Video Article

Sample Preparation for Mass Spectrometry-based Identification of RNA-binding Regions

Robert Warneford-Thomson[1,4], Chongsheng He[1,2], Simone Sidoli[1,3], Benjamin A. Garcia[1,3], Roberto Bonasio[1,2]

[1]Epigenetics Institute, University of Pennsylvania Perelman School of Medicine

[2]Department of Cell and Developmental Biology, University of Pennsylvania Perelman School of Medicine

[3]Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine

[4]Graduate Group in Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine

Correspondence to: Roberto Bonasio at rbon@mail.med.upenn.edu

URL: https://www.jove.com/video/56004
DOI: doi:10.3791/56004

Keywords: Biochemistry, Issue 127, Protein-RNA interactions, RNA-binding domain, RNA-binding region, mass spectrometry, noncoding RNAs, UV crosslinking

Date Published: 9/28/2017

Citation: Warneford-Thomson, R., He, C., Sidoli, S., Garcia, B.A., Bonasio, R. Sample Preparation for Mass Spectrometry-based Identification of RNA-binding Regions. *J. Vis. Exp.* (127), e56004, doi:10.3791/56004 (2017).

Abstract

Noncoding RNAs play important roles in several nuclear processes, including regulating gene expression, chromatin structure, and DNA repair. In most cases, the action of noncoding RNAs is mediated by proteins whose functions are in turn regulated by these interactions with noncoding RNAs. Consistent with this, a growing number of proteins involved in nuclear functions have been reported to bind RNA and in a few cases the RNA-binding regions of these proteins have been mapped, often through laborious, candidate-based methods.

Here, we report a detailed protocol to perform a high-throughput, proteome-wide unbiased identification of RNA-binding proteins and their RNA-binding regions. The methodology relies on the incorporation of a photoreactive uridine analog in the cellular RNA, followed by UV-mediated protein-RNA crosslinking, and mass spectrometry analyses to reveal RNA-crosslinked peptides within the proteome. Although we describe the procedure for mouse embryonic stem cells, the protocol should be easily adapted to a variety of cultured cells.

Video Link

The video component of this article can be found at https://www.jove.com/video/56004/

Introduction

The purpose of the RBR-ID method is to identify novel RNA-binding proteins (RBPs) and map their RNA-binding regions (RBRs) with peptide-level resolution to facilitate the design of RNA-binding mutants and the investigation of the biological and biochemical functions of protein-RNA interactions.

RNA is unique among biomolecules as it can both act as a messenger carrying genetic information and also fold into complex three-dimensional structures with biochemical functions more akin to those of proteins[1,2]. A growing body of evidence suggests that noncoding RNAs (ncRNAs) play important roles in various gene regulatory and epigenetic pathways[3,4,5] and, typically, these regulatory functions are mediated in concert with proteins that interact specifically with a given RNA. Of particular relevance, a set of interacting proteins was recently identified for the intensely studied long ncRNA (lncRNA) Xist, providing valuable insight into how this lncRNA mediates X-chromosome inactivation in female cells[6,7,8]. Notably, several of these Xist-interacting proteins do not contain any canonical RNA-binding domains[9], and therefore their RNA-binding activity could not be predicted *in silico* based on their primary sequence alone. Considering that thousands of lncRNAs are expressed in any given cell[10], it is reasonable to assume that many of them might act via interactions with yet to be discovered RNA-binding proteins (RBPs). An experimental strategy to identify these novel RBPs would therefore greatly facilitate the task of dissecting the biological function of ncRNAs.

Previous attempts to identify RBPs empirically have relied on polyA+ RNA selection coupled to mass spectrometry (MS)[11,12,13,14,15]. Although these experiments added many proteins to the list of putative RBPs, by design they could only detect proteins bound to polyadenylated transcripts. However, most small RNAs and many lncRNAs are not polyadenylated.[16,17] and their interacting proteins would likely have been missed in these experiments. A recent study applied machine learning to protein-protein interactome databases to identify proteins that co-purified with multiple known RBPs and showed that these recurrent RBP partners were more likely to possess RNA-binding activities[18]. However, this approach relies on mining existing large interaction databases and can only identify proteins that can be co-purified in non-denaturing conditions with known RBPs, thus excluding from the analysis insoluble, membrane-embedded, and scarce proteins.

The identification of a protein as a *bona fide* RBP often does not automatically yield information on the biological and/or biochemical function of the protein-RNA interaction. To address this point, it is typically desirable to identify the protein domain and amino acid residues involved in the interaction so that specific mutants can be designed to test the function of RNA binding in the context of each novel RBP[19,20]. Previous efforts by our group and others have used recombinant protein fragments and deletion mutants to identify RNA binding regions (RBRs)[19,20,21,22]; however,

Copyright © 2017 Journal of Visualized Experiments

such approaches are labor-intensive and incompatible with high-throughput analyses. More recently, a study described an experimental strategy to map RNA-binding activities in a high-throughput fashion using mass spectrometry[23]; however, this approach relied on a double polyA+ RNA selection, and thus carried the same limitations as the RBP identification approaches described above.

We developed a technique, termed RNA binding region identification (RBR-ID), which exploits protein-RNA photocrosslinking and quantitative mass spectrometry to identify proteins and protein regions interacting with RNA in live cells without making assumption on the RNA polyadenylation status, thus including RBPs bound to polyA- RNAs[24]. Moreover, this method relies exclusively on crosslinking and has no requirements on protein solubility or accessibility and is thus suitable to map RNA-binding activities within membranes (*e.g.* the nuclear envelope) or poorly soluble compartments (*e.g.* the nuclear matrix). We describe the experimental steps to perform RBR-ID for the nuclei of mouse embryonic stem cells (mESCs) but with minor modifications this protocol should be suitable for a variety of cell types, provided that they can efficiently incorporate 4SU from the culture medium.

Protocol

1. Culture and Expansion of mESCs

NOTE: Mouse embryonic stem cells are easy to culture and can be quickly expanded to the large numbers required by biochemical experiments thanks to their fast cycling time. Healthy mESCs double every 12 h.

1. Expand mESCs to the desired number on gelatinized plates in mESC medium (see below) in a tissue culture incubator kept at 37 °C, 5% CO_2, and >95% humidity.
   NOTE: Enough plates should be prepared for 3 - 5 biological replicates for the +4SU sample and the -4SU control. One 10-cm plate (*i.e.* 5 - 10 million cells) is more than sufficient for each biological replicate.
2. The medium for mESCs is composed of Dulbecco's modified Eagle's medium (DMEM) supplemented with 15% fetal bovine serum (FBS), 2 mM L-glutamine, 0.5x penicillin/streptomycin, 1x non-essential amino acids, 100 µM beta-mercaptoethanol, 1,000 U/mL leukemia inhibitory factor (LIF), 3 µM CHIR99021, 1 µM PD0325901.
3. **Before passaging mESCs, prepare the desired number of 10 cm plates.**
   1. Add 4 mL of 0.1% gelatin in water making sure that the entire surface of the plate is covered; incubate for 10 min at room temperature (RT); remove gelatin solution.
4. **Passage cells when they reach ~10 million cells per plate, typically ~48 h following inoculation.**
   NOTE: As mESