HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2009 May 27.

HITS-CLIP yields genome-wide insights into brain alternative RNA processing

Donny D. Licatalosi[1], **Aldo Mele**[1], **John J. Fak**[1], **Jernej Ule**[2], **Melis Kayikci**[2], **Sung Wook Chi**[1], **Tyson A. Clark**[3], **Anthony C. Schweitzer**[3], **John E. Blume**[3], **Xuning Wang**[4], **Jennifer C. Darnell**[1], and **Robert B. Darnell**[1]

[1]Laboratory of Molecular Neuro-Oncology and Howard Hughes Medical Institute, The Rockefeller University, 1230 York Ave, NY, NY 10021 USA [2]MRC Laboratory of Molecular Biology, Cambridge, England [3]Expression Research, Affymetrix, Inc. Santa Clara, CA 95051 [4]Biocomputing, Information Technology, The Rockefeller University, 1230 York Ave, NY, NY 10021 USA

Summary

Protein-RNA interactions play critical roles in all aspects of gene expression. Here we develop a genome-wide means of mapping protein-RNA binding sites *in vivo*, by high throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP). HITS-CLIP analysis of the neuron-specific splicing factor Nova2 revealed extremely reproducible RNA binding maps in multiple mouse brains. These maps provide genome-wide *in vivo* biochemical footprints confirming the previous prediction that the position of Nova binding determines the outcome of alternative splicing; moreover, they are sufficiently powerful to predict Nova action *de novo*. HITS-CLIP revealed a large number of Nova-RNA interactions in 3′ UTRs, leading to the discovery that Nova regulates alternative polyadenylation in the brain. HITS-CLIP, therefore, provides a robust, unbiased means to identify functional protein-RNA interactions *in vivo*.

The discovery of RNA molecules with catalytic activity[1] led to the hypothesis that from the earliest life forms, RNA regulation evolved to play critical roles in living organisms[2-5]. Efforts to develop comprehensive understanding of protein-RNA interactions *in vivo* have combined genetics, bioinformatics, microarray-profiling, and biochemical approaches. However, the latter have been hampered by methodologic problems;[6,7] for example, co-immunoprecipitations can lead to re-association of protein-RNA complexes *in vitro*[8], non-specifically bound RNAs, and additional co-precipitating RNA binding proteins (RNABPs)[9].

We have taken a different approach toward understanding protein-RNA interactions by developing a crosslinking protocol that works in tissues, and can therefore be applied prior to protein purification. This method, termed CLIP10,11, uses UV-irradiation to induce covalent crosslinks between protein-RNA complexes *in situ*, allowing rigorous purification of RNABPs along with small fragments of RNA, which can be amplified and sequenced. CLIP has been used to study direct protein-RNA interactions extant in living cells11-13, including identification of RNA targets11 for the KH-type RNABP Nova14,15, and the discovery of hnRNPA1-dependent regulation of a miRNA12.

Genome-wide efforts to understand Nova function, using exon junction microarrays and bioinformatic analysis of Nova binding sites (YCAY clusters15, characterized biochemically16 and crystallographically17) suggested that the position of Nova binding to pre-mRNA predicted its action to enhance or inhibit alternate exon inclusion18. To identify direct Nova-RNA interactions *in vivo*, we applied high throughput sequencing methods to CLIP. Here we demonstrate that this approach uncovers new biology in the brain, identifying functional interactions that mediate tissue-specific alternative RNA processing.

Genome-wide protein-RNA interaction maps

We studied Nova-RNA interactions in the mouse neocortex, which expresses the Nova2 protein19. We identified 2,481 Nova-bound RNAs (CLIP tags)10,11 from five experiments using traditional CLIP strategies10,11 and 412,686 CLIP tags from three experiments using high throughput pyrosequencing. Tags were filtered to eliminate those with imperfect (<80%) matches to genomic sequences, with multiple genomic hits, or that were exact duplicates. The resulting set of 168,632 unique tags included 123,734 tags mapping to mRNA-encoding genes and 44,898 tags mapping to intergenic regions (Supplemental Fig. 2).

As negative controls, we repeated Nova CLIP using Nova KO brain or an irrelevant antibody, but were unable to amplify PCR products. We also sequenced 43,000 crosslinked RNA tags remaining after Nova immunoprecipitation, corresponding to a sample of all remaining RNABP-RNA interactions in the brain, and compared the frequency of Nova binding sites15 in these control tags with that in Nova HITS-CLIP tags. Only the latter showed enrichment for YCAY sequences (observed:expected YCAY frequency was 3.56 for Nova tags, compared to 0.99 for the control tags, determined by Chi-square distribution, $p < 10^{-227}$; see Methods), demonstrating the specificity of CLIP.

HITS-CLIP reproducibility and cluster analysis

We reasoned that one way to distinguish between biologically robust and transient Nova-RNA interactions would be to assess the reproducibility with which HITS CLIP tags were identified in individual mice. Tags obtained from the neocortex of two P13 littermates showed a remarkable degree of similarity; when equal numbers were aligned across the entire mouse genome in 10kb windows, a high correlation was evident both graphically (Fig. 1a) and statistically ($R^2 = 0.75$). To more accurately assess reproducibility, we focused on sites containing overlapping tags ("clusters"). 19,156 clusters had at least two tags, 508 had 20 or more tags (Supplemental Fig. 3a), and 608 RefSeq transcripts had ≥ 6 clusters

(Supplemental Fig. 2). Inter-animal clusters—sites containing at least one tag from each littermate—were highly reproducible; over 90% (9,697/10,740) of sites containing tags from one animal also had at least one littermate tag. Finally, inspection of individual chromosomes and genes revealed that Nova clusters were highly reproducible in both position and extent of crosslinking, and specific to a subset of brain-expressed RefSeq genes (Fig.1 and Supplemental Fig. 2).

