

- Miller, P.S., Da Silva, H.M., and Smart, T.G. (2005). *J. Biol. Chem.* 280, 37877–37884.
- Palmiter, R.D., and Huang, L. (2004). *Pflugers Arch.* 447, 744–751.
- Paoletti, P., Ascher, P., and Neyton, J. (1997). *J. Neurosci.* 17, 5711–5725.
- Smart, T.G., Hosie, A.M., and Miller, P.S. (2004). *Neuroscientist* 10, 432–442.
- Suwa, H., Saint-Amant, L., Triller, A., Drapeau, P., and Legendre, P. (2001). *J. Neurophysiol.* 85, 912–925.
- Vogt, K., Mellor, J., Tong, G., and Nicoll, R. (2000). *Neuron* 26, 187–196.

DOI 10.1016/j.neuron.2006.11.007

## Maps, Codes, and Sequence Elements: Can We Predict the Protein Output from an Alternatively Spliced Locus?

**Alternative splicing choices are governed by splicing regulatory protein interactions with splicing silencer and enhancer elements present in the pre-mRNA. However, the prediction of these choices from genomic sequence is difficult, in part because the regulators can act as either enhancers or silencers. A recent study describes how for a particular neuronal splicing regulatory protein, Nova, the location of its binding sites is highly predictive of the protein's effect on an exon's splicing.**

In eukaryotic cells, the formation of a mature mRNA requires the removal of introns from the precursor mRNA and splicing of its exons. This is a key step in determining the protein output from a gene. Alternative splicing allows joining of exons in different patterns, enabling a single gene to produce multiple protein isoforms (Black, 2003). This form of regulation is particularly common in the mammalian nervous system where a large percentage (40%–60%) of neuronal pre-mRNAs undergo alternative splicing (Yeo et al., 2004). Many of these splicing events produce proteins important for neuronal development (e.g. neurexins, EphA7) and mature neuronal function (e.g. NMDA receptor 1, CaV2), and changes in the splicing of their pre-mRNAs result in multiple functional variants (Lipscombe, 2005). A long-term goal of genomic research is the prediction of the protein products of a gene under different cellular conditions. However, the frequency with which gene transcripts show alternative splicing in the nervous system and the large number of potential products from some of these genes make this prediction extremely difficult.

The information that determines an alternative splicing pattern is usually encoded within the sequence of a regulated exon and its flanking introns in the form of intronic or exonic splicing silencer elements (ISS or ESS) and intronic or exonic splicing enhancer elements (ISE or ESE). Enhancer elements promote the inclusion of an exon, and silencers promote its skipping or exclusion from the final mRNA. Many of these elements are bound by known RNA-binding proteins (RBPs), such

as the SR proteins and members of the hnRNP group of proteins (Black, 2003). Many more elements have been identified, but their protein mediators are unknown. Most alternative exons are controlled by multiple splicing enhancer and silencer elements. Moreover, many elements are not strict silencers or enhancers, rather the position of an element relative to an alternative exon can determine whether it acts positively or negatively (Hui et al., 2005). We need to know much more about the rules that determine the positive or negative activity of these regulatory elements. What is the code for exon use? What are the mechanisms by which these elements act on the splicing apparatus? A recent paper in *Nature* shows for one splicing regulator that such a code exists and can successfully predict the regulatory properties of the exon targets for this protein (Ule et al., 2006). Moreover, this paper and recent work from others have started to identify the specific steps in spliceosome assembly affected by a regulatory protein.

Genomic analyses of alternative splicing indicate that the brain has the highest frequency of alternative splicing (Sugnet et al., 2006; Yeo et al., 2004). The neuron-specific inclusion or skipping of alternative exons is primarily achieved by tissue-specific expression of particular RBPs (Matlin et al., 2005). Over a dozen neuron-specific RBPs have been identified, including the Hu family, members of the CELF family, neural PTB, the Fox protein family, and Nova—the subject of the *Nature* paper (McKee et al., 2005; Ule et al., 2006). These regulators affect a wide range of target gene transcripts, encoding proteins involved in cytoskeletal rearrangement, vesicular transport, cell adhesion, signal transduction pathways, and synaptic activity (Lipscombe, 2005). In earlier work, the Darnell lab identified Nova protein as a YCAY element binding factor. They went on to use a variety of biochemical, genomic, and genetic approaches to identify a large number of exons in genes affecting synaptic function whose splicing was affected by Nova (Ule et al., 2006, and references therein).

The known Nova-regulated exons showed a variety of dependencies, some enhanced by the protein and others repressed. In the present study, the authors carried out statistical analyses of the positions of YCAY clusters relative to the known Nova target exons (Ule et al., 2006). They noted the positions of these clusters relative to the target exon and created a general map relating the frequency of YCAY clusters surrounding Nova-regulated exons. These clusters fall into relatively few defined positions upstream, downstream, or within the exon (Figure 1). Interestingly, certain positions strongly correlate with splicing enhancement by Nova, while other positions correlate with Nova acting to repress splicing. The authors tested the predictive value of the map by searching a genomic database for additional YCAY cluster-containing exons that had not been previously identified as Nova targets. The splicing of these exons was compared in wild-type and Nova knockout mice. They found that the positions of Nova-dependent enhancers and silencers indicated by the map were strongly predictive of the changes observed between the presence and absence of Nova. This both validated the value of the Nova RNA map and identified a new set of Nova target exons.

