# RNA Structure Switches RBP Binding

Zheng Luo,[1,2] Qin Yang,[1,2] and Li Yang[1,2,3,*]
[1]Key Laboratory of Computational Biology, CAS Center for Excellence in Brain Science and Intelligence Technology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]School of Life Science and Technology, ShanghaiTech University, Shanghai 200031, China
*Correspondence: liyang@picb.ac.cn
http://dx.doi.org/10.1016/j.molcel.2016.10.006

**RNA sequence motifs are not sufficient for association with RBPs. In this issue of *Molecular Cell*, Taliaferro et al. (2016) demonstrate that, other than sequence motif, RNA secondary structure plays a repressive role on RBP binding, both in vitro and in vivo.**

The association of RNA molecules with RNA binding proteins (RBPs) is crucial for co-/post-transcriptional RNA processing, such as splicing in eukaryotes. Splicing is catalyzed stepwise by a large ribonucleoprotein complex, referred to as the spliceosome, to join exons and remove introns from precursor-messenger RNAs (pre-mRNAs). In higher eukaryotes, alternative splicing (AS) of a single pre-mRNA generates multiple mRNA isoforms and therefore multiple protein products (Nilsen and Graveley, 2010). AS is affected by short *cis*-acting sequence motifs nearby and *trans*-acting RBPs that recognize these short motifs (Nilsen and Graveley, 2010; Ray et al., 2013). Thus, identification of RNA sequence motifs and their association with RBPs is important to allow the prediction of regulated AS events. Systematic evolution of ligands by exponential enrichment (SELEX) has been used to identify short RNA sequence motifs for RBP binding (Jin et al., 2003), but it lacks quantifiability even with a high-throughput version (Lambert et al., 2014). Recently, multiple high-throughput assays combined with computational approaches have been developed to determine the atlas of RBP binding motifs in vitro (e.g., RNAcompete [Ray et al., 2009] and RNA Bind-n-Seq [RBNS; Lambert et al., 2014]) and motifs associated with RBPs in vivo, such as by ultraviolet (UV) crosslinking and immunoprecipitation (CLIP)-seq and individual nucleotide-resolution CLIP (iCLIP)-seq (König et al., 2012). However, obvious discrepancies have been observed between the wide occurrence of predicted RNA binding motifs and the only small subset of them occupied in vivo (Van Nostrand et al., 2016). In this issue of *Molecular Cell*, Taliaferro et al. (2016) show, using systematic analysis combining multiple biochemical, computational, and evolutionary methods, that, other than RNA sequence motif, local RNA secondary structure plays a determinant and repressive role in RBP binding.

RNA motifs in introns flanking conserved AS exons in mammals are more likely to be bound in vivo than those in introns flanking constitutive exons (Merkin et al., 2012). Taliaferro et al. (2016) speculate that other intrinsic context features must be required for RBP binding and have evolved to regulate conserved AS. By focusing on two well-studied RBPs, muscleblind (MBNL1) and RBFOX2, the authors perform upgraded RBNS to detect additional contextual features that may be required to regulate RBP binding. Different from the random oligo pools used in RBNS (Lambert et al., 2014), Taliaferro et al. (2016) use approximately 10,000 naturally occurring intronic RNA sequences for RBP binding ("natural sequence" RBNS or nsRBNS for simplicity). These "natural sequences" are about 110 nt in length, flanking ~3,000 constitutive and alternative spliced exons with varying evolutionary ages (Taliaferro et al., 2016). Compared to other in vitro methods with complex/diverse random oligos, nsRBNS contains a more concentrated pool, thus empowering measurement of natural sequence enrichment (R value) after binding analysis. Also, the 110-nt-long sequences enable assessment of other contextual features, such as RNA secondary structure. In addition, in vitro nsRBNS excludes other uncontrolled effects of a complex cellular environment, allowing a focus on contextual features that regulate RBP binding. Importantly, the application of natural sequences largely reflects binding contexts that occur endogenously, as shown by a good association between in vitro binding and in vivo AS regulation during development (Lambert et al., 2014; Taliaferro et al., 2016).

In agreement with early in vivo observations, Taliaferro et al. (2016) demonstrate that, although having identical core motifs, natural sequences flanking mammalian-wide AS exons are more likely to be bound than those flanking constitutive exons in nsRBNS binding assays with MBNL1 or RBFOX2. Sequence conservation analysis with PhastCons suggests that sequences flanking mammalian-wide AS exons are more conserved than those flanking species-specific AS exons. Computational analysis with RNAstructure and data mining from selective 2′ hydroxyl acylation analyzed by primer extension (SHAPE)-seq indicate that introns flanking mammalian-wide AS exons are less structured and preferentially in single-stranded character in close proximity to RNA motifs, which implies a repressive role of RNA secondary structure in RBP binding. Using a combinatorial mutagenesis RBNS (cmRBNS) assay with comprehensive mutagenesis at each base of examined intronic sequences (except cognate core motifs), they argue that the regulatory impact of motif-centric RNA secondary structure changes on RBP binding (Figure 1): RNA secondary structure with the inclusion of the RNA motif in a double-stranded stem region switches off RBP binding to the

RNA motif, whereas RNA secondary structure releasing the RNA motif to a single-stranded or loop region switches on RBP binding. The decisive role of RNA secondary structure in regulating RBP binding was further confirmed by splicing minigene reporter assays.

