Sone, M., Hayashi, T., Tarui, H., Agata, K., Takeichi, M., and Nakagawa, S. (2007). J. Cell Sci. *120*, 2498–2506.

Weber, J.D., Taylor, L.J., Roussel, M.F., Sherr, C.J., and Bar-Sagi, D. (1999). Nat. Cell Biol. *1*, 20–26.

West, J.A., Mito, M., Kurosaka, S., Takumi, T., Tanegashima, C., Chujo, T., Yanaka, K., Kingston, R.E., Hirose, T., Bond, C., et al. (2016). J. Cell Biol. *214*, 817–830.

Wilusz, J.E. (2016). Biochim. Biophys. Acta *1859*, 128–138.

Wu, H., Yin, Q.-F., Luo, Z., Yao, R.-W., Zheng, C.-C., Zhang, J., Xiang, J.-F., Yang, L., and Chen, L.-L. (2016). Mol. Cell *64*, this issue, 534–548.

Yin, Q.F., Yang, L., Zhang, Y., Xiang, J.F., Wu, Y.W., Carmichael, G.G., and Chen, L.L. (2012). Mol. Cell *48*, 219–230.

From Protein-RNA Predictions toward a Peptide-RNA Code

Kristopher W. Brannan[1,2] and Gene W. Yeo[1,2,3,4,*]
[1]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92093, USA
[2]Stem Cell Program and Institute for Genomic Medicine, University of California, San Diego, La Jolla, CA 92093, USA
[3]Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
[4]Molecular Engineering Laboratory, A*STAR, Singapore
*Correspondence: geneyeo@ucsd.edu
http://dx.doi.org/10.1016/j.molcel.2016.10.023

The RNA field is undergoing a renaissance, with a deluge of proteins being identified to bind RNA. Two reports now introduce proteome-wide approaches that identify the peptides that are crosslinked to RNA (Castello et al., 2016; He et al., 2016).

Ultraviolet-induced crosslinking causes covalent, irreversible interactions between specific protein domains and RNA molecules (Piñol-Roma et al., 1989). Technologies that leverage crosslinking have been used to identify RNA targets of RNA binding proteins (RBPs) (Jensen and Darnell, 2008). Immobilization of crosslinked RNA-RBP complexes used in combination with mass spectrometry (RNA interactome capture) has enabled identification of the mRNA-bound proteome (Castello et al., 2012), and hybridization of RNA probes followed by mass spectrometry can be used to discover the RBPs bound to specific RNA targets (McHugh et al., 2015). These techniques reveal and confirm known and candidate RBPs, but they lack resolution about the specific peptide regions within these RBPs that directly interact with RNA. To obtain RNA binding domain (RBD) identities at peptide-level resolution, two recent studies utilize shifts in mass spectra induced by UV-crosslinking of protein to RNA to identify RBDs across whole proteomes (Castello et al., 2016; He et al., 2016).

Castello et al. modified RBP interactome capture to include limited protease digestion that releases peptides that are not crosslinked to polyadenylated RNA (Figure 1). This strategy, termed RBDmap, allowed identification of 1,174 RNA binding sites (RBS) contained by 529 RBPs in human cells. The list of RBPs identified by RBDmap largely overlapped with previously defined RNA interactomes and included both canonical and non-canonical RBPs. Peptides enriched in the crosslinked fractions had known characteristics of RBDs, including charged and aromatic residues, and more than half contained disordered regions. Hundreds of the RBDmap-identified RBPs have no previous annotation of RNA associations, excitingly expanding the list of RBPs to include enzymes and chaperones. Newly identified RBDs for these factors are enriched for residues subject to posttranslational modifications, implying new regulatory mechanisms governing RNA-protein interactions. RNA binding domains within enzymatic cores or protein-protein interaction surfaces point to new roles for RNA in competitively influencing enzymatic activity or protein-protein interaction networks. The RBDmap method proves extremely powerful for the identification of RNA binding sites, but its scope is limited to interactions with mature, polyadenylated mRNA.

In a parallel effort, He et al. utilized 4SU incorporation to identify peptide regions that directly contact RNA (Figure 1). They found that changes in peptide mass caused by 4SU incorporation and UV-mediated crosslinking selectively deplete peptides in direct contact with RNA from mass spectrometry peak signals. The authors applied this technique, referred to as RBR-ID, to isolated nuclei from mouse embryonic stem cells and report the identification of 800 RBPs. More than half of these candidate RBPs were not previously annotated to bind RNA. Overlap of RNA binding residues called by RBR-ID with known structures of RNA-protein complexes shows that the method is robust. The expanded list of RBPs from this work predicts new RNA binding activity for many chromatin-associated factors. A key strength of the method is the lower (10–100 times) requirement for input material compared to RBDmap.

