e.g., cancer, immunologic or neurologic disease) using a range of human tissue (tumor (Vanharanta et al. 2014), blood (Loeb et al. 2012) and brain (Tollervey et al. 2011; Scheckel et al. 2016), respectively). The value of CLIP has also been demonstrated to unravel the dynamics of RNPs under disease conditions by comparing tissue from disease patients and controls (Tollervey et al. 2011), or monitoring the impact of viral infection on host transcripts in biology and human disease (Flynn et al. 2015; Luna et al. 2015, 2017; Moore et al. 2018).

Demarcation of RNA structure using CLIP was originally pioneered with hiCLIP to analyze Staufen-RNA interactions (Sugimoto et al. 2015), and subsequently developed with a variety of complementary approaches such as PARIS (psoralen analysis of RNA interactions and structures) (Lu, Gong, and Zhang 2018; Lu et al. 2016) as discussed above. Such analyses are important in their own right, and in their potential to allow more general understanding of structures accessible to RBP regulation, as determined from CLIP analyses. The combined understanding of protein-binding sites, RNA structure and evolutionary constraints promises to expand our ability to predict genetic variants using Bayesian (C. Zhang et al. 2010) and machine-learning tools (Barash et al. 2010), and ultimately apply such understanding to interpretation of those variants that have impact on disease phenotypes.

Since the very initial development of CLIP, a common theme of many CLIP studies has been that they combine discovery of fundamental mechanisms of gene regulation along with biologic and clinically relevant insights. CLIP holds great promise for studies aimed at defining the functionally important RNA sites across the transcriptome, unravelling the mechanisms of RNP assembly and function, and uncovering the role of these mechanisms in diseases.

Acknowledgements

References

Attig, Jan, Igor Ruiz de Los Mozos, Nejc Haberman, Zhen Wang, Warren Emmett, Kathi Zarnack, Julian König, and Jernej Ule. 2016. "Splicing Repression Allows the Gradual Emergence of New Alu-Exons in Primate Evolution." *eLife* 5 (November). https://doi.org/10.7554/eLife.19545.

Barash, Yoseph, John A. Calarco, Weijun Gao, Qun Pan, Xinchen Wang, Ofer Shai, Benjamin J. Blencowe, and Brendan J. Frey. 2010. "Deciphering the Splicing Code." *Nature* 465 (7294): 53–59.

Bhate, Amruta, Darren J. Parker, Thomas W. Bebee, Jaegyoon Ahn, Waqar Arif, Edrees H. Rashan, Sandip Chorghade, et al. 2015. "ESRP2 Controls an Adult Splicing Programme in Hepatocytes to Support Postnatal Liver Maturation." *Nature Communications* 6 (November): 8768.

Burger, Kaspar, Bastian Mühl, Markus Kellner, Michaela Rohrmoser, Anita Gruber-Eber, Lukas Windhager, Caroline C. Friedel, Lars Dölken, and Dirk Eick. 2013. "4-Thiouridine Inhibits rRNA Synthesis and Causes a Nucleolar Stress Response." *RNA Biology* 10 (10): 1623–30.

Cardona, Astrid E., Deren Huang, Margaret E. Sasse, and Richard M. Ransohoff. 2006. "Isolation of Murine Microglial Cells for RNA Analysis or Flow Cytometry." *Nature Protocols* 1 (4): 1947–51.

Castello, Alfredo, Bernd Fischer, Katrin Eichelbaum, Rastislav Horos, Benedikt M. Beckmann, Claudia Strein, Norman E. Davey, et al. 2012. "Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins." *Cell* 149 (6): 1393–1406.

Chakrabarti, Anob M., Nejc Haberman, Arne Praznik, Nicholas M. Luscombe, and Jernej Ule. 2017. "Data Science Issues in Understanding Protein-RNA Interactions." https://doi.org/10.1101/208124.

Chen, Kai, Zhike Lu, Xiao Wang, Ye Fu, Guan-Zheng Luo, Nian Liu, Dali Han, et al. 2015. "High-Resolution N(6) -Methyladenosine (m(6) A) Map Using Photo-Crosslinking-Assisted m(6) A Sequencing." *Angewandte Chemie* 54 (5): 1587–90.

Chi, Sung Wook, Gregory J. Hannon, and Robert B. Darnell. 2012. "An Alternative Mode of microRNA Target Recognition." *Nature Structural & Molecular Biology* 19 (3): 321–27.

Chi, Sung Wook, Julie B. Zang, Aldo Mele, and Robert B. Darnell. 2009. "Argonaute HITS-CLIP Decodes microRNA-mRNA Interaction Maps." *Nature* 460 (7254): 479–86.

Conlon, Erin G., and James L. Manley. 2017. "RNA-Binding Proteins in Neurodegeneration: Mechanisms in Aggregate." *Genes & Development* 31 (15): 1509–28.

Danan, Charles, Sudhir Manickavel, and Markus Hafner. 2016. "PAR-CLIP: A Method for Transcriptome-Wide Identification of RNA Binding Protein Interaction Sites." *Methods in Molecular Biology* 1358: 153–73.

Darnell, Jennifer C., Sarah J. Van Driesche, Chaolin Zhang, Ka Ying Sharon Hung, Aldo Mele, Claire E. Fraser, Elizabeth F. Stone, et al. 2011. "FMRP Stalls Ribosomal Translocation on mRNAs Linked to Synaptic Function and Autism." *Cell* 146 (2): 247–61.

Darnell, Robert. 2012. "CLIP (cross-Linking and Immunoprecipitation) Identification of RNAs Bound by a Specific Protein." *Cold Spring Harbor Protocols* 2012 (11): 1146–60.

Darnell, Robert B. 2010. "HITS-CLIP: Panoramic Views of Protein-RNA Regulation in Living Cells." *Wiley Interdisciplinary Reviews. RNA* 1 (2): 266–86.

———. 2013. "RNA Protein Interaction in Neurons." *Annual Review of Neuroscience* 36 (July): 243–70.

Darnell, Robert B., Shengdong Ke, and James E. Darnell Jr. 2018. "Pre-mRNA Processing Includesmethylation of Adenosine Residues That Are Retained in mRNA Exons and the Fallacy of 'RNA Epigenetics.'" *RNA* 24 (3): 262–67.

Eom, Taesun, Chaolin Zhang, Huidong Wang, Kenneth Lay, John Fak, Jeffrey L. Noebels, and Robert B. Darnell. 2013. "NOVA-Dependent Regulation of Cryptic NMD Exons

Controls Synaptic Protein Levels after Seizure." *eLife* 2 (January): e00178.

Flynn, Ryan A., Lance Martin, Robert C. Spitale, Brian T. Do, Selena M. Sagan, Brian Zarnegar, Kun Qu, et al. 2015. "Dissecting Noncoding and Pathogen RNA-Protein Interactomes." *RNA* 21 (1): 135–43.

Gerfen, Charles R., Ronald Paletzki, and Nathaniel Heintz. 2013. "GENSAT BAC Cre-Recombinase Driver Lines to Study the Functional Organization of Cerebral Cortical and Basal Ganglia Circuits." *Neuron* 80 (6): 1368–83.

Granneman, Sander, Grzegorz Kudla, Elisabeth Petfalski, and David Tollervey. 2009. "Identification of Protein Binding Sites on U3 snoRNA and Pre-rRNA by UV Cross-Linking and High-Throughput Analysis of cDNAs." *Proceedings of the National Academy of Sciences of the United States of America* 106 (24): 9613–18.

Haberman, Nejc, Ina Huppertz, Jan Attig, Julian König, Zhen Wang, Christian Hauer, Matthias W. Hentze, et al. 2017. "Insights into the Design and Interpretation of iCLIP Experiments." *Genome Biology* 18 (1): 7.

Hafner, Markus, Markus Landthaler, Lukas Burger, Mohsen Khorshid, Jean Hausser, Philipp Berninger, Andrea Rothballer, et al. 2010. "Transcriptome-Wide Identification of RNA-Binding Protein and microRNA Target Sites by PAR-CLIP." *Cell* 141 (1): 129–41.

Heiman, Myriam, Ruth Kulicke, Robert J. Fenster, Paul Greengard, and Nathaniel Heintz. 2014. "Cell Type-Specific mRNA Purification by Translating Ribosome Affinity Purification (TRAP)." *Nature Protocols* 9 (6): 1282–91.

He, Miao, Yu Liu, Xiaowo Wang, Michael Q. Zhang, Gregory J. Hannon, and Z. Josh Huang. 2012. "Cell-Type-Based Analysis of microRNA Profiles in the Mouse Brain." *Neuron* 73 (1): 35–48.

