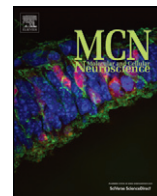




Contents lists available at SciVerse ScienceDirect

Molecular and Cellular Neuroscience

journal homepage: www.elsevier.com/locate/ymcne

CLIPing the brain: Studies of protein–RNA interactions important for neurodegenerative disorders[☆]

Miha Modic^{a,b}, Jernej Ule^{a,c,*}, Christopher R. Sibley^{a,c,*}

^a MRC Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 0QH, UK

^b Gene Center and Department of Biochemistry, Ludwig-Maximilians-Universität München, Feodor-Lynen-Strasse 25, 81377 Munich, Germany

^c Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK

ARTICLE INFO

Article history:

Received 14 December 2012

Accepted 3 April 2013

Available online xxxx

Keywords:

CLIP

Neurodegeneration

FUS

Muscleblind

RNA binding proteins

ABSTRACT

The fate of an mRNA is largely determined by its interactions with RNA binding proteins (RBPs). Post-transcriptional processing, RNA stability, localisation and translation are some of the events regulated by the plethora of RBPs present within cells. Mutations in various RBPs cause several diseases of the central nervous system, including frontotemporal lobar degeneration, amyotrophic lateral sclerosis and fragile X syndrome. Here we review the studies that integrated UV-induced cross-linked immunoprecipitation (CLIP) with other genome-wide methods to comprehensively characterise the function of diverse RBPs in the brain. We discuss the technical challenges of these studies and review the strategies that can be used to reliably identify the RNAs bound and regulated by an RBP. We conclude by highlighting how CLIP and related techniques have been instrumental in addressing the role of RBPs in neurologic diseases.

This article is part of a Special Issue entitled: RNA and splicing regulation in neurodegeneration.

Crown Copyright © 2013 Published by Elsevier Inc. All rights reserved.

Introduction

As an mRNA progresses through the various regulatory stages, from transcription to translation and degradation, it is associated with a myriad of multifunctional RNA-binding proteins (RBPs). Indeed, recent studies suggest that hundreds of RBPs interact with mRNAs to collectively determine their fate (Baltz et al., 2012; Castello et al., 2012). For example, many RBPs are components of multi-protein complexes such as the spliceosome or ribosome that mediate specific aspects of gene expression. Others can directly interact with pre-mRNAs, mRNAs and other non-coding RNAs to control their processing, localisation, stability and structure. The importance of RBPs is further highlighted by the fact that aberrant RBP interactions with RNA can lead to several human pathologies, including neurodegenerative diseases, making their study of great importance. Here we briefly discuss what the genome-wide technologies have taught us about post-transcriptional regulation by RBPs that are implicated in neurodegenerative disease.

Identifying the RNA targets of an RBP

A crucial challenge of researchers studying the functions of RBPs is to identify the RNAs regulated by a specific RBP. CLIP enables study of

RBP–RNA interactions that occur in intact cells or tissues since the biological samples are UV cross-linked, forming a covalent bond between the RNAs in direct contact with the RBP. The RBP-of-interest is then immuno precipitated, and the co-purified RNA is ligated to adapters that enable production of cDNA libraries for sequencing (Ule et al., 2003) (Fig. 1A). CLIP allows unbiased transcriptome-wide analysis of protein–RNA interactions, which is a crucial step towards unravelling the mechanisms of post-transcriptional regulatory networks. However, it is important to appreciate the challenges of CLIP, and the need to integrate it with other technologies.

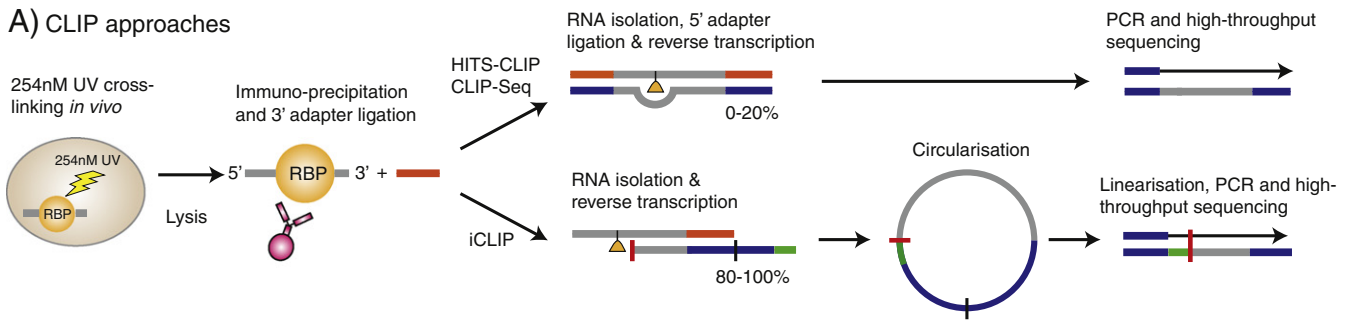
A common expectation from the CLIP studies is that they will directly generate a shortlist of the functionally relevant RNA targets, enabling the study of pathological relevance of these RNAs and the development of new RNA-based therapies. CLIP is clearly a step towards this goal. However, each RBP can regulate a great multitude of RNAs by diverse mechanisms which presents a greater challenge than appreciated at first sight. The frequency of CLIP sequence reads that identify a specific site depends not only on the affinity of the RBP for this site, but also on the RNA abundance. It is clear that RNA abundance varies between genes, and can vary also within each gene due to variable kinetics of transcription, splicing and intronic RNA degradation. It remains challenging to adequately account for these factors in data analysis. The tendency to report the absolute number of RNA targets should thus be approached with caution, since these numbers depend on the variable quantitative quality of CLIP cDNA libraries and require the use of arbitrary binding thresholds. In appreciation of these points, we describe several approaches that can be used to increase the confidence in assigning functionally relevant RNA targets of RBPs.

[☆] This article is part of a Special Issue entitled: RNA and splicing regulation in neurodegeneration.