However, a limitation by RBR-ID is the reliance on 4SU incorporation for detection of efficient protein-RNA crosslinks. Other nucleotide analogs and crosslinking strategies may lower the false-negative rate.

Taken together, these studies offer insights into the increasingly diverse peptide regions that associate with RNA. The results corroborate recent work identifying enzymes and chromatin-associated proteins that "moonlight" as RBPs (Beckmann et al., 2015; Brannan et al., 2016; Conrad et al., 2016). Because many of these factors found in these studies lack known RBDs, methods such as RBDmap and RBR-ID will be useful tools to identify the specific portions of these proteins that directly interact with RNA. Future applications of these methods may eventually lead to the characterization of new classes of RBDs within non-canonical RBPs. By combining such systematic proteomic classification of new RBDs with RNA-binding site motif information obtained from increasingly more exhaustive maps of RBP-RNA sites at nucleotide level resolution, we are moving closer toward cracking the code of regulatory RNA-peptide interactions.

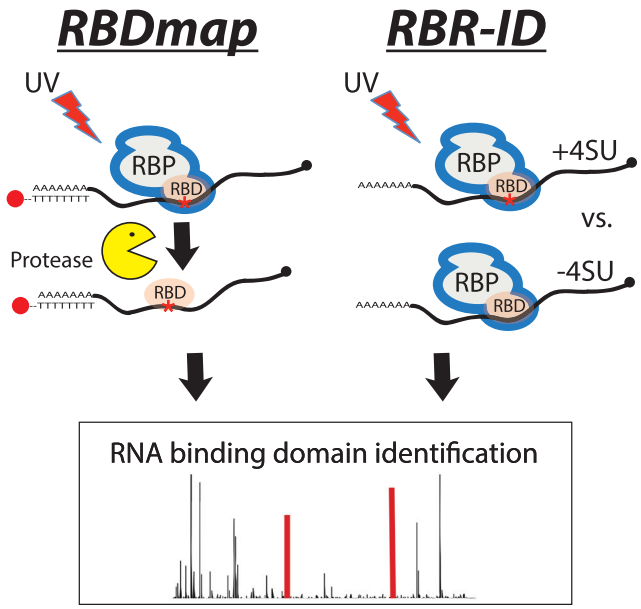

Figure 1. RBDmap and RBR-ID Are Two Methods that Utilize UV-Crosslinking of Protein to RNA Followed by Mass Spectrometry to Identify RNA Binding Domains Proteome-wide

RBDmap uses serial interactome capture along with protease digestion steps to isolate RNA binding sites from crosslinked RBP-RNA complexes. RBR-ID discovers RBDs by depleting 4SU-mediated crosslinked peptides from mass spectra.

REFERENCES

Beckmann, B.M., Horos, R., Fischer, B., Castello, A., Eichelbaum, K., Alleaume, A.M., Schwarzl, T., Curk, T., Foehr, S., Huber, W., et al. (2015). Nat. Commun. *6*, 10127.

Brannan, K.W., Jin, W., Huelga, S.C., Banks, C.A.S., Gilmore, J.M., Florens, L., Washburn, M.P., Van Nostrand, E.L., Pratt, G.A., Schwinn, M.K., et al. (2016). Mol. Cell *64*, 282–293.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., et al. (2012). Cell *149*, 1393–1406.

Castello, A., Fischer, B., Frese, C.K., Horos, R., Alleaume, A.M., Foehr, S., Curk, T., Krijgsveld, J., and Hentze, M.W. (2016). Mol. Cell *63*, 696–710.

Conrad, T., Albrecht, A.S., de Melo Costa, V.R., Sauer, S., Meierhofer, D., and Ørom, U.A. (2016). Nat. Commun. *7*, 11212.

He, C., Sidoli, S., Warneford-Thomson, R., Tatomer, D.C., Wilusz, J.E., Garcia, B.A., and Bonasio, R. (2016). Mol. Cell *64*, 416–430.

Jensen, K.B., and Darnell, R.B. (2008). Methods Mol. Biol. *488*, 85–98.

McHugh, C.A., Chen, C.K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., et al. (2015). Nature *521*, 232–236.

Piñol-Roma, S., Adam, S.A., Choi, Y.D., and Dreyfuss, G. (1989). Methods Enzymol. *180*, 410–418.