To determine how faithfully Nova CLIP tag clusters reflect previously defined consensus Nova binding sites15, we analyzed them for consensus motifs by MEME analysis, and found they were significant enriched in YCAY motifs (Supplemental Fig. 3). This was evident across all 19,156 Nova CLIP tag clusters (3.9-fold; $p < 10^{-227}$), and in tags associated with functional Nova interactions (see below). Taken together, these observations indicate that HITS-CLIP reproducibly identifies discrete, YCAY-rich, Nova binding sites in mouse brain RNAs, and suggests that these binding sites may point to positions of functional Nova-RNA interactions.

HITS-CLIP analysis of splicing targets

HITS-CLIP offered a chance to compare predicted sites of Nova-RNA regulation derived from bioinformatic and microarray analysis11,18,20 with interaction sites observed by *in vivo* crosslinking. 39 previously validated20 Nova2-regulated transcripts harbored Nova CLIP tags (ranging from 1 to 96 tags) within 3 kb of the alternative exon local region (bounded by the constitutive splice donor and acceptor exons) and 34 of these harbored CLIP-tag clusters. The position and YCAY content (4.1-fold enrichment; $p < 10^{-156}$) of these clusters was consistent with the predicted Nova bioinformatic map18. For example, YCAY-rich HITS-CLIP clusters were present downstream of the known Nova2 target *Grin1* exon 19 (E19; Fig 1b-c(ii))20, in a position previously predicted by the Nova bioinformatic map18 (Supplemental Fig. 4).

We also observed HITS-CLIP tags in *Grin1* upstream of an alternative exon (exon 4; E4) that was not a previously known Nova target. The position of these tags predicted Nova-dependent inhibition of E4 inclusion, which was confirmed experimentally (Fig. 1b-c(i)), suggesting that HITS-CLIP might provide a general means to identify new sites of protein-RNA regulation. Six additional transcripts with Nova HITS-CLIP clusters near regulated splice sites were tested; each was aberrantly spliced in *Nova2* KO compared to WT brain in a manner conforming to the Nova bioinformatic map (Supplemental Fig. 5).

To further assess how the position of Nova binding related to the outcome of such splicing events, we analyzed Nova HITS-CLIP tags in Nova-regulated exons newly identified using an updated version of exon-junction microarrays20 harboring probesets for exon junctions in ~145,000 transcripts. Arrays were interrogated with RNA from WT or Nova2 null neocortex, and results analyzed with ASPIRE2, a revision of the ASPIRE algorithm20 that searches for reciprocal changes in exon-included and exon-excluded probesets. We identified 32/45 previously validated20 Nova2-dependent exons, and 46 new candidates with |ΔI| values ranging from 0.19 - 0.60 and with characteristics seen previously20 (Supplemental Fig. 6, Supplemental Tables 1-2). To simplify subsequent analysis, we

focused on 35 cassette exons, and confirmed that alternative splicing was Nova2-dependent in 7/7 (Supplemental Fig. 4).

We generated a map in which we placed all 1,085 Nova CLIP tags identified from a total of 71 Nova2-regulated cassette exons (43 validated targets, and 28 newly predicted targets with ΔI>0.2 and ΔI-tTest>25; see Methods) onto a single composite pre-mRNA (Fig. 2a; Supplemental Fig. 7). These tags spanned 11.5kb, but were very heavily concentrated around splice sites, in positions that corresponded extremely well with the bioinformatically predicted Nova map[18], and with prior biochemical analysis of Nova-dependent splicing[21] [22] [23] (Fig. 2a). Furthermore, clusters in these regions showed a 3.4-fold enrichment in YCAY elements ($p<10^{-174}$), with 72 of 123 clusters containing at least 3 YCAY elements within 30 nt, consistent with prior biochemical data[21] [22] [23].

We also noted some HITS-CLIP tags in unanticipated regions. For example, we observed frequent binding of Nova in intronic sequences upstream of Nova-regulated exons. However, binding to these sites was only robust in a limited number of transcripts (Fig. 2a; Supplemental Fig. 7). To generate a map representative of consensus Nova action, we normalized our data, first to the number and distribution of CLIP tags between transcripts, and then to the number of different transcripts with tags at a given position (complexity). This allowed us to focus on potential regulatory binding sites common to several transcripts. This "normalized complexity" map (Fig. 2b) demonstrated that Nova CLIP tags corresponded very precisely to the bioinformatically predicted sites of Nova action (Fig. 2b, insets). We conclude that HITS-CLIP confirms the hypothesis that Nova binding occurs directly on YCAY-rich elements near splice sites *in vivo*, and that the position of such Nova binding determines the outcome of Nova-dependent splicing regulation.