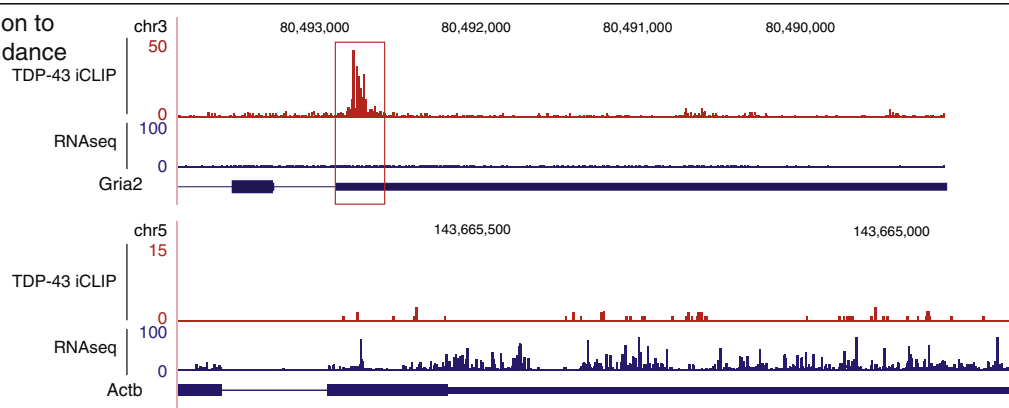
* Corresponding authors at: Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK. Fax: +44 1223213556.

E-mail addresses: j.ule@ucl.ac.uk (J. Ule), c.sibley@ucl.ac.uk (C.R. Sibley).

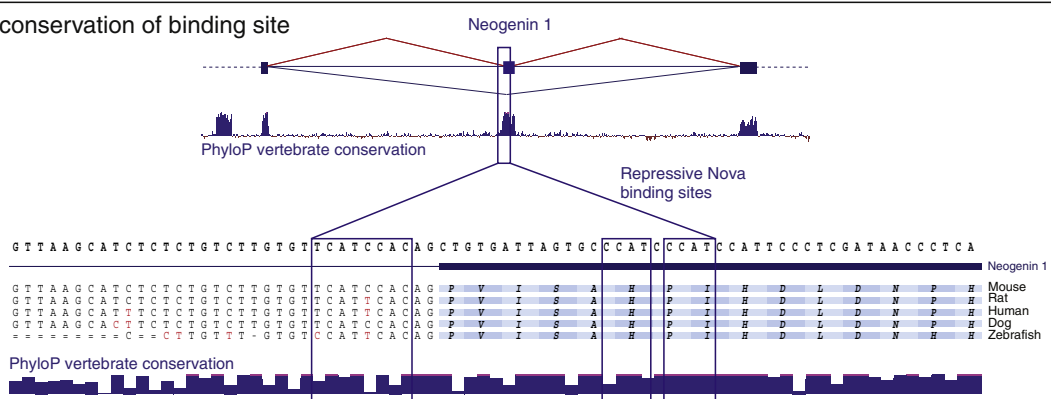
A) CLIP approaches



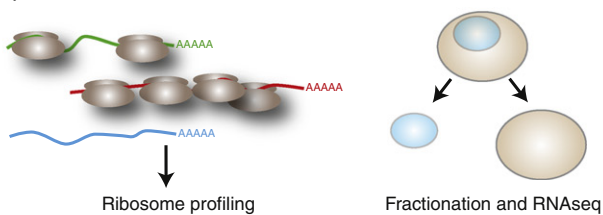
B) Normalisation to transcript abundance



C) Evolutionary conservation of binding site



D) Independent high-throughput approaches



E) RNA maps relating RBP binding to function

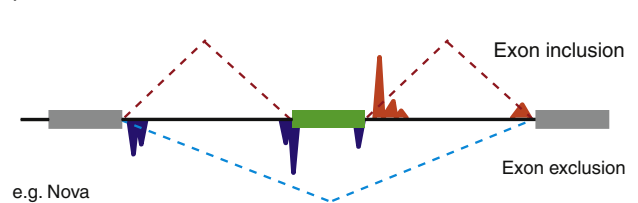


Fig. 1. Identifying RNA targets of an RBP with CLIP. **A)** Schematic representation of HITS-CLIP, CLIP-Seq and iCLIP procedures. Samples are initially cross-linked with 254 nm UV before being lysed and the RBP-of-interest immunoprecipitated together with bound RNA. A 3' adapter is ligated to the RNA and integrity of RBP-RNA complexes examined following isotopic labelling and denaturing electrophoresis (not shown). RNA is released from the RBP by proteinase digestion. In standard CLIP protocol (used in HITS-CLIP, CLIP-Seq and PAR-CLIP), a 5' adapter is ligated to the released RNA – therefore, only the cDNAs that pass across the cross-link site can be identified. In iCLIP, the second adapter is included as an overhang on the primer used for reverse transcription. Circularisation of cDNAs, followed by linearisation, enables iCLIP to identify the cDNAs truncating at the cross-link site. **B)** Normalisation to RNAseq reveals enriched clusters of TDP-43 binding to the *Gria2* 3'UTR above background in the embryonic day 18 mouse brain (red box), and shows that binding to the 3'UTR of the β -actin transcript is an artefact of high RNA abundance. Scales on y-axis represent number of CLIP/RNAseq tags detected. **C)** Evolutionary conservation of repressive Nova binding sites around an alternative exon of *Neogenin 1*. **D)** Confidence in the identification of CLIP targets can be improved when CLIP data is compared to other high-throughput approaches in which the RBP is manipulated. This can include ribosome profiling and knockdown RNAseq analysis of different cell fractions to assess changes in locations. **E)** CLIP data can be integrated into RNA maps which predict how RBP binding determines mechanistic outcomes on an RNA target.

Produce CLIP data with high resolution and quantitative nature

The original CLIP method (Ule et al., 2003), which was used also in HITS-CLIP and CLIP-Seq protocols (Licatalosi et al., 2008; Yeo et al.,

2009), had not been optimised for the purpose of high-throughput sequencing. Therefore, when a limited amount of starting material is available, this method may generate cDNA libraries that contain a limited number of unique sequence reads. UV cross-linking employed by all

CLIP methods leads to formation of a covalent bond between the RNA and protein. Upon proteinase digestion to release the RNA after immunoprecipitation, a short poly-peptide is left at the cross-link site which subsequently stalls reverse transcription in 80–100% of cDNAs (Sugimoto et al., 2012). The original CLIP method relies on read-through across the deposited polypeptide during reverse transcription into a 5' adapter, and therefore doesn't identify the majority of cDNAs that truncate at cross-link sites. This loss is avoided by the individual nucleotide resolution CLIP (iCLIP), a method that captures the truncated cDNAs resulting from stalled reverse transcription and thereby enables a more efficient and comprehensive analysis of protein–RNA interactions (Konig et al., 2010) (Fig. 1A). Importantly, the truncation sites in iCLIP enable nucleotide-resolution analysis of cross-link sites.