Huppertz, Ina, Jan Attig, Andrea D'Ambrogio, Laura E. Easton, Christopher R. Sibley, Yoichiro Sugimoto, Mojca Tajnik, Julian König, and Jernej Ule. 2014. "iCLIP: Protein-RNA Interactions at Nucleotide Resolution." *Methods* 65 (3): 274–87.

Hussain, Shobbir, Abdulrahim A. Sajini, Sandra Blanco, Sabine Dietmann, Patrick Lombard, Yoichiro Sugimoto, Maike Paramor, et al. 2013. "NSun2-Mediated Cytosine-5 Methylation of Vault Noncoding RNA Determines Its Processing into Regulatory Small RNAs." *Cell Reports* 4 (2): 255–61.

Hwang, Hun-Way, Yuhki Saito, Christopher Y. Park, Nathalie E. Blachère, Yoko Tajima, John J. Fak, Ilana Zucker-Scharff, and Robert B. Darnell. 2017. "cTag-PAPERCLIP Reveals Alternative Polyadenylation Promotes Cell-Type Specific Protein Diversity and Shifts Araf Isoforms with Microglia Activation." *Neuron* 95 (6): 1334–49.e5.

Imig, Jochen, Andreas Brunschweiger, Anneke Brümmer, Boris Guennewig, Nitish Mittal, Shivendra Kishore, Panagiota Tsikrika, André P. Gerber, Mihaela Zavolan, and Jonathan Hall. 2014. "miR-CLIP Capture of a miRNA Targetome Uncovers a lincRNA H19–miR-106a Interaction." *Nature Chemical Biology* 11 (December). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 107.

Jereb, Sasa, Hun-Way Hwang, Eric Van Otterloo, Eve-Ellen Govek, John J. Fak, Yuan Yuan, Mary E. Hatten, and Robert B. Darnell. 2018. "Differential 3′ Processing of Specific Transcripts Expands Regulatory and Protein Diversity Across Neuronal Cell Types." https://doi.org/10.1101/245886.

Jonkhout, Nicky, Julia Tran, Martin A. Smith, Nicole Schonrock, John S. Mattick, and Eva Maria Novoa. 2017. "The RNA Modification Landscape in Human Disease." *RNA* 23 (12): 1754–69.

Ke, Shengdong, Endalkachew A. Alemu, Claudia Mertens, Emily Conn Gantman, John J. Fak, Aldo Mele, Bhagwattie Haripal, et al. 2015. "A Majority of m6A Residues Are in the Last Exons, Allowing the Potential for 3′ UTR Regulation." *Genes & Development* 29 (19): 2037–53.

Ke, Shengdong, Amy Pandya-Jones, Yuhki Saito, John J. Fak, Cathrine Broberg Vågbø, Shay Geula, Jacob H. Hanna, Douglas L. Black, James E. Darnell Jr, and Robert B. Darnell. 2017. "mA mRNA Modifications Are Deposited in Nascent Pre-mRNA and Are Not Required for Splicing but Do Specify Cytoplasmic Turnover." *Genes & Development* 31 (10): 990–1006.

König, Julian, Kathi Zarnack, Gregor Rot, Tomaz Curk, Melis Kayikci, Blaz Zupan, Daniel J. Turner, Nicholas M. Luscombe, and Jernej Ule. 2010. "iCLIP Reveals the Function of hnRNP Particles in Splicing at Individual Nucleotide Resolution." *Nature Structural & Molecular Biology* 17 (7): 909–15.

Kudla, Grzegorz, Sander Granneman, Daniela Hahn, Jean D. Beggs, and David Tollervey. 2011. "Cross-Linking, Ligation, and Sequencing of Hybrids Reveals RNA–RNA Interactions in Yeast." *Proceedings of the National Academy of Sciences of the United States of America* 108 (24). National Academy of Sciences: 10010–15.

Lee, Flora C. Y., and Jernej Ule. 2018. "Advances in CLIP Technologies for Studies of Protein-RNA Interactions." *Molecular Cell* 69 (3): 354–69.

Lerner, M. R., and J. A. Steitz. 1979. "Antibodies to Small Nuclear RNAs Complexed with Proteins Are Produced by Patients with Systemic Lupus Erythematosus." *Proceedings of the National Academy of Sciences of the United States of America* 76 (11): 5495–99.

Licatalosi, D. D., A. Mele, J. J. Fak, J. Ule, M. Kayikci, and S. W. Chi. 2008. "HITS-CLIP Yields Genome-Wide Insights into Brain Alternative RNA Processing." *Nature*. nature.com. https://www.nature.com/articles/nature07488.