Nova regulates alternative polyadenylation

We next explored whether other HITS-CLIP clusters might reveal new Nova functions. Analysis of the genomic position of Nova clusters revealed that 23% of Nova HITS-CLIP tags mapped to intergenic regions (Fig. 3a). To examine the possibility that these tags may correspond to previously undescribed isoforms of RefSeq genes with alternative terminal exons, we examined the distance between intergenic clusters and neighboring RefSeq genes. There was an exponential increase in the cumulative number of tags within 10kb downstream of known stop codons, compared to linear increases beyond 10kb (717 versus 101 clusters within 10kb of the stop or start codon, respectively; Fig. 3b), or upstream of known start codons. This suggests that in addition to binding known 3′ UTR's (Fig 3a), Nova binds to unannotated 3′ UTR extensions of known genes. Within 3′ UTRs, tags were enriched near poly(A) sites, and to a lesser degree near stop codons (Fig. 3c). A large number of clusters were positioned within a few hundred nt of poly(A) sites (Fig. 3d), a region that contains core and potential auxiliary elements controlling transcript termination and poly(A) site utilization[24,25].

These observations suggested that Nova might function in a second pre-mRNA processing event in the mouse brain, regulated poly(A) site utilization (alternative polyadenylation), a process about which little is known. We analyzed alternative polyadenylation by hybridizing

Affymetrix Exon Arrays with *Nova2* WT versus KO brain RNA, and screened for changes in alternate 3′ UTR relative to total mRNA abundance (Supplemental Fig. 8). We identified 297 transcripts with such differences (1.5-fold; p<0.05); 43 contained 100 3′ UTR CLIP tag clusters, and these were preferentially present near poly(A) sites (Fig. 3d).

We tested poly(A) site use in two candidates, *Cugbp2* and *Slc8a1*. Both have microarray-predicted Nova-dependent changes in 3′ UTR usage (1.5 and 2 fold, respectively), and both contained CLIP tags near poly(A) sites (Fig. 4a; Supplemental Fig. 9). RNase protection analysis (RPA) demonstrated that utilization of these poly(A) sites was increased in *Nova2* KO brain (Fig. 4a, 4e; Supplemental Fig. 9); C (the change in percent transcripts cleaved at the relevant poly(A) site, analogous to I)18 for these transcripts was 0.22-0.25 (for example, 41% to 66% utilization of pA2 in *Cugbp2* transcripts in WT vs. *Nova2* KO brain; Fig. 4a), comparable in magnitude to Nova-dependent changes in alternative exon usage. Furthermore, the increase in proximal poly(A) use in *Cugbp2* and *Slc8a1* transcripts in *Nova2* KO brain was associated with reciprocal decreases in processing at distal poly(A) sites, suggesting that changes in the relative levels of alternatively polyadenylated *Cugbp2* and *Scl8a1* mRNAs are not due to differences in isoform stability, but result directly from aberrant poly(A) site utilization in the *Nova2* KO.

We used qRT-PCR to measure the relative abundance of alternative poly(A) isoforms of 29 additional candidate Nova targets (from Fig. 3d). 12 transcripts had significant changes in levels of alternatively polyadenylated transcripts (p<0.04 in 11/12; Fig. 4b). These transcripts did not change in overall transcript abundance (data not shown); 9 were consistent with a Nova-dependent action to block, and 3 to enhance utilization of the adjacent poly(A) site. 17 transcripts had either no change in poly(A) site usage in *Nova2* KO brain (most of these were low abundance (at least in alternate 3′ isoforms)), and/or had confounding changes in overall steady-state transcript levels, and 2 transcripts had 3′ UTR changes as well as alternate splicing of terminal exons. We mapped Nova CLIP tags from the 12 Nova-regulated 3′ UTRs onto a composite transcript containing an alternative polyadenylation site (Fig 4c), and onto a normalized complexity Nova-RNA 3′ UTR interaction map (Fig. 4d). This revealed reproducible Nova binding to discrete YCAY-rich (3.5-fold, $p<10^{-227}$) regions flanking Nova-regulated alternative poly(A) sites. Taken together, the quantitative analysis of transcript levels and the HITS-CLIP map demonstrate that Nova binds to YCAY-rich elements flanking poly(A) sites and is necessary for their proper regulation in mouse brain.

To test whether Nova binding to these 3′ UTR's is sufficient to suppress poly(A) site utilization, we generated a GFP reporter construct containing parts of the *Slc8a1* 3′ UTR harboring alternative poly(A) sites (Supplemental Fig. 9b), as well as mutant constructs in which YCAY elements were mutated to YACY (a sequence to which Nova does not bind15). Co-transfection of Nova2 expressing constructs with these reporters into 293T cells (which do not express Nova22) demonstrated a Nova and YCAY-dependent reduction in alternative poly(A) site usage of the same magnitude and direction as seen in WT versus *Nova2* KO neocortex (Fig. 4e). Taken together, these results identify direct Nova-RNA interactions in the 3′ UTR that regulate brain-specific alternative RNA processing.

Discussion

Genome-wide screens have been used to establish correlations between the action of RNABPs and biologic diversity6,7,26-30, but are unable to identify direct sites of RNA regulation. HITS-CLIP provides a general solution to this problem by generating a transcriptome-wide biochemical "footprint" of protein-RNA interactions in living tissues. This in turn allows a direct comparison of predicted (e.g. microarray or bioinformatically derived) and observed (HITS-CLIP) sites of action, and thereby provides a new platform for deriving functional RNABP maps and for assessing models of protein-RNA regulation.