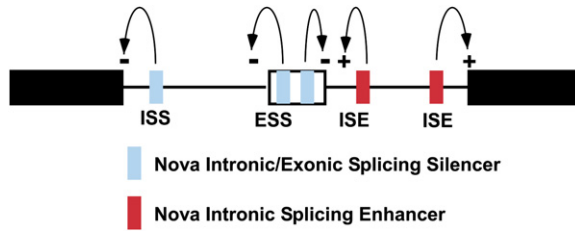


Figure 1. The Nova Binding Site Map Shows that the Position of YCAY Clusters Can Be Correlated with Their Effect on Alternative Splicing

The position of splicing regulatory elements has often been shown to affect their activity as enhancers or silencers. This has been best studied for the SR proteins, which enhance splicing when bound in an exon but are inhibitory as intron-binding proteins (Ibrahim et al., 2005). HnRNPs H and L seem to repress splicing when bound in an exon, but activate it from the intron downstream (Black, 2003; House and Lynch, 2006; Hui et al., 2005). Fox proteins bound upstream of an exon repress its splicing, whereas downstream Fox binding enhances it (Jin et al., 2003). The Nova results also fit with genome-wide analyses that most splicing regulatory elements lie in the vicinity of 5' and 3' splice sites, or upstream of the branch point (Sugnet et al., 2006, and references therein). Nova enhancer elements are intronic, most commonly in the intron downstream of the regulated exon, but some are upstream. Nova silencer elements, on the other hand, are either within the regulated exon or close to the 5' splice site of the upstream intron (Figure 1). What is new in the Ule paper is the precise definition of the regions where YCAY clusters can act and their clear predictive power for identifying Nova-regulated exons and their positive or negative response. It is likely that similar predictive maps can be generated for other splicing elements and their regulatory factors.

All major class introns are excised by a dynamic macromolecular complex called the spliceosome (Will and Luhrmann, 2006). The spliceosome assembles through the sequential binding of five small nuclear ribonucleoprotein particles (snRNPs) and multiple auxiliary proteins. The ordered association of the snRNPs and proteins with the target intron leads to the formation of discrete intermediate complexes (Figure 2A). During this assembly, the two splice sites are recognized by specific components, brought into close juxtaposition, and ultimately subjected to the two catalytic reactions that result in intron excision and exon ligation (Figures 2A and 2B). The stepwise nature of spliceosome assembly provides multiple potential points at which it can be regulated. Indeed, different regulatory proteins have been shown to affect different steps in spliceosome assembly (Black, 2003, and references therein). Splicing repressors can act by directly blocking the binding of spliceosomal components with the pre-mRNA, by blocking the interaction of splicing enhancers with the basal splicing machinery, or by blocking the transition from one stage of spliceosome assembly to the next (House and Lynch, 2006; Izquierdo et al., 2005; Sharma et al., 2005). In each case, the interactions of the repressor protein with basal splicing components can lead to

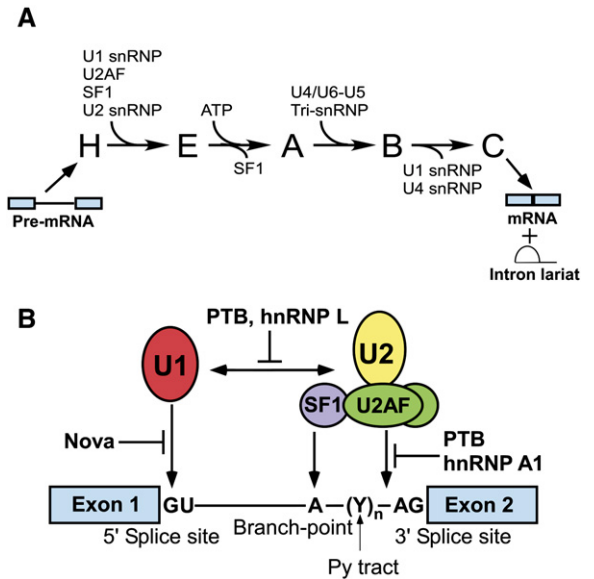


Figure 2. Pre-mRNA Splicing Is Catalyzed by the Spliceosome (A) The spliceosome assembles from five snRNPs, referred to as U1, U2, U4, U5, and U6, and auxiliary proteins. The multiple steps in this assembly lead to the formation of distinct H, E, A, B, and C complexes. (B) The E complex has the U1 snRNP bound to the 5' splice site and the SF1 and U2AF proteins bound to the BP and polypyrimidine tract, respectively. Splicing regulatory proteins interfere with the binding of core components of the splicing machinery or prevent transitions between spliceosome assembly intermediates.