Licatalosi, Donny D., and Robert B. Darnell. 2010. "RNA Processing and Its Regulation: Global Insights into Biological Networks." *Nature Reviews. Genetics* 11 (1): 75–87.

Linder, Bastian, Anya V. Grozhik, Anthony O. Olarerin-George, Cem Meydan, Christopher E. Mason, and Samie R. Jaffrey. 2015. "Single-Nucleotide-Resolution Mapping of m6A and m6Am throughout the Transcriptome." *Nature Methods* 12 (8): 767–72.

Loeb, Gabriel B., Aly A. Khan, David Canner, Joseph B. Hiatt, Jay Shendure, Robert B. Darnell, Christina S. Leslie, and Alexander Y. Rudensky. 2012. "Transcriptome-Wide miR-155 Binding Map Reveals Widespread Noncanonical microRNA Targeting." *Molecular Cell* 48 (5): 760–70.

Luna, Joseph M., Juan M. Barajas, Kun-Yu Teng, Hui-Lung Sun, Michael J. Moore, Charles M. Rice, Robert B. Darnell, and Kalpana Ghoshal. 2017. "Argonaute CLIP Defines a Deregulated miR-122-Bound Transcriptome That Correlates with Patient Survival in Human Liver Cancer." *Molecular Cell* 67 (3): 400–410.e7.

Luna, Joseph M., Troels K. H. Scheel, Tal Danino, Katharina S. Shaw, Aldo Mele, John J. Fak, Eiko Nishiuchi, et al. 2015. "Hepatitis C Virus RNA Functionally Sequesters miR-122." *Cell* 160 (6): 1099–1110.

Lu, Zhipeng, Jing Gong, and Qiangfeng Cliff Zhang. 2018. "PARIS: Psoralen Analysis of RNA Interactions and Structures with High Throughput and Resolution." *Methods in Molecular Biology* 1649: 59–84.

Lu, Zhipeng, Qiangfeng Cliff Zhang, Byron Lee, Ryan A. Flynn, Martin A. Smith, James T. Robinson, Chen Davidovich, et al. 2016. "RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure." *Cell* 165 (5): 1267–79.

Mili, Stavroula, and Joan A. Steitz. 2004. "Evidence for Reassociation of RNA-Binding Proteins after Cell Lysis: Implications for the Interpretation of Immunoprecipitation Analyses." *RNA* 10 (11): 1692–94.

Modic, Miha, Jernej Ule, and Christopher R. Sibley. 2013. "CLIPing the Brain: Studies of Protein-RNA Interactions Important for Neurodegenerative Disorders." *Molecular and Cellular Neurosciences* 56 (September): 429–35.

Moore, Michael J., Nathalie E. Blachere, John J. Fak, Christopher Y. Park, Kirsty Sawicka, Salina Parveen, Ilana Zucker-Scharff, Bruno Moltedo, Alexander Rudensky, and Robert Darnell. 2018. "ZFP36 RNA-Binding Proteins Restrain T-Cell Activation and Anti-Viral Immunity." https://doi.org/10.1101/247668.

Moore, Michael J., Troels K. H. Scheel, Joseph M. Luna, Christopher Y. Park, John J. Fak, Eiko Nishiuchi, Charles M. Rice, and Robert B. Darnell. 2015. "miRNA-Target Chimeras Reveal miRNA 3'-End Pairing as a Major Determinant of Argonaute Target Specificity." *Nature Communications* 6 (November): 8864.

Niranjanakumari, Somashe, Erika Lasda, Robert Brazas, and Mariano A. Garcia-Blanco. 2002. "Reversible Cross-Linking Combined with Immunoprecipitation to Study RNA–protein Interactions in Vivo." *Methods* 26 (2): 182–90.

Okaty, Benjamin W., Ken Sugino, and Sacha B. Nelson. 2011. "A Quantitative Comparison of Cell-Type-Specific Microarray Gene Expression Profiling Methods in the Mouse Brain." *PloS One* 6 (1): e16493.

Patil, Deepak P., Brian F. Pickering, and Samie R. Jaffrey. 2018. "Reading mA in the Transcriptome: mA-Binding Proteins." *Trends in Cell Biology* 28 (2): 113–27.

Racca, Claudia, Alejandra Gardiol, Taesun Eom, Jernej Ule, Antoine Triller, and Robert B. Darnell. 2010. "The Neuronal Splicing Factor Nova Co-Localizes with Target RNAs in the Dendrite." *Frontiers in Neural Circuits* 4 (March): 5.

Rot, Gregor, Zhen Wang, Ina Huppertz, Miha Modic, Tina Lenče, Martina Hallegger, Nejc Haberman, Tomaž Curk, Christian von Mering, and Jernej Ule. 2017. "High-Resolution RNA Maps Suggest Common Principles of Splicing and Polyadenylation Regulation by TDP-43." *Cell Reports* 19 (5): 1056–67.

Ruggiu, Matteo, Ruth Herbst, Natalie Kim, Marko Jevsek, John J. Fak, Mary Anne Mann, Gerald Fischbach, Steven J. Burden, and Robert B. Darnell. 2009. "Rescuing Z+ Agrin Splicing in Nova Null Mice Restores Synapse Formation and Unmasks a Physiologic Defect in Motor Neuron Firing." *Proceedings of the National Academy of Sciences of the United States of America* 106 (9): 3513–18.

Saito, Yuhki, Yuan Yuan, Ilana Zucker-Scharff, John J. Fak, Yoko Tajima, Donny D. Licatalosi, and Robert B. Darnell. 2018. "Neuron-Specific cTag-CLIP Reveals Cell-Specific Diversity of Functional RNA Regulation in the Brain." https://doi.org/10.1101/244905.

Scheckel, Claudia, Elodie Drapeau, Maria A. Frias, Christopher Y. Park, John Fak, Ilana Zucker-Scharff, Yan Kou, et al. 2016. "Regulatory Consequences of Neuronal ELAV-like Protein Binding to Coding and Non-Coding RNAs in Human Brain." *eLife* 5 (February). https://doi.org/10.7554/eLife.10421.

Shalem, Ophir, Neville E. Sanjana, and Feng Zhang. 2015. "High-Throughput Functional Genomics Using CRISPR-Cas9." *Nature Reviews. Genetics* 16 (5): 299–311.

Sibley, Christopher R., Lorea Blazquez, and Jernej Ule. 2016. "Lessons from Non-Canonical Splicing." *Nature Reviews. Genetics* 17 (7): 407–21.

Singh, Guramrit, Gabriel Pratt, Gene W. Yeo, and Melissa J. Moore. 2015. "The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion." *Annual Review of Biochemistry* 84 (March): 325–54.

Sugimoto, Yoichiro, Anob M. Chakrabarti, Nicholas M. Luscombe, and Jernej Ule. 2017. "Using hiCLIP to Identify RNA Duplexes That Interact with a Specific RNA-Binding Protein." *Nature Protocols* 12 (3): 611–37.

Sugimoto, Yoichiro, Julian König, Shobbir Hussain, Blaž Zupan, Tomaž Curk, Michaela Frye, and Jernej Ule. 2012. "Analysis of CLIP and iCLIP Methods for Nucleotide-Resolution Studies of Protein-RNA Interactions." *Genome Biology* 13 (8): R67.

Sugimoto, Yoichiro, Alessandra Vigilante, Elodie Darbo, Alexandra Zirra, Cristina Militti, Andrea D'Ambrogio, Nicholas M. Luscombe, and Jernej Ule. 2015. "hiCLIP Reveals the in Vivo Atlas of mRNA Secondary Structures Recognized by Staufen 1." *Nature* 519 (7544): 491–94.

Tasic, Bosiljka, Vilas Menon, Thuc Nghi Nguyen, Tae Kyung Kim, Tim Jarsky, Zizhen Yao, Boaz Levi, et al. 2016. "Adult Mouse Cortical Cell Taxonomy Revealed by Single Cell Transcriptomics." *Nature Neuroscience* 19 (2): 335–46.

Tollervey, James R., Tomaž Curk, Boris Rogelj, Michael Briese, Matteo Cereda, Melis Kayikci, Julian König, et al. 2011. "Characterizing the RNA Targets and Position-Dependent Splicing Regulation by TDP-43." *Nature Neuroscience* 14 (4): 452–58.

Ule, Jernej, Kirk B. Jensen, Matteo Ruggiu, Aldo Mele, Aljaz Ule, and Robert B. Darnell. 2003. "CLIP Identifies Nova-Regulated RNA Networks in the Brain." *Science* 302 (5648): 1212–15.