The challenge to identify functionally relevant RBP targets of an RBP is particularly great for RBPs that bind to pre-mRNAs in the nucleus. Due to the great length of pre-mRNAs, a binding site for an RBP is present in most expressed pre-mRNAs. If sequencing is carried out to sufficient depth, CLIP studies typically identify multiple binding sites of each RBP in most expressed pre-mRNAs. It is therefore crucial that CLIP data are quantitative, in order to identify discrete regions where high cross-linking density corresponds to high-affinity binding. iCLIP was developed to preserve the quantitative information by using primers for reverse transcription that contain a random barcode. Analysis of this random sequence enables computational filtering of high-throughput sequencing data to remove the artefacts caused by variable PCR amplification of different sequences (Sugimoto et al., 2012).

An alternative method to increase the efficiency of CLIP employs photoactivatable ribonucleoside enhanced cross-linking and immunoprecipitation (PAR-CLIP) (Hafner et al., 2010). However, this method has yet to be used in the brain due to challenges facing the required incorporation of synthetic nucleosides *in vivo*. A recent review on these technologies provides a more detailed discussion on the quality-control steps that can be taken when using these and other high-throughput studies for studying protein–RNA interactions (Konig et al., 2011).

Produce a rank-ordered list of RNA-abundance normalised data

A crucial step in the analysis of CLIP data is to identify sites where RBPs bind with high-affinity, which can be identified by clusters of reads (in CLIP) or clusters of cross-link events (in iCLIP or PAR-CLIP). This in turn acts to reduce the prediction of false positive RNA targets. The most common approach used in the past studies is to calculate the probability that a cluster of reads that is present in a specific genomic region (such as an intron or a gene) is statistically significant when real CLIP data are compared to randomised CLIP data within the same region (Konig et al., 2010; Yeo et al., 2009). This approach is particularly appropriate for studying intronic binding events, since intronic RNA abundance has rarely been comprehensively quantified with current high-throughput sequencing methods. One way to approximately measure intronic abundance is to use the average intronic read density of CLIP data.

To normalise mRNA abundance for RBPs that bind to mRNAs, mRNA sequencing (RNAseq) should be carried out simultaneously with CLIP experiments, and used to normalise the number of CLIP reads by the RNAseq reads (Kishore et al. 2011) (Fig. 1B). The primary objective of this approach is to filter the regions with enriched CLIP reads from those which have widespread CLIP reads due to high RNA abundance. This normalisation is particularly important when CLIPing from heterogenous populations of cells in the brain, where expression of mRNA targets can vary greatly, and often the mRNAs with low expression have important cell-type specific functions. The CLIP/RNAseq ratio can then be used to produce a rank-ordered list of RNAs (Darnell et al., 2011; Konig et al., 2010; Polymenidou et al., 2011). Past studies have defined groups of high, medium and low

confidence targets for further analyses (Darnell et al., 2011; Konig et al., 2010; Polymenidou et al., 2011). An alternative approach would be to avoid the arbitrary thresholds, and instead use rank-ordered lists as in the Sylamer method, which detects microRNA target sites from a rank-ordered list of genes (van Dongen et al., 2008). A similar approach has recently been used to evaluate the likely mRNA targets of FMRP (Ascano et al., 2012).

Use comparative genomics to identify the conserved binding sites

Evolutionary conservation of the sequence motifs bound by the protein can support the functional importance of the RNA binding site (Fig. 1C). For instance, analysis of the RNA motifs that confer Nova binding around its regulated exons demonstrated their high conservation between vertebrate species in CLIP identified targets (Jelen et al., 2007). It was found that the most highly conserved Nova binding sites were in genes with indispensable functions in synaptic development. Nevertheless, it must be considered that the analysis of conserved sites will bias against the sites that are functional and contribute to phenotypic diversity within and between species.

Use independent functional methods

Genome-wide studies have revealed that RBPs often control diverse aspects of post-transcriptional regulation. Initial biochemical studies were based on a small number of regulated RNAs, therefore RBPs that regulate splicing were often classified as splicing repressors or enhancers. Studies of Nova showed that it can enhance or repress exon inclusion in many different RNA targets (Ule et al., 2005), and this has now been seen also for most other RBPs within the subset that are implicated in splicing regulation (Witten and Ule, 2011). Moreover, Nova governs the usage of alternative polyadenylation sites (Licatalosi et al., 2008), and can also control mRNA localisation (Racca et al., 2010). Similarly, a comprehensive functional evaluation of Mbnl1 demonstrated its functions in pre-mRNA processing, translation and control of RNA localisation (Wang et al., 2012). Specifically it was found that Mbnl1 bound transcripts showed a change in ribosomal footprints and redistribution from the membrane fraction to the insoluble fraction upon Mbnl1/Mbnl2 knockdown. Due to this multifunctionality of RBPs, the full functional validation of CLIP datasets will remain a challenge in the years to come. Integration of CLIP with multiple genome-wide methods including RNAseq, ribosome profiling (Ingolia et al., 2011), mRNA localisation studies (Wang et al., 2012) and others will be required to determine the full set of functionally relevant RNA targets of an RBP (Fig. 1D).

Define position-dependent regulatory principles, or “RNA maps”

Integration of CLIP data with analysis of global changes of splicing has been used to identify the global positional principles governing the splicing functions of Nova (Licatalosi et al., 2008; Ule et al., 2006). The combined findings excitingly identified a position-dependent regulation of these transcripts by Nova proteins and were used to define an RNA map in which exon exclusion is facilitated by binding close to branch points, splice sites or within exons, and exon inclusion favoured by Nova binding downstream of exons (Fig. 1E). Notably, binding sites close to the flanking exons also mediated predictable outcomes (Ule et al., 2006). Among others, this approach has now been used for TDP-43/TARDBP (herein referred to as TDP-43) (Polymenidou et al., 2011; Tollervy et al., 2011a), TIA1/TIAL1 (Wang et al., 2010), FUS/TLS (herein referred to as FUS) (Lagier-Tourenne et al., 2012; Rogelj et al., 2012), FOX2 (Yeo et al., 2009), PTBP1 (Xue et al., 2009) and Mbnl1 (Wang et al., 2012), and has shown that many RBPs have common positional principles in the regulation of splicing. Moreover, an RNA map also underlined the position-dependent activity of Nova proteins in alternative polyadenylation (Licatalosi et al., 2008).