The elucidation of RNA secondary structure functioning as a mechanistic determinant for RBP binding by Taliaferro et al. (2016) helps to solve the puzzle regarding the observation that the occurrence of cognate motifs is not sufficient for RBP binding. Thus, the prediction of splicing codes can be significantly improved using additional RNA secondary structure information. It will be of interest to survey the regulation of RNA secondary structure changes on AS changes and their subsequent consequences. For instance, as temperature influences RNA secondary structure, it is possible that AS events can be altered with body temperature change, such as in fever. In addition, it remains to be determined whether RNA secondary structure regulates RBP binding for AS conserved beyond mammals, what additional contextual features account for the determination of AS regulation, and how RNA secondary structures may influence other RBP affinity to their cognate motifs. Due to the limitation of in vitro synthesis, sequences with only 110 nt in length were obtained for the nsRBNS experiment. It is worth noting that mammalian introns generally span hundreds to thousands of nucleotides. Hence, elucidating long-range RNA interactions/structures within the same introns, anchoring adjacent introns/exons or across different introns, will provide additional insights into the RNA-RBP interaction. Indeed, RNA secondary structures may play more complex than repressive roles in AS regulation, suggested by the finding that competition of RNA pairing across different sets of flanking introns is associated with alternative back-splicing (Zhang et al., 2016). Nevertheless, this work by Taliaferro et al. (2016) sheds new light on the role of RNA structure in switching RBP binding, indicating an unexplored "3D-transcriptome" for dynamic gene regulation at the pre-(m)RNA level.

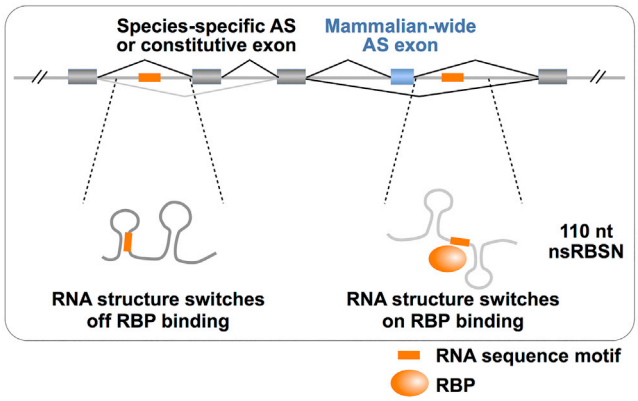


**Figure 1. Local RNA Secondary Structure Plays a Determinant and Repressive Role in RBP Binding**

Together with other approaches, the upgraded "natural sequences" RNA Bind-n-Seq (nsRBNS) in vitro binding assay argues for a regulatory impact of RNA secondary structure on RBP binding. Intronic sequences from species-specific AS or constitutive exons likely form a local RNA secondary structure that includes RNA motif in a double-stranded stem region, leading to a switch-off for RBP binding (left), whereas intronic sequences from mammalian-wide AS exons form a RNA secondary structure that releases RNA motif to a single-stranded or loop region, resulting in a switch-on for RBP binding (right).

## REFERENCES

Jin, Y., Suzuki, H., Maegawa, S., Endo, H., Sugano, S., Hashimoto, K., Yasuda, K., and Inoue, K. (2003). EMBO J. *22*, 905–912.

König, J., Zarnack, K., Luscombe, N.M., and Ule, J. (2012). Nat. Rev. Genet. *13*, 77–83.

Lambert, N., Robertson, A., Jangi, M., McGeary, S., Sharp, P.A., and Burge, C.B. (2014). Mol. Cell *54*, 887–900.

Merkin, J., Russell, C., Chen, P., and Burge, C.B. (2012). Science *338*, 1593–1599.

Nilsen, T.W., and Graveley, B.R. (2010). Nature *463*, 457–463.

Ray, D., Kazan, H., Chan, E.T., Peña Castillo, L., Chaudhry, S., Talukder, S., Blencowe, B.J., Morris, Q., and Hughes, T.R. (2009). Nat. Biotechnol. *27*, 667–670.

Ray, D., Kazan, H., Cook, K.B., Weirauch, M.T., Najafabadi, H.S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., et al. (2013). Nature *499*, 172–177.

Taliaferro, J.M., Lambert, N.J., Sudmant, P.H., Dominguez, D., Merkin, J.J., Alexis, M.S., Bazile, C.A., and Burge, C.B. (2016). Mol. Cell *64*, this issue, 294–306.

Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., et al. (2016). Nat. Methods *13*, 508–514.

Zhang, X.O., Dong, R., Zhang, Y., Zhang, J.L., Luo, Z., Zhang, J., Chen, L.L., and Yang, L. (2016). Genome Res. *26*, 1277–1287.