Ule, Jernej, Giovanni Stefani, Aldo Mele, Matteo Ruggiu, Xuning Wang, Bahar Taneri, Terry Gaasterland, Benjamin J. Blencowe, and Robert B. Darnell. 2006. "An RNA Map Predicting Nova-Dependent Splicing Regulation." *Nature* 444 (7119): 580–86.

Ule, Jernej, Aljaz Ule, Joanna Spencer, Alan Williams, Jing-Shan Hu, Melissa Cline, Hui

Wang, et al. 2005. "Nova Regulates Brain-Specific Splicing to Shape the Synapse." *Nature Genetics* 37 (8): 844–52.

Vanharanta, Sakari, Christina B. Marney, Weiping Shu, Manuel Valiente, Yilong Zou, Aldo Mele, Robert B. Darnell, and Joan Massagué. 2014. "Loss of the Multifunctional RNA-Binding Protein RBM47 as a Source of Selectable Metastatic Traits in Breast Cancer." *eLife* 3 (June). https://doi.org/10.7554/eLife.02734.

Van Nostrand, Eric L., Gabriel A. Pratt, Alexander A. Shishkin, Chelsea Gelboin-Burkhart, Mark Y. Fang, Balaji Sundararaman, Steven M. Blue, et al. 2016. "Robust Transcriptome-Wide Discovery of RNA-Binding Protein Binding Sites with Enhanced CLIP (eCLIP)." *Nature Methods* 13 (6): 508–14.

Weyn-Vanhentenryck, Sebastien M., Aldo Mele, Qinghong Yan, Shuying Sun, Natalie Farny, Zuo Zhang, Chenghai Xue, et al. 2014. "HITS-CLIP and Integrative Modeling Define the Rbfox Splicing-Regulatory Network Linked to Brain Development and Autism." *Cell Reports* 6 (6): 1139–52.

Witten, Joshua T., and Jernej Ule. 2011. "Understanding Splicing Regulation through RNA Splicing Maps." *Trends in Genetics: TIG* 27 (3): 89–97.

Wright, Addison V., James K. Nuñez, and Jennifer A. Doudna. 2016. "Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering." *Cell* 164 (1-2): 29–44.

Yano, Masato, Yoshika Hayakawa-Yano, Aldo Mele, and Robert B. Darnell. 2010. "Nova2 Regulates Neuronal Migration through an RNA Switch in Disabled-1 Signaling." *Neuron* 66 (6): 848–58.

Yap, Karen, Yixin Xiao, Brad A. Friedman, H. Shawn Je, and Eugene V. Makeyev. 2016. "Polarizing the Neuron through Sustained Co-Expression of Alternatively Spliced Isoforms." *Cell Reports* 15 (6): 1316–28.

Zarnack, Kathi, Julian König, Mojca Tajnik, Iñigo Martincorena, Sebastian Eustermann, Isabelle Stévant, Alejandro Reyes, Simon Anders, Nicholas M. Luscombe, and Jernej Ule. 2013. "Direct Competition between hnRNP C and U2AF65 Protects the Transcriptome from the Exonization of Alu Elements." *Cell* 152 (3): 453–66.

Zarnegar, Brian J., Ryan A. Flynn, Ying Shen, Brian T. Do, Howard Y. Chang, and Paul A. Khavari. 2016. "irCLIP Platform for Efficient Characterization of Protein-RNA Interactions." *Nature Methods* 13 (6): 489–92.

Zhang, Chaolin, and Robert B. Darnell. 2011. "Mapping in Vivo Protein-RNA Interactions at Single-Nucleotide Resolution from HITS-CLIP Data." *Nature Biotechnology* 29 (7): 607–14.

Zhang, Chaolin, Maria A. Frias, Aldo Mele, Matteo Ruggiu, Taesun Eom, Christina B. Marney, Huidong Wang, Donny D. Licatalosi, John J. Fak, and Robert B. Darnell. 2010. "Integrative Modeling Defines the Nova Splicing-Regulatory Network and Its Combinatorial Controls." *Science* 329 (5990): 439–43.

Zhang, Xiaochang, Ming Hui Chen, Xuebing Wu, Andrew Kodani, Jean Fan, Ryan Doan, Manabu Ozawa, et al. 2016. "Cell-Type-Specific Alternative Splicing Governs Cell Fate in the Developing Cerebral Cortex." *Cell* 166 (5): 1147–62.e15.