The integrative analysis has three main advantages. Firstly, the RNA map can validate the high quality of data by demonstrating that binding position correlates with the effect of the protein (enhanced vs. repressed), and that binding at these positions in the regulated RNAs is significantly enriched relative to the control RNAs. Secondly, since the RNA map defines the positions where RBPs bind to have functional impact, it enables separation of those changes observed in RNAseq data that are the result of direct effects of the RBP from those that may result from indirect effects. This is particularly important for studies in the brain, since manipulation of RBPs can lead to developmental defects, cell stress, changes in neuronal networks or glia–neuron interactions, siRNA off-targets, and other effects that can indirectly change gene expression. Thirdly, the positional pattern in the RNA map can reveal new regulatory mechanisms employed by the RBPs. For instance, RNA maps demonstrated that the enhancer elements are most often located immediately downstream of exons. Moreover, they showed that certain RBPs can also regulate splicing via distal binding sites (Witten and Ule, 2011).

CLIPing the brain in health and disease

To date, RNA binding profiles of several RBPs involved in neurologic disorders have been investigated either in cell culture, transgenic animal models, or human post-mortem tissue (Table 1). Here we choose representative RBPs to demonstrate how CLIP studies helped understand their activity in the CNS, and thereby provide the foundation for understanding their role in disease pathologies.

Nova

Paraneoplastic opsoclonus myoclonus ataxia is a neurological disorder mediated by autoimmune attack against onconeural disease antigens including Nova-1 (Buckanovich et al., 1993) and Nova-2 (Yang et al., 1998), and was the study of initial CLIP experiments. Low-throughput (Ule et al., 2003) and, more recently, high-throughput (Licatalosi et al., 2008) sequencing of Nova CLIP libraries revealed the RNA map in which exon exclusion is facilitated by binding close to

branch points, splice sites or within exon, and exon inclusion favoured by binding downstream of exons (Ule et al., 2006). Activity is specifically directed by binding to YCAY clusters within target genes, with nuclear-cytoplasmic shuttling of certain targets possible due to Nova's nuclear localisation and nuclear export sequences (Racca et al., 2010). This has been followed up by showing that Nova can act outside of splicing to regulate alternative polyadenylation through binding of clusters flanking poly-A sites, although the relevance to disease remains unknown (Licatalosi et al., 2008). By identifying the mRNAs that are directly regulated by Nova proteins (Licatalosi et al., 2008; Ule et al., 2006), the Nova RNA map enabled studies that revealed how specific mRNAs mediate the function of Nova proteins in the brain. In particular, defective splicing of *Agrn* and *Dab1* leads to defects in the formation of the neuromuscular junction and in neuronal migration in the CNS of the *Nova1*^{-/-}/*Nova2*^{-/-} mice, respectively (Ruggiu et al., 2009; Yano et al., 2010).

FMRP

Loss of function of fragile-X mental retardation protein (FMRP) is responsible for fragile-X syndrome, the most common inherited form of mental retardation. FMRP is an important regulator of translation (Khandjian et al., 2004; Lagerbauer et al., 2001; Li et al., 2001). A study using CLIP in mouse brain reported that FMRP binds the coding sequence of mRNAs with little sequence specificity. By uncovering FMRP interacting targets in polyribosomal enriched samples, it was shown that FMRP binding is enriched on open reading frames of a collection of mRNAs encoding synaptic transcripts, including multiple receptor complexes and G-protein signalling pathways (Darnell et al., 2011). A recent study using FMRP PAR-CLIP in HEK293 cells additionally reported two binding motifs recognised by FMRP: ACUK and WGGA (in which K = G or U and W = A or U). An integration of PAR-CLIP data with motif analysis was used to determine a ranked enrichment of FMRP mRNA targets. Interestingly, 93 of the top ranked genes are implicated in autism spectrum disorders (Angelman, Prader–Willi, Rett, and Cornelia de Lange syndromes). FMRP was found to regulate protein levels of several of these genes in

Table 1
CLIP studies on RBPs implicated in neurological function in health and disease. Abbreviations: NOVA1/2 – neuro-oncological ventral antigen 1/2, TARBP – TAR DNA binding protein, FMRP – fragile-X mental retardation protein, MBNL1/2 – muscleblind-like protein 1/2, FUS/TLS – fused in sarcoma/translocated in liposarcoma, PTBP2 – polypyrimidine tract binding protein 2, PARK7 – Parkinson protein 7, ELAVL1 – embryonic lethal, abnormal vision, Drosophila-like 1, CELF4 – CUGBP Elav-like family member 4.

Symbol	Disease	Key findings	Reference
CELF4	Epilepsy Hyperactivity	Binds UGU motifs in 3'UTRs. Controls stability of mRNAs encoding synaptic proteins.	Wagnon et al. (2012)
ELAVL1	Epilepsy	Recognises U-rich stretches interspersed with Gs. Regulates transcript stability. Controls the synthesis of glutamate.	Ince-Dunn et al. (2012)
FMR1	Fragile-X mental retardation Autism spectrum disorders	Represses the translation of target mRNAs. Preferred binding to the coding region of exons. Increased association with transcripts encoding synaptic proteins. Binds ACUK and WGGA (in which K = G or U and W = A or U) motifs.	Darnell et al. (2011) Ascano et al. (2012)
FUS	Frontotemporal lobar degeneration, amyotrophic lateral sclerosis	Binds along the full length of pre-mRNAs. Regulates alternative splicing of many neuronal development genes. Knockdown leads to decreased expression of long genes in the brain.	Ishigaki et al. (2012) Rogelj et al. (2012) Lagier-Tourenne et al. (2012)
MBNL1/2	Myotonic dystrophy (DM)	Recognises UGC or GCU-containing 4-mer clusters. Regulates DM-related alternative splicing. Contributes to mRNA localisation and translation by binding to 3'UTRs.	Wang et al. (2012) Charzaniis et al. (2012)
NOVA1/2	Paraneoplastic opsoclonus-myoclonus-ataxia (POMA)	Binds YCAY clusters to regulate alternative splicing. Controls synaptogenesis and neuronal migration via specific mRNAs. Regulates alternative poly-adenylation in the brain.	Licatalosi et al. (2008) Ruggiu et al. (2009) Yano et al. (2010)
PARK7	Parkinson's disease	Recognises CC/GG rich regions. Inhibits translation of target mRNAs.	van der Brug et al. (2008)
PTBP2		Recognises UCU-rich motifs to regulate alternative splicing. Regulates neural stem cell polarity in developing brain. Involved in mRNA trafficking stability and translation.	Licatalosi et al. (2012)
TARDBP	Frontotemporal lobar degeneration, amyotrophic lateral sclerosis	Recognises UG repeats and UG-rich motifs in introns and 3' UTRs. Regulates alternative splicing of many neuronal development genes. Knockdown leads to decreased expression of long genes in the brain.	Tollervy et al. (2011a) Polymenidou et al. (2011)