exon skipping. To examine the mechanism by which Nova repressed exon inclusion, Ule and colleagues examined spliceosome assembly for an RNA containing an inhibitory YCAY cluster within the regulated exon. Splicing complex analysis on this RNA indicated that Nova blocked assembly of the E complex and more specifically the association of the U1 snRNP with the 5' splice site (see Figure 2A for the specifics of spliceosome assembly). This is different from splicing repressors, such as PTB and hnRNP L, which generally do not block the initial binding of snRNPs or other basal components, but instead interfere with their interactions within subsequent prespliceosome complexes. In light of the diverse nature of the Nova silencer and enhancer elements, it is likely that they will use additional mechanisms to alter spliceosome assembly beyond the one identified by Ule et al.

To examine Nova function in vivo, the authors also examine the partially spliced RNA surrounding Nova-dependent exons in wild-type and Nova knockout mice. Using RT/PCR to amplify RNAs where either the upstream or the downstream intron has been excised, but not the other, they find the build up or decrease of only one of the partially spliced RNAs in the Nova knockout mice. For intronic YCAY clusters, Nova affects the extent of splicing of the YCAY-containing intron, and not the opposite intron across the regulated exon. If the YCAY cluster is exonic, Nova affects the removal of the more proximal intron. This agrees with the observed effect on U1 binding for the in vitro experiments but also indicates that Nova likely affects other factors when the binding site is not adjacent to the 5' splice

site. Moreover, most tissue-specific exons are regulated by multiple positive and negative factors. When Nova is absent, these exons may be more strongly affected by other factors than is normally seen. Thus, it is not clear whether there will be a general rule for the order of intron removal.

A long-term goal of genomics is to predict the tissue-specific expression of each protein encoded within a genome. The prevalence of alternative splicing in the mammalian nervous system and the diverse factors controlling it make this much more difficult. However, the success of the Ule study in predicting Nova regulation from the global mapping of its binding sites is a very promising step forward and can be usefully applied to other splicing regulators. With these maps in hand, we can begin the even more daunting task of examining how the splicing apparatus can integrate the multiple positive and negative inputs for each exon.

**Shalini Sharma<sup>1</sup> and Douglas L. Black<sup>1</sup>**

<sup>1</sup>Department of Microbiology, Immunology,  
and Molecular Genetics  
Howard Hughes Medical Institute  
University of California, Los Angeles  
Los Angeles, California 90095

#### Selected Reading

- Black, D.L. (2003). *Annu. Rev. Biochem.* 72, 291–336.
- House, A.E., and Lynch, K.W. (2006). *Nat. Struct. Mol. Biol.* 13, 937–944.
- Hui, J., Hung, L.H., Heiner, M., Schreiner, S., Neumuller, N., Reither, G., Haas, S.A., and Bindereif, A. (2005). *EMBO J.* 24, 1988–1998.
- Ibrahim el, C., Schaal, T.D., Hertel, K.J., Reed, R., and Maniatis, T. (2005). *Proc. Natl. Acad. Sci. USA* 102, 5002–5007.
- Izquierdo, J.M., Majos, N., Bonnal, S., Martinez, C., Castelo, R., Guigo, R., Bilbao, D., and Valcarcel, J. (2005). *Mol. Cell* 19, 475–484.
- Jin, Y., Suzuki, H., Maegawa, S., Endo, H., Sugano, S., Hashimoto, K., Yasuda, K., and Inoue, K. (2003). *EMBO J.* 22, 905–912.
- Lipscombe, D. (2005). *Curr. Opin. Neurobiol.* 15, 358–363.
- Matlin, A.J., Clark, F., and Smith, C.W. (2005). *Nature* 6, 386–398.
- McKee, A.E., Minet, E., Stern, C., Riahi, S., Stiles, C.D., and Silver, P.A. (2005). *BMC Dev. Biol.* 5, 14.
- Sharma, S., Falick, A.M., and Black, D.L. (2005). *Mol. Cell* 19, 485–496.
- Sugnet, C.W., Srinivasan, K., Clark, T.A., O'Brien, G., Cline, M.S., Wang, H., Williams, A., Kulp, D., Blume, J.E., Haussler, D., and Ares, M., Jr. (2006). *PLoS Comput. Biol.* 2, e4. Published online January 20, 2006. 10.1371/journal.pcbi.0020004.
- Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J., and Darnell, R.B. (2006). *Nature*, in press. Published online October 25, 2006. 10.1038/nature05304.
- Will, C.L., and Luhrmann, R. (2006). In *RNA World*, Third Edition, R.F. Gesteland, T.R. Cech, and J.F. Atkins, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 369–400.
- Yeo, G., Holste, D., Kreiman, G., and Burge, C.B. (2004). *Genome Biol.* 5, R74.