HITS-CLIP extends our transcriptome-wide understanding of Nova-RNA interactions, which was previously limited to bioinformatic analysis of YCAY clusters within 200nt of alternate or bounding constitutive exons18. Analysis of HITS-CLIP tags mapping to 71 Nova-regulated exons (Fig. 2) yielded a more refined map of Nova action. Over 91% of the normalized Nova binding associated with exon inclusion (Fig. 2b) occurred within 500 nt of either the alternative 5′ or constitutive 3′ splice sites, while 74% of the normalized Nova binding associated with exon exclusion (Fig. 2b) occurred within 500 nt of the constitutive 5′ splice donor or surrounding the alternate exon. This strengthens the conclusion that the position of Nova-RNA interaction determines the outcome of splicing, an observation that may extend to splicing factors more generally31. Importantly, the strength of these correlations suggests that the HITS-CLIP map is sufficiently robust to predict protein-RNA regulation, as shown for seven new Nova splicing targets (Fig. 1d and Supplemental Fig. 5).

While the majority of the Nova regulated sites conform to a general set of rules based on direct Nova binding, there are also clear exceptions. For example, Nova binds robustly but in atypical positions in several regulated transcripts (e.g. *Brsk2*, *Rap1gap*) (Fig. 2a-b; Supplemental Fig. 7). Such examples may point to new mechanisms of Nova action, which may include interactions with other RNABPs. For example, PTBP2 interacts with Nova and other RNABPs such as KSRP to modulate the outcome of alternative splicing32,33. In addition, RNA structure may regulate or be impacted by interaction with other factors, as suggested by the ability of the splicing factor MBNL1 to stabilize RNA hairpins34, from analysis of splicing defects in *MAPT* that underlie frontotemporal dementia with parkinsonism35, and from structural studies of competition between hnRNP F and PTB binding to the *src* transcript (F. Allain, personal communication).

The unbiased nature of HITS-CLIP led to the unexpected identification of Nova binding near poly(A) sites and the recognition of its role in regulating alternative polyadenylation in the brain. The presence of such tissue-specific factors was postulated after the recognition of differential polyadenylation of IgM heavy chain transcripts in B cells36 and of calcitonin/ CGRP pre-mRNA in neurons37. Alternative poly(A) sites are present in ∼50% of human genes38, and their regulation is believed to play an important role in tissue and developmental mRNA regulation39,40 and in human disease41. In particular, brain mRNAs appear to be preferentially processed at promoter-distal poly(A) sites to generate long 3′ UTRs42,43. Interestingly, in 9 of 12 instances examined (Fig. 4b) Nova promoted the production of mRNAs with long 3′ UTRs. Thus one important action of Nova may be to

generate long 3′ UTRs in neurons, which may be subject to regulation by miRNAs or other RNABPs.

Numerous links have been made between pre-mRNA splicing and 3′-end processing[44,45], such as the observation that the splicing factor sex-lethal can regulate polyadenylation by competing with CstF64 for RNA binding[46]. While we found two transcripts that had both Nova-dependent changes in splicing and polyadenylation, Nova can mediate splicing-independent alternative polyadenylation. For example, Nova suppresses the *Slc8a1* pA2 site in an intronless transcript (Fig. 4), and *Cugbp2* and *Slc8a1* alternative polyadenylation was not coupled to alternative splicing in brain (unpublished data).

The Nova HITS-CLIP map offers insight into the mechanism of poly(A) site selection in the brain. Changes in the accessibility of core (e.g. CPSF and CstF) or auxiliary factors to interact with *cis* elements surrounding the poly(A) site underlie the regulation of alternative polyadenylation[24,25,47,48]. We find no evidence that Nova regulates transcripts encoding such factors (including subunits of CPSF, CSTF, CF-1 and CF-2). Instead, our data point to Nova as a *trans*-acting factor that binds YCAY elements flanking regulated poly(A) sites, and that the position of Nova binding may determine whether it acts to promote or inhibit poly(A) site use (Fig. 4). For example, Nova CLIP tags overlap the canonical CPSF and/or CstF binding sites within 30nt of the *Cugbp2* and *Slc8a1* poly(A) sites, which are suppressed by Nova. In contrast, in transcripts where Nova enhances poly(A) site use, it binds to more distal elements, where it may antagonize the action of auxiliary factors. Therefore the position of Nova 3′ UTR binding may determine the outcome of poly(A) site selection in a manner analogous to its action on splicing regulation (Supplemental Fig. 1).

In summary, HITS-CLIP offers a powerful new platform for studying RNA regulation *in vivo*. This genome-wide biochemical approach complements bioinformatic, microarray and genetic studies. HITS-CLIP is able to identify biologically relevant interactions, providing a focus on direct protein-RNA contacts as critical points for understanding RNABP function. The unbiased nature of the platform holds the potential for new discovery, including the elucidation of preferred binding sequences and the identification of regulated RNA substrates. Identifying Nova as the first vertebrate factor to regulate alternative polyadenylation in mouse brain demonstrates that a single factor can regulate different aspects of tissue-specific RNA metabolism. Finally, the reproducible nature of HITS-CLIP suggests that it provides a robust platform to explore RNABP-dependent mechanisms of gene expression in complex and dynamic scenarios.

Methods Summary

HITS-CLIP

CLIP was performed on mouse *Nova2* WT and KO (CD1) brains as described[11]. After PCR amplification, high throughput sequencing was performed (454 Life Sciences).