human cell culture, mouse ovaries and human brain (Ascano et al., 2012). Interestingly, a recent study of CELF4, another neuronal RBP, showed that it shares more than 30% of its RNA targets with FMRP, and is even more strongly enriched in transcripts linked to autism spectrum disorders (Wagnon et al., 2012).

Mbnl

Muscleblind-like proteins (Mbnl1 and Mbnl2) are sequestered in various tissues within the neuromuscular diseases of myotonic dystrophy types 1 (DM1) and 2 (DM2) in response to expansions of CUG repeats in the 3'UTR of the myotonic dystrophy protein kinase (DMPK) (Harley et al., 1992; Mahadevan et al., 1992), or CCUG repeats in the zinc finger protein 9 (ZNF9) (Ranum et al., 1998), respectively. Previously known to be regulators of splicing, two groups have now integrated CLIP, RNAseq and ribosome profiling to confirm that nuclear Mbnl proteins bind introns, coding sequences and 3'UTRs at UGC or GCU-containing 4-mers to dictate splicing changes and identify direct targets for splicing regulation based on CLIP-identified binding (Charizanis et al., 2012; Wang et al., 2012). The RNA map reveals that Mbnl proteins bind in upstream introns and alternative exons to promote exclusion, and in downstream introns to promote inclusion of exons. Further, a previously described role of Mbnl proteins in regulating localisation in the cytoplasm (Adereth et al., 2005) was extended transcriptome-wide, whilst the aforementioned regulation of the translation state of 3'UTRs carrying Mbnl1 binding sites provides further insight into disease mechanisms when this protein is sequestered (Wang et al., 2012). It will now be important to assess if these perturbations in patient samples with CUG repeats correlate with disease severity.

FUS and TDP-43

ALS and FTLN are two diseases with phenotypic and mechanistic overlap (Ferrari et al., 2011; Lagier-Tourenne and Cleveland, 2009). Familial mutations to TDP-43 and FUS have been identified in each disease (Broustal et al., 2010; Kabashi et al., 2008; Kwiatkowski et al., 2009; Van Langenhove et al., 2010; Vance et al., 2009), and these proteins aggregate abnormally in neurons during disease progression. Despite these common links, CLIP studies of TDP-43 and FUS did not find a significant overlap in their RNA binding sites. Whereas TDP-43 has a strong specificity to bind distinct UG-rich RNA regions (Polymenidou et al., 2011; Tollervey et al., 2011a), FUS binds with mild specificity to G-rich motifs along the full-length of pre-mRNA transcripts (Ishigaki et al., 2012; Lagier-Tourenne et al., 2012; Rogelj et al., 2012).

Another observation made by FUS CLIP was a saw-tooth binding pattern along long introns, similar to that seen in total RNAseq (Ameur et al., 2011; Lagier-Tourenne et al., 2012; Rogelj et al., 2012). However, whether this binding profile is distinct and disease-relevant remains a matter of debate. This pattern is readily observed for FUS in long introns (Ameur et al., 2011; Lagier-Tourenne et al., 2012), but can also be observed in shorter introns using the increased coverage and quantitative ability of the iCLIP method (Rogelj et al., 2012). Although suggested to be unique to FUS (Lagier-Tourenne et al., 2012), iCLIP analysis showed that TDP-43, U2AF65 and FUS all had a linear increase in crosslinking as the distance from the 3' splice site increases, similar to the increase in total RNAseq (Rogelj et al., 2012). This suggests that the saw-tooth pattern may be an indirect result of the variable abundance of the intronic RNA, which increases linearly with the distance from the 3' splice site.

Interestingly, FUS and TDP-43 regulate splicing of different alternative exons (Lagier-Tourenne et al., 2012; Rogelj et al., 2012). However, analysis of RNAseq data showed an overlap in the long genes, which were sensitive to loss of function of FUS and TDP-43 in transgenic models and in neurons derived from patients with familial

mutations to FUS (Lagier-Tourenne et al., 2012; Polymenidou et al., 2011). It will be important to understand the mechanisms behind changes in expression of long genes, and their role in disease pathology. Whereas the effect on long genes may involve novel transcriptional functions of TDP-43 and FUS, it could also result from accumulated defects in pre-mRNA processing. RBPs can promote stability of mRNAs by preventing recognition of cryptic exons; for example, knockdown of *hnRNPC* leads to widespread expression of cryptic exons, which can result in decreased gene expression. In the absence of hnRNP C, cryptic exons are recognised by U2AF65, a splicing factor that initiates the recognition of 3' splice sites (Zarnack et al., 2012). RNA splicing maps demonstrate that similar to hnRNP C, FUS and TDP-43 bind close to 3' splice sites to repress splicing, where they might displace U2AF65 (Rogelj et al., 2012; Tollervey et al., 2011a). Due to the saw-tooth pattern, U2AF65 has increased binding in long introns (Rogelj et al., 2012), which could increase the danger of recognising cryptic exons, and thereby decreasing expression of long genes. Aberrant binding of U2AF65 could be caused by loss of other RBPs that regulate splicing in the brain. Therefore, it will be important to study gene expression in transgenic models of other RBPs to assess if the effect on long genes is specific to TDP-43 and FUS.

