RNA Processing: Redrawing the Map of Charted Territory

Anna Corrionero[1,2] and Juan Valcárcel[1,2,3,*]
[1]Centre de Regulació Genòmica
[2]Universitat Pompeu Fabra
[3]Institució Catalana de Recerca i Estudis Avançats
Dr. Aiguader 88, 08010 Barcelona, Spain
*Correspondence: juan.valcarcel@crg.es
DOI 10.1016/j.molcel.2009.12.004

Using genome-wide RNA-binding data, Xue et al. (2009) draw a regulatory map in this issue of *Molecular Cell* for the much-studied polypyrimidine tract-binding protein (PTB) that reveals a unique paradigm in posttranscriptional gene regulation.

When great explorers like Chen Ho, Columbus, or Shackleton set sail to uncharted land, they were prepared to endure hardship and danger but were also confident that, if successful, their expeditions would bring back goods, stories, and detailed maps of unheard-of wonders. In contrast, few explorers would be thrilled to chart well-mapped terrain. In this issue of *Molecular Cell*, Xue et al. (2009) reinvestigate one of the most studied posttranscriptional regulators of gene expression, the polypyrimidine tract-binding protein (PTB), by analyzing its binding sites in the transcriptome of human HeLa cells. Their results show that novel principles remain to be discovered even in well-charted territory.

PTB (also known as hnRNP I) is an RNA-binding protein expressed in multiple tissues, with binding preference for CU-rich sequences (Singh et al., 2005; Oberstrass et al., 2005). Like other members of the hnRNP family, PTB plays a variety of roles in RNA metabolism, including regulation of alternative pre-mRNA splicing, polyadenylation, mRNA stability, and translation. For example, PTB has emerged as a repressor of neuron-specific alternative splicing in nonneural cells. Indeed, downregulation of PTB and upregulation of a functionally distinct neuron-specific paralog (nPTB) may be involved in about 25% of the splicing changes observed during neuron differentiation (Boutz et al., 2007). A similar switch has been reported between PTB and its hematopoietic cells paralog ROD1 (Spellman et al., 2007). Given the variety of functions and targets associated with PTB, genome-wide identification of its binding sites is of obvious general interest.

For this, Xue et al. (2009) employed an in vivo UV crosslinking and immunoprecipitation (CLIP) technique developed by Darnell and coworkers (Ule et al., 2003). This method captures close interactions between proteins and RNA molecules in living cells. After UV irradiation, cells are lysed and the crosslinked RNA-protein complexes immunoprecipitated. Treatment with RNase digests away the RNA molecules except for those fragments protected by their interaction with the protein. Sequencing of these RNA tags not only identifies target transcripts, but also informs about the location of protein binding within the transcript. When combined with high-throughput sequencing technologies, the method allows one to exhaustively characterize binding sites in whole transcriptomes, thus providing a panoramic view of targets and possible functions of the RNA-binding protein (Licatalosi et al., 2008; Yeo et al., 2009).

The results of Xue et al. (2009) significantly expand the catalog of PTB-binding sites, confirm known gene targets of PTB-mediated regulation, refine their possible mechanisms of regulation, identify novel regulated alternative splicing events, and open intriguing new possibilities for PTB function. PTB-binding sites are found in nearly 50% of human protein-coding transcripts and in 20% of all annotated alternatively spliced regions, confirming the potential genome-wide effects of the protein. Consistent with the variety of functions that PTB plays in gene regulation, only 30% of PTB sites are associated with alternative splicing events, and additional roles in constitutive splicing and in repression of cryptic sites can be envisioned. Nearly 30% of PTB sites map to intergenic regions, perhaps indicative of a function in the metabolism of noncoding transcripts. Of relevance, most, but not all, binding sites could be predicted bioinformatically by previous knowledge of PTB RNA binding preferences, and only a small fraction of the predicted sites were detected experimentally. Sequence context effects associated with cooperative or antagonistic interactions with other factors are likely to explain these differences, arguing for the added value of experimental data.

When the binding landscape provided by high-throughput methods like CLIP is combined with functional information, for example obtained using splicing-sensitive microarrays, correlations between the location of binding sites and regulatory outputs emerge. Such representations are known as RNA maps (Figure 1). RNA maps have been reported for the tissue-specific splicing regulators Nova and Fox2 (Ule et al., 2006; Licatalosi et al., 2008; Yeo et al., 2009). These maps share interesting features: binding of Nova/Fox2 upstream of the regulated exon correlates with Nova/Fox2-mediated exon skipping, whereas binding to the intronic region downstream of the exon correlates with exon inclusion (Figures 1A and 1B). Although the mechanisms for these position-dependent effects remain to be determined, these RNA maps suggest that Nova and Fox2

Cell PRESS

use common molecular principles to modulate splice site choice.