Microarrays

For analysis of Nova-dependent alternate splice and alternate 3′ UTR variants, a custom exon junction array (Affymetrix) or MoEx 1.0 ST Affymetrix Exon arrays, respectively, were used.

Bioinformatics

CLIP tags and clusters were analyzed with BED or WIG formatted custom tracks using the UCSC Genome Browser and Genome Graph tools (genome.ucsc.edu). Composite maps were generated by determining the distance between tags and closest splice sites within the alternative exon local region and converted to coordinates in a BED format custom track, with tags from each gene assigned different colors. MEME sequence analysis was done using tools available at meme.sdsc.edu. ASPIRE2 was based on ASPIRE20.

Biochemical and transfection assays

Biochemical assays were done using biologic triplicate sibling mice, unless otherwise noted. RPAIII kits from Ambion were used, and RT-PCR were done as described21 22 23, with modifications described in Methods; qPCR was done with a MyIQ BioRad thermal cycler and data analyzed as described in Methods. WT or mutant GFP alternative polyadenylation reporters were transfected into 293T cells in the presence or absence of pNova2 (described in Methods).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Zaug AJ, Cech TR. The intervening sequence RNA of Tetrahymena is an enzyme. Science. 1986; 231:470–475. [PubMed: 3941911]

2. de Duve C. Co-chairman's remarks: the RNA world: before and after. Gene. 1993; 135:29–31. [PubMed: 7506221]

3. Maizels N, Weiner AM. The 'last ribo-organism' was no breakthrough. Nature. 1987; 330:616. [PubMed: 2446146]

4. Gilbert W. The RNA world. Nature. 1986; 319:618.

5. Sharp PA. On the origin of RNA splicing and introns. Cell. 1985; 42:397–400. [PubMed: 2411416]

6. David CJ, Manley JL. The search for alternative splicing regulators: new approaches offer a path to a splicing code. Genes Dev. 2008; 22:279–285. [PubMed: 18245441]

7. Moore MJ, Silver PA. Global analysis of mRNA splicing. RNA. 2008; 14:197–203. [PubMed: 18083834]

8. Mili S, Steitz JA. Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. RNA. 2004; 10:1692–1694. [PubMed: 15388877]

9. Darnell JC, Mostovetsky O, Darnell RB. FMRP RNA targets: identification and validation. Genes Brain Behav. 2005; 4:341–349. [PubMed: 16098133]

10. Ule J, Jensen K, Mele A, Darnell RB. CLIP: a method for identifying protein-RNA interaction sites in living cells. Methods. 2005; 37:376–386. [PubMed: 16314267]

11. Ule J, et al. CLIP identifies Nova-regulated RNA networks in the brain. Science. 2003; 302:1212–1215. [PubMed: 14615540]

12. Guil S, Caceres JF. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. Nat Struct Mol Biol. 2007; 14:591–596. [PubMed: 17558416]

13. van der Brug MP, et al. RNA binding activity of the recessive parkinsonism protein DJ-1 supports involvement in multiple cellular pathways. Proc Natl Acad Sci U S A. 2008; 105:10244–10249. [PubMed: 18626009]

14. Buckanovich RJ, Posner JB, Darnell RB. Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. Neuron. 1993; 11:657–672. [PubMed: 8398153]

15. Darnell RB. Developing global insight into RNA regulation. Cold Spring Harb Symp Quant Biol. 2006; 71:321–327. [PubMed: 17381312]

16. Buckanovich RJ, Yang YY, Darnell RB. The onconeural antigen Nova-1 is a neuron-specific RNA-binding protein, the activity of which is inhibited by paraneoplastic antibodies. J Neurosci. 1996; 16:1114–1122. [PubMed: 8558240]

17. Lewis HA, et al. Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. Cell. 2000; 100:323–332. [PubMed: 10676814]

18. Ule J, et al. An RNA map predicting Nova-dependent splicing regulation. Nature. 2006; 444:580–586. [PubMed: 17065982]

19. Yang YYL, Yin GL, Darnell RB. The neuronal RNA binding protein Nova-2 is implicated as the autoantigen targeted in POMA patients with dementia. Proc Natl Acad Sci. 1998; 95:13254–13259. [PubMed: 9789075]

20. Ule J, et al. Nova regulates brain-specific splicing to shape the synapse. Nat Genet. 2005; 37:844–852. [PubMed: 16041372]

21. Jensen KB, et al. Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. Neuron. 2000; 25:359–371. [PubMed: 10719891]

22. Dredge BK, Darnell RB. Nova regulates GABA(A) receptor gamma2 alternative splicing via a distal downstream UCAU-rich intronic splicing enhancer. Mol Cell Biol. 2003; 23:4687–4700. [PubMed: 12808107]

23. Dredge BK, Stefani G, Engelhard CC, Darnell RB. Nova autoregulation reveals dual functions in neuronal splicing. EMBO J. 2005; 24:1608–1620. [PubMed: 15933722]

24. Hu J, Lutz CS, Wilusz J, Tian B. Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. RNA. 2005; 11:1485–1493. [PubMed: 16131587]

25. Zhao J, Hyman L, Moore C. Formation of mRNA 3′ ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol Mol Biol Rev. 1999; 63:405–445. [PubMed: 10357856]

26. Ben-Dov C, Hartmann B, Lundgren J, Valcarcel J. Genome-wide analysis of alternative pre-mRNA splicing. J Biol Chem. 2008; 283:1229–1233. [PubMed: 18024428]