Future considerations for CLIP studies in the CNS

CLIP and its variants are now the primary method for studying protein–RNA interactions and we have described how several early high-throughput studies have provided invaluable insights into the mechanisms of neurodegenerative diseases. This includes the identification of putative RNA targets and understanding the way in which an RBP acts through looking at its specificity and the regions bound within a transcript. It is likely that CLIP-based approaches will remain popular, in part due to recent method developments making it a more robust, and also following the recent identification of many new candidate RBPs using the global crosslinking-based purification (Baltz et al., 2012; Castello et al., 2012).

We explain the limitations in reporting a precise number of RNA targets of an RBP based on CLIP. CLIP tags alone are not sufficient to determine the functionally relevant binding sites of RBPs. Therefore it is important to integrate CLIP with other approaches in order to identify the RNAs that are bound and directly regulated by the RBPs. For RBPs interacting with mRNAs we encourage the use of normalisation of CLIP clusters to transcript abundance either using RNAseq, whereas other approaches are necessary for CLIP data normalisation of RBPs binding pre-mRNAs. Evaluation of evolutionary conservation, integration of complementary high-throughput methods, and generation of RNA maps can help to unify the identification of functionally relevant RNA targets for subsequent studies.

One application of CLIP that remains to be widely applied is the study of protein–RNA interactions in human postmortem tissues (Tollervey et al., 2011a). It is important to appreciate that cell stress and neuronal loss can obscure the causative mechanisms in postmortem tissues. This has become particularly apparent from the study of gene expression of FTLN and Alzheimer disease temporal cortex, which found widespread yet identical changes in gene expression (Tollervey et al., 2011b). Whilst CLIP techniques are compatible with partially degraded postmortem tissue, it is important to appreciate that there are many variables in sample preparation that can affect the quality of the sample. This can include the length to autopsy and sample collection, as well as the fixation and freezing methods employed. Therefore special care needs to be taken when analysing data produced from samples collected under variable conditions.

The studies of postmortem tissues are also inevitably correlative, since it is not possible to go back to the human tissue to manipulate the function of the RBP. Therefore, it is important to additionally study the relevant protein–RNA interactions in disease model systems. Transgenic animal models have been used most often but a

new opportunity is arising in the study of pluripotent stem cell (PSC) disease models (Patani et al., 2012). Induced PSCs (iPSCs) of TDP-43 M337V mutation have now been generated which have elevated levels of TDP-43 protein and decreased survival following neuronal differentiation (Bilican et al., 2012), whilst FUS has already been knocked down in differentiated neural stem cells and displayed comparative changes to both transgenic models and patient tissue (Lagier-Tourenne et al., 2012). If disease phenotype continues to be recapitulated for other disease-associated proteins, then iPSCs will represent an invaluable sample source for future CLIP studies.

Over the next years, we can expect more CLIP studies to assess RNA interactions of multiple RBPs at a specific stage of brain development, which will enable comprehensive comparative studies. It will be possible at that point to define the phenomena specific to RBPs involved in neurodegenerative disorders, and separate these from the more general aspects of protein–RNA interactions.

Acknowledgements

The authors thank all members of the Ule lab for discussion of the manuscript. We would like to thank Nejc Haberman for computational identification of enriched CLIP clusters and Dr Michael Briese for supplying embryonic mouse brain RNAseq data. This work was supported by Likar Foundation, IMPRS and DAAD fellowship to M.M, an MRC career development fellowship to C.R.S and the European Research Council (grant number 206726-CLIP), and the Medical Research Council (grant number U105185858).