In contrast, the RNA map drawn for PTB by Xue et al. (2009) is remarkably different. PTB-induced exon inclusion or skipping correlates with the location of PTB-binding sites relative to the constitutive or alternative splice sites involved in the regulated event (Figure 1C). As expected from previous knowledge of PTB-mediated regulation, PTB binding near an alternative exon correlates with skipping. Unexpectedly, PTB binding near the constitutive distal splice sites correlates with enhanced exon inclusion. Although not without exceptions, these results greatly populate the catalog of genes in which PTB promotes exon inclusion and reveal an internal logic for the dual activity of this protein in alternative splicing regulation.

A variety of mechanisms have been proposed for PTB-mediated effects on splice site selection, including steric hindrance of splice sites, interference with intron or exon definition, looping out RNA regions, or cooperative spreading of PTB from high-affinity sites (Spellman and Smith, 2006). None of these models can easily explain PTB-mediated exon inclusion. As argued by Xue et al., PTB binding may make neighboring sites less competitive by interfering with thermodynamic or kinetic parameters of their recognition by basal splicing factors. PTB could therefore serve as a valve that diverts the flow of splicing activity from a region of the pre-mRNA and thus modulates splice site competition.

The results of Xue et al. eloquently argue that high-throughput-enabling technologies provide panoramic views, illuminating mechanistic concepts that escaped the attention of researchers doing detailed work on more limited sets of genes. They also bring hope that RNA maps will one day allow us to navigate through genome sequences to predict their posttranscriptional outputs in different tissues or physiological situations.

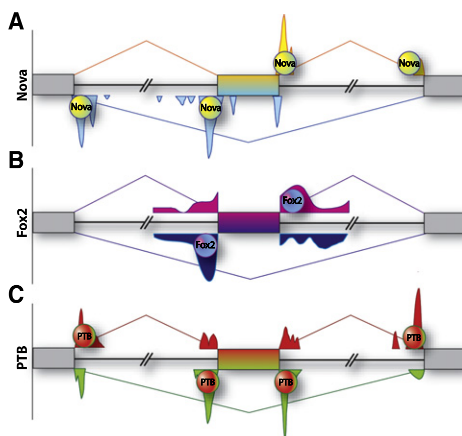

Figure 1. RNA Maps of Three Splicing Regulators
RNA maps depict the correlation between sites of regulatory protein binding (determined by CLIP) relative to a model alternative exon and the outcome of splicing regulation (exon inclusion or skipping). Peaks represent regions with higher density of CLIP reads. Binding of Nova or Fox2 upstream of the exon correlates with exon skipping, whereas binding downstream of the exon correlates with inclusion. For PTB, skipping correlates with binding of the protein near the alternative exon, whereas inclusion correlates with binding near the distal, constitutive sites.
(A) Nova RNA map.
(B) Fox2 RNA map.
(C) PTB RNA map.

REFERENCES

Boutz, P.L., Stoilov, P., Li, Q., Lin, C.-H., Chawla, G., Ostrow, K., Shiue, L., Ares, M., Jr., and Black, D.L. (2007). Genes Dev. *21*, 1636–1652.

Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., et al. (2008). Nature *456*, 464–469.

Oberstrass, F.C., Auweter, S.D., Erat, M., Hargous, Y., Henning, A., Wenter, P., Reymond, L., Amir-Ahmady, B., Pitsch, S., Black, D.L., and Allain, F.H. (2005). Science *309*, 2054–2057.

Singh, R., Valcárcel, J., and Green, M.R. (2005). Science *268*, 1173–1176.

Spellman, R., and Smith, C.W.J. (2006). Trends Biochem. Sci. *31*, 73–76.

Spellman, R., Llorian, M., and Smith, C.W.J. (2007). Mol. Cell *27*, 420–434.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). Science *302*, 1212–1215.

Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J., and Darnell, R.B. (2006). Nature *444*, 580–586.

Xue, Y., Zhou, Y., Wu, T., Zu, T., Ji, X., Kwon, Y.-S., Zhang, C., Yeo, G., Black, D.L., Sun, H., et al. (2009). Mol. Cell *36*, this issue, 996–1006.

Yeo, G.W., Coufal, N.G., Liang, T.Y., Peng, G.E., Fu, X.-D., and Gage, F.H. (2009). Nat. Struct. Mol. Biol. *16*, 130–137.