27. Blencowe BJ. Alternative splicing: new insights from global analyses. Cell. 2006; 126:37–47. [PubMed: 16839875]

28. Wang GS, Cooper TA. Splicing in disease: disruption of the splicing code and the decoding machinery. Nat Rev Genet. 2007; 8:749–761. [PubMed: 17726481]

29. Johnson JM, et al. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. Science. 2003; 302:2141–2144. [PubMed: 14684825]

30. Keene JD. RNA regulons: coordination of post-transcriptional events. Nat Rev Genet. 2007; 8:533–543. [PubMed: 17572691]

31. Zhang C, et al. Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2. Genes Dev. 2008; 22:2550–2563. [PubMed: 18794351]

32. Markovtsov V, et al. nPTB, a tissue-specific homolog of polypyrimidine-tract binding protein. Mol Cell Biol. 2000 submitted.

33. Polydorides AD, Okano HJ, Yang YY, Stefani G, Darnell RB. A brain-enriched polypyrimidine tract-binding protein antagonizes the ability of Nova to regulate neuron-specific alternative splicing. Proc Natl Acad Sci U S A. 2000; 97:6350–635. [PubMed: 10829067]

34. Yuan Y, et al. Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. Nucleic Acids Res. 2007; 35:5474–5486. [PubMed: 17702765]

35. Grover A, et al. 5′ splice site mutations in tau associated with the inherited dementia FTDP-17 affect a stem-loop structure that regulates alternative splicing of exon 10. J Biol Chem. 1999; 274:15134–15143. [PubMed: 10329720]

36. Early P, et al. Two mRNAs can be produced from a single immunoglobulin mu gene by alternative RNA processing pathways. Cell. 1980; 20:313–319. [PubMed: 6771020]

37. Rosenfeld MG, et al. Production of a novel neuropeptide encoded by the calcitonin gene via tissue-specific RNA processing. Nature. 1983; 304:129–135. [PubMed: 6346105]

38. Iseli C, et al. Long-range heterogeneity at the 3′ ends of human mRNAs. Genome Res. 2002; 12:1068–1074. [PubMed: 12097343]

39. Edwalds-Gilbert G, Veraldi KL, Milcarek C. Alternative poly(A) site selection in complex transcription units: means to an end? Nucleic Acids Res. 1997; 25:2547–2561. [PubMed: 9185563]

40. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. Proliferating cells express mRNAs with shortened 3′ untranslated regions and fewer microRNA target sites. Science. 2008; 320:1643–1647. [PubMed: 18566288]

41. Danckwardt S, Hentze MW, Kulozik AE. 3′ end mRNA processing: molecular mechanisms and implications for health and disease. Embo J. 2008; 27:482–498. [PubMed: 18256699]

42. Burge CB. Nature. 2008 al., e. in press.

43. Zhang H, Lee JY, Tian B. Biased alternative polyadenylation in human tissues. Genome Biol. 2005; 6:R100. [PubMed: 16356263]

44. Maniatis T, Reed R. An extensive network of coupling among gene expression machines. Nature. 2002; 416:499–506. [PubMed: 11932736]

45. Proudfoot NJ, Furger A, Dye MJ. Integrating mRNA processing with transcription. Cell. 2002; 108:501–512. [PubMed: 11909521]

46. Gawande B, Robida MD, Rahn A, Singh R. Drosophila Sex-lethal protein mediates polyadenylation switching in the female germline. Embo J. 2006; 25:1263–1272. [PubMed: 16511567]

47. Takagaki Y, Seipelt RL, Peterson ML, Manley JL. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell. 1996; 87:941–952. [PubMed: 8945520]

48. Veraldi KL, et al. hnRNP F influences binding of a 64-kilodalton subunit of cleavage stimulation factor to mRNA precursors in mouse B cells. Mol Cell Biol. 2001; 21:1228–1238. [PubMed: 11158309]