References

- Adereth, Y., Dammai, V., Kose, N., Li, R., Hsu, T., 2005. RNA-dependent integrin alpha3 protein localization regulated by the Muscleblind-like protein MLP1. *Nat. Cell Biol.* 7, 1240–1247.
- Ameur, A., Zaghool, A., Halvardson, J., Wetterbom, A., Gyllensten, U., Cavelier, L., Feuk, L., 2011. Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat. Struct. Mol. Biol.* 18, 1435–1440.
- Ascano Jr., M., Mukherjee, N., Bandaru, P., Miller, J.B., Nusbaum, J.D., Corcoran, D.L., Langlois, C., Munschauer, M., Dewell, S., Hafner, M., Williams, Z., Ohler, U., Tuschl, T., 2012. FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature* 492, 382–386.
- Baltz, A.G., Munschauer, M., Schwannhauser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., Wyler, E., Bonneau, R., Selbach, M., Dieterich, C., Landthaler, M., 2012. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* 46, 674–690.
- Bilican, B., Serio, A., Barmada, S.J., Nishimura, A.L., Sullivan, G.J., Carrasco, M., Phatnani, H.P., Puddifoot, C.A., Story, D., Fletcher, J., Park, I.H., Friedman, B.A., Daley, G.Q., Wyllie, D.J., Hardingham, G.E., Wilmut, I., Finkbeiner, S., Maniatis, T., Shaw, C.E., Chandran, S., 2012. Mutant induced pluripotent stem cell lines recapitulate aspects of TDP-43 proteinopathies and reveal cell-specific vulnerability. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5803–5808.
- Broustal, O., Camuzat, A., Guillot-Noel, L., Guy, N., Millicamps, S., Deffond, D., Lacomblez, L., Golfier, V., Hannequin, D., Salachas, F., Camu, W., Didic, M., Dubois, B., Meininger, V., Le Ber, I., Brice, A., 2010. FUS mutations in frontotemporal lobar degeneration with amyotrophic lateral sclerosis. *J. Alzheimers Dis.* 22, 765–769.
- Buckanovich, R.J., Posner, J.B., Darnell, R.B., 1993. Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. *Neuron* 11, 657–672.
- Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., Krijgsvelde, J., Hentze, M.W., 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393–1406.
- Charizanis, K., Lee, K.Y., Batra, R., Goodwin, M., Zhang, C., Yuan, Y., Shiue, L., Cline, M., Scotti, M.M., Xia, G., Kumar, A., Ashizawa, T., Clark, H.B., Kimura, T., Takahashi, M.P., Fujimura, H., Jinnai, K., Yoshikawa, H., Gomes-Pereira, M., Gourdon, G., Sakai, N., Nishino, S., Foster, T.C., Ares Jr., M., Darnell, R.B., Swanson, M.S., 2012. Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. *Neuron* 75, 437–450.
- Darnell, J.C., Van Driesche, S.J., Zhang, C., Hung, K.Y., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W., Licatalosi, D.D., Richter, J.D., Darnell, R.B., 2011. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* 146, 247–261.
- Ferrari, R., Kapogiannis, D., Huey, E.D., Momeni, P., 2011. FTD and ALS: a tale of two diseases. *Curr. Alzheimer Res.* 8, 273–294.
- Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano Jr., M., Jungkamp, A.C., Munschauer, M., Ulrich, A., Wardle, G.S., Dewell, S., Zavolan, M., Tuschl, T., 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141, 129–141.
- Harley, H.G., Brook, J.D., Rundle, S.A., Crow, S., Reardon, W., Buckler, A.J., Harper, P.S., Housman, D.E., Shaw, D.J., 1992. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* 355, 545–546.
- Ince-Dunn, G., Okano, H.J., Jensen, K.B., Park, W.Y., Zhong, R., Ule, J., Mele, A., Fak, J.J., Yang, C., Zhang, C., Yoo, J., Herre, M., Okano, H., Noebels, J.L., Darnell, R.B., 2012. Neuronal Elav-like (Hu) proteins regulate RNA splicing and abundance to control glutamate levels and neuronal excitability. *Neuron* 75, 1067–1080.
- Ingolia, N.T., Lareau, L.F., Weissman, J.S., 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 147, 789–802.
- Ishigaki, S., Masuda, A., Fujioka, Y., Iguchi, Y., Katsuno, M., Shibata, A., Urano, F., Sobue, G., Ohno, K., 2012. Position-dependent FUS–RNA interactions regulate alternative splicing events and transcriptions. *Sci. Rep.* 2, 529.
- Jelen, N., Ule, J., Zivin, M., Darnell, R.B., 2007. Evolution of Nova-dependent splicing regulation in the brain. *PLoS Genet.* 3, 1838–1847.
- Kabashi, E., Valdmanis, P.N., Dion, P., Spiegelman, D., McConkey, B.J., Vande Velde, C., Bouchard, J.P., Lacomblez, L., Pochigaeva, K., Salachas, F., Pradat, P.F., Camu, W., Meininger, V., Dupre, N., Rouleau, G.A., 2008. TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat. Genet.* 40, 572–574.
- Khandjian, E.W., Huot, M.E., Tremblay, S., Davidovic, L., Mazroui, R., Bardoni, B., 2004. Biochemical evidence for the association of fragile X mental retardation protein with brain polyribosomal ribonucleoproteins. *Proc. Natl. Acad. Sci. U. S. A.* 101, 13357–13362.
- Kishore, S., Jaskiewicz, L., Burger, L., Hausser, J., Khorshid, M., Zavolan, M., 2011. A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nat. Methods.* 8 (7), 559–564.
- Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., Ule, J., 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* 17, 909–915.
- Konig, J., Zarnack, K., Luscombe, N.M., Ule, J., 2011. Protein–RNA interactions: new genomic technologies and perspectives. *Nat. Rev. Genet.* 13, 77–83.
- Kwiatkowski Jr., T.J., Bosco, D.A., Leclerc, A.L., Tamrazian, E., Vanderburg, C.R., Russ, C., Davis, A., Gilchrist, J., Kasarskis, E.J., Munsat, T., Valdmanis, P., Rouleau, G.A., Hosler, B.A., Cortelli, P., de Jong, P.J., Yoshinaga, Y., Haines, J.L., Pericak-Vance, M.A., Yan, J., Ticozzi, N., Siddique, T., McKenna-Yasek, D., Sapp, P.C., Horvitz, H.R., Landers, J.E., Brown Jr., R.H., 2009. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* 323, 1205–1208.
- Laggerbauer, B., Ostareck, D., Keidel, E.M., Ostareck-Lederer, A., Fischer, U., 2001. Evidence that fragile X mental retardation protein is a negative regulator of translation. *Hum. Mol. Genet.* 10, 329–338.
- Lagier-Tourenne, C., Cleveland, D.W., 2009. Rethinking ALS: the FUS about TDP-43. *Cell* 136, 1001–1004.
- Lagier-Tourenne, C., Polymenidou, M., Hutt, K.R., Vu, A.Q., Baughn, M., Huelga, S.C., Clutario, K.M., Ling, S.C., Liang, T.Y., Mazur, C., Wanciewicz, E., Kim, A.S., Watt, A., Freier, S., Hicks, G.G., Donohue, J.P., Shiue, L., Bennett, C.F., Ravits, J., Cleveland, D.W., Yeo, G.W., 2012. Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat. Neurosci.* 15, 1488–1497.
- Li, Z., Zhang, Y., Ku, L., Wilkinson, K.D., Warren, S.T., Feng, Y., 2001. The fragile X mental retardation protein inhibits translation via interacting with mRNA. *Nucleic Acids Res.* 29, 2276–2283.
- 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., Darnell, J.C., Darnell, R.B., 2008. HTS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 456, 464–469.
- Licatalosi, D.D., Yano, M., Fak, J.J., Mele, A., Grabinski, S.E., Zhang, C., Darnell, R.B., 2012. Ptb2 represses adult-specific splicing to regulate the generation of neuronal precursors in the embryonic brain. *Genes Dev.* 26, 1626–1642.
- Mahadevan, M., Tsilfidis, C., Sabourin, L., Shuttler, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barcelo, J., O'Hoy, K., et al., 1992. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* 255, 1253–1255.
- Patani, R., Sibley, C.R., Chandran, S., Ule, J., 2012. Using human pluripotent stem cells to study post-transcriptional mechanisms of neurodegenerative diseases. *Brain Res.* 1462, 129–138.
- Polymenidou, M., Lagier-Tourenne, C., Hutt, K.R., Huelga, S.C., Moran, J., Liang, T.Y., Ling, S.C., Sun, E., Wanciewicz, E., Mazur, C., Kordasiewicz, H., Sedaghat, Y., Donohue, J.P., Shiue, L., Bennett, C.F., Yeo, G.W., Cleveland, D.W., 2011. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* 14, 459–468.
- Racca, C., Gardiol, A., Eom, T., Ule, J., Triller, A., Darnell, R.B., 2010. The neuronal splicing factor Nova co-localizes with target RNAs in the dendrite. *Front. Neural Circuits* 4, 5.
- Ranum, L.P., Rasmussen, P.F., Benzow, K.A., Koob, M.D., Day, J.W., 1998. Genetic mapping of a second myotonic dystrophy locus. *Nat. Genet.* 19, 196–198.
- Rogelj, B., Easton, L.E., Bogu, G.K., Stanton, L.W., Rot, G., Curk, T., Zupan, B., Sugimoto, Y., Modic, M., Haberman, N., Tollervey, J., Fujii, R., Takumi, T., Shaw, C.E., Ule, J., 2012. Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. *Sci. Rep.* 2, 603.
- Ruggiu, M., Herbst, R., Kim, N., Jevsek, M., Fak, J.J., Mann, M.A., Fischbach, G., Burden, S.J., Darnell, R.B., 2009. Rescuing Z + agrin splicing in Nova null mice restores synapse formation and unmasks a physiologic defect in motor neuron firing. *Proc. Natl. Acad. Sci. U. S. A.* 106, 3513–3518.
- Sugimoto, Y., Konig, J., Hussain, S., Zupan, B., Curk, T., Frye, M., Ule, J., 2012. Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein–RNA interactions. *Genome Biol.* 13, R67.
- Tollervey, J.R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., Konig, J., Hortobagyi, T., Nishimura, A.L., Zupunski, V., Patani, R., Chandran, S., Rot, G.,

- Zupan, B., Shaw, C.E., Ule, J., 2011a. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* 14, 452–458.
- Tollervey, J.R., Wang, Z., Hortobagyi, T., Witten, J.T., Zarnack, K., Kayikci, M., Clark, T.A., Schweitzer, A.C., Rot, G., Curk, T., Zupan, B., Rogelj, B., Shaw, C.E., Ule, J., 2011b. Analysis of alternative splicing associated with aging and neurodegeneration in the human brain. *Genome Res.* 21, 1572–1582.
- Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., Darnell, R.B., 2003. CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302, 1212–1215.
- Ule, J., Ule, A., Spencer, J., Williams, A., Hu, J.S., Cline, M., Wang, H., Clark, T., Fraser, C., Ruggiu, M., Zeeberg, B.R., Kane, D., Weinstein, J.N., Blume, J., Darnell, R.B., 2005. Nova regulates brain-specific splicing to shape the synapse. *Nat. Genet.* 37, 844–852.
- Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J., Darnell, R.B., 2006. An RNA map predicting Nova-dependent splicing regulation. *Nature* 444, 580–586.
- van der Brug, M.P., Blackinton, J., Chandran, J., Hao, L.Y., Lal, A., Mazan-Mamczarz, K., Martindale, J., Xie, C., Ahmad, R., Thomas, K.J., Beilina, A., Gibbs, J.R., Ding, J., Myers, A.J., Zhan, M., Cai, H., Bonini, N.M., Gorospe, M., Cookson, M.R., 2008. RNA binding activity of the recessive parkinsonism protein DJ-1 supports involvement in multiple cellular pathways. *Proc. Natl. Acad. Sci. U. S. A.* 105, 10244–10249.
- van Dongen, S., Abreu-Goodger, C., Enright, A.J., 2008. Detecting microRNA binding and siRNA off-target effects from expression data. *Nat. Methods* 5, 1023–1025.
- Van Langenhove, T., van der Zee, J., Slegers, K., Engelborghs, S., Vandenbergh, R., Gijssels, I., Van den Broeck, M., Mattheijssens, M., Peeters, K., De Deyn, P.P., Cruts, M., Van Broeckhoven, C., 2010. Genetic contribution of FUS to frontotemporal lobar degeneration. *Neurology* 74, 366–371.
- Vance, C., Rogelj, B., Hortobagyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P., Ganesalingam, J., Williams, K.L., Tripathi, V., Al-Saraj, S., Al-Chalabi, A., Leigh, P.N., Blair, I.P., Nicholson, G., de Bellerche, J., Gallo, J.M., Miller, C.C., Shaw, C.E., 2009. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323, 1208–1211.
- Wagnon, J.L., Briese, M., Sun, W., Mahaffey, C.L., Curk, T., Rot, G., Ule, J., Frankel, W.N., 2012. CELF4 regulates translation and local abundance of a vast set of mRNAs, including genes associated with regulation of synaptic function. *PLoS Genet.* 8, e1003067.
- Wang, Z., Kayikci, M., Briese, M., Zarnack, K., Luscombe, N.M., Rot, G., Zupan, B., Curk, T., Ule, J., 2010. iCLIP predicts the dual splicing effects of TIA–RNA interactions. *PLoS Biol.* 8, e1000530.
- 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., Lecuyer, E., Burge, C.B., 2012. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* 150, 710–724.
- Witten, J.T., Ule, J., 2011. Understanding splicing regulation through RNA splicing maps. *Trends Genet.* 27, 89–97.
- Xue, Y., Zhou, Y., Wu, T., Zhu, T., Ji, X., Kwon, Y.S., Zhang, C., Yeo, G., Black, D.L., Sun, H., Fu, X.D., Zhang, Y., 2009. Genome-wide analysis of PTB–RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Mol. Cell* 36, 996–1006.
- Yang, Y.Y., Yin, G.L., Darnell, R.B., 1998. The neuronal RNA-binding protein Nova-2 is implicated as the autoantigen targeted in POMA patients with dementia. *Proc. Natl. Acad. Sci. U. S. A.* 95, 13254–13259.
- Yano, M., Hayakawa-Yano, Y., Mele, A., Darnell, R.B., 2010. Nova2 regulates neuronal migration through an RNA switch in disabled-1 signaling. *Neuron* 66, 848–858.
- Yeo, G.W., Coufal, N.G., Liang, T.Y., Peng, G.E., Fu, X.D., Gage, F.H., 2009. An RNA code for the FOX2 splicing regulator revealed by mapping RNA–protein interactions in stem cells. *Nat. Struct. Mol. Biol.* 16, 130–137.
- Zarnack, K., Konig, J., Tajnik, M., Martincorina, I., Stevant, M., Reyes, A., Anders, S., Luscombe, N.M., Ule, J., 2012. Direct competition between hnRNP C and U2AF65 protects the transcriptome from the uncontrolled exonization of Alu elements. *Cell* 152, 453–466.