49. Clark TA, et al. Discovery of tissue-specific exons using comprehensive human exon microarrays. Genome Biol. 2007; 8:R64. [PubMed: 17456239]
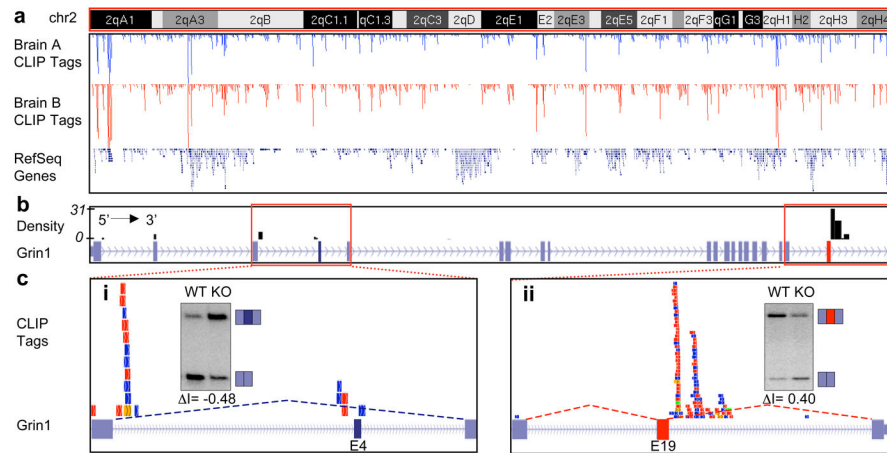
Figure 1. HITS-CLIP genome-wide map of Nova-RNA binding sites
a, Chromosome 2 RefSeq genes and CLIP tags from the neocortex of two mouse littermates (Brain A, 46,106 tags in 10,740 clusters; Brain B, 100,874 tags in 15,805 clusters). **b**, Cluster Density (#tags/cluster length; black) in the *Grin1* transcript; **c**, (i) *Grin1* E4 and (ii) E19 tags, one color per biologic replicate, predict Nova-dependent exon skipping and inclusion, respectively (experimental validation20 is shown).
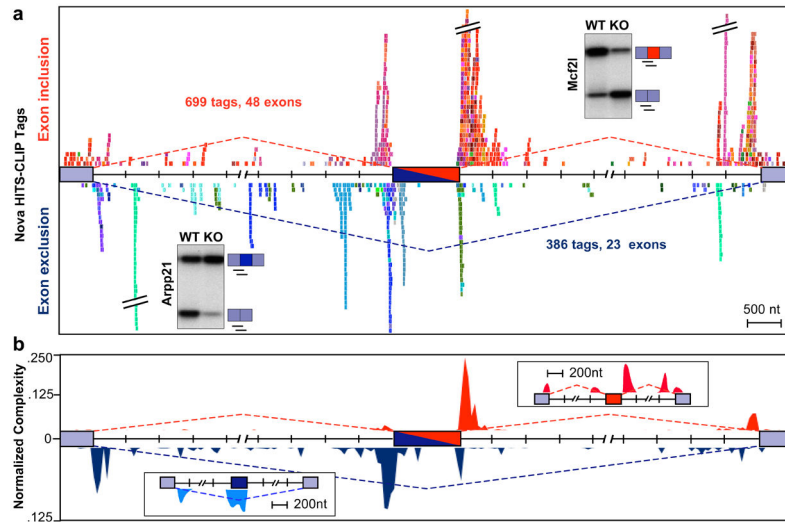
Figure 2. Nova-RNA interaction maps associated with the Nova-dependent splicing regulation
a, CLIP tags around all known Nova-regulated cassette exons; one color per transcript. Tags were mapped onto a composite transcript containing an alternative (dark blue/red box) and flanking constitutive (light blue box) exons. Tags are from transcripts showing Nova-dependent exon inclusion (top panel) or exclusion (bottom panel); experimental validation is shown (insets). **b,** Normalized complexity map (see methods) of Nova-RNA interactions recapitulate predicted maps18 (insets) for Nova-dependent exon inclusion (red) or exclusion (blue).
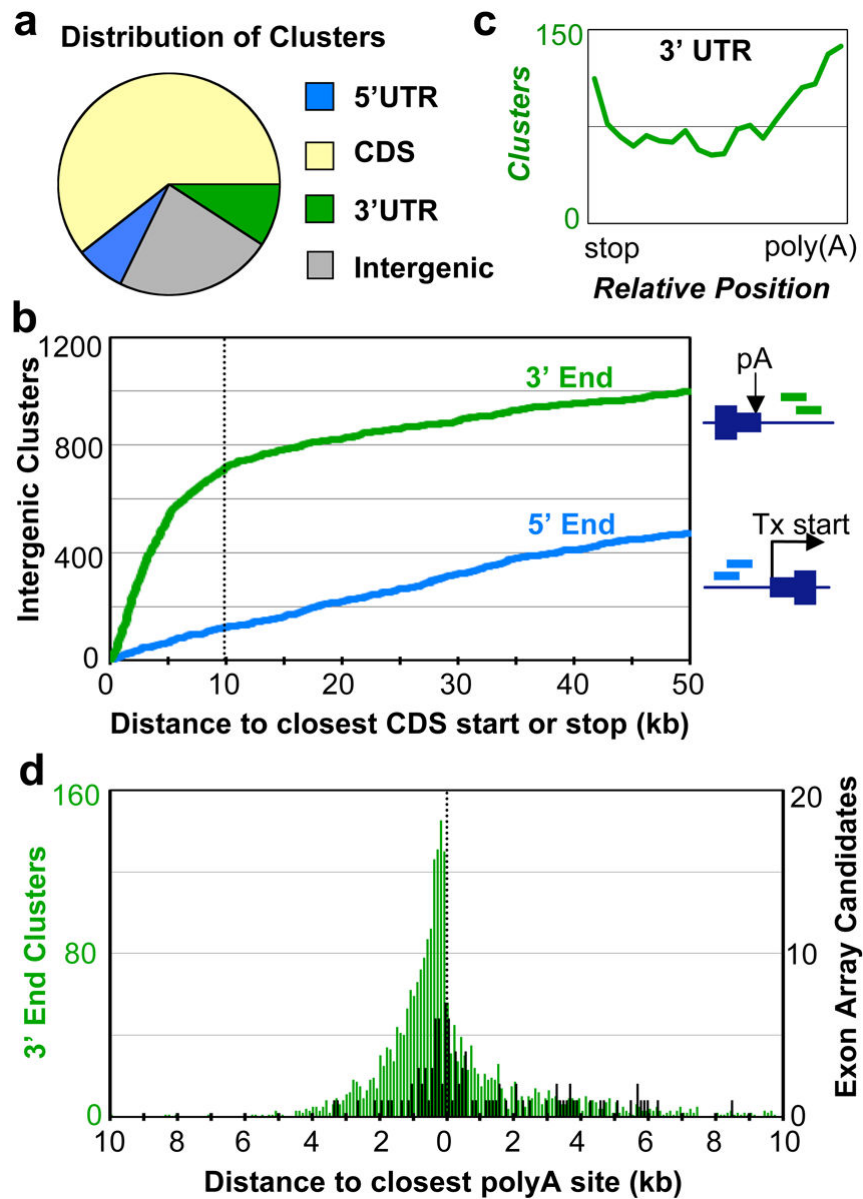
Figure 3. Nova CLIP tags cluster near polyadenylation sites
a, Distribution of Nova CLIP tag clusters (RefSeq mm8; CDS refers to introns/exons between UTRs). **b,** Intergenic clusters were plotted relative to the closest annotated start (5′ end) or stop (3′ end) codon. **c,** Distribution of 2,465 clusters in 3′ UTRs (defined as distance from stop codon to closest transcript termination site). **d,** CLIP tag clusters relative to the closest poly(A) site (transcript end in UCSC known genes), plotted as clusters per 50 nt in 2,465 3′ UTRs (green bars) or in 43 Nova-dependent 3′ UTRs identified with exon arrays (black bars).
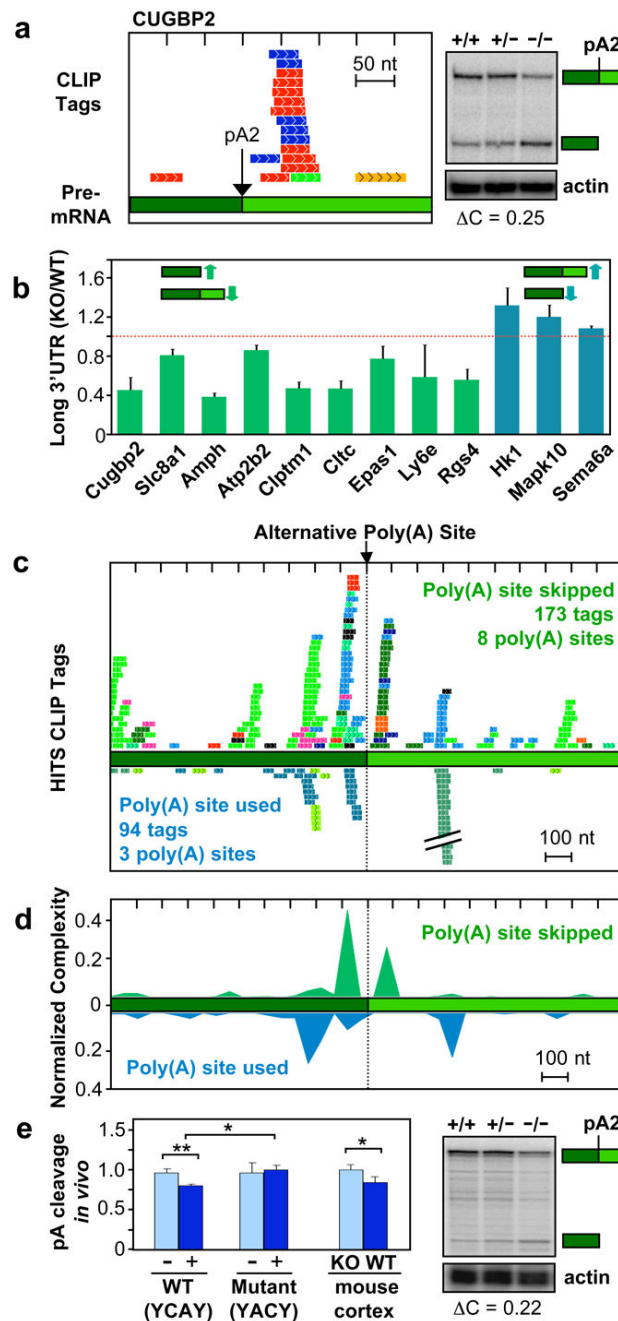
Figure 4. Nova regulates alternative polyadenylation

a, Nova CLIP tags near *Cugbp2* poly(A) site 2 (pA2); each color represents a biologic replicate. Right panel: RPA measuring cleavage at pA2 in WT (+/+), *Nova2* heterozygous (+/−) and KO (−/−) P10 neocortex, with actin as a normalization control and ΔC (see text) shown. **b,** qRT-PCR validation of Nova-dependent poly(A) regulation, using RNA from three WT or *Nova2* KO littermates, presented as long 3′ UTR isoform relative to total transcript abundance, normalized to WT (1.0; dotted line); error bars represent standard deviation. **c-d,** Composite (c) and normalized complexity (d) maps for Nova-regulated

alternative polyadenylation sites (as in Fig. 2). **e,** Nova binds YCAY elements to directly regulate alternative polyadenylation. Left panel: qRT-PCR analysis of cells transfected with WT or Mutant (3 YACY mutations) poly(A) reporter constructs cotransfected with a control (−) or Nova2 expressing (+) plasmid, or WT vs. *Nova2* KO neocortex (* = $p < 0.02$, ** = $p < 0.01$); error bars represent standard deviation. Right panel: RPA for cleavage of *Slc8a1* poly(A) site 2, otherwise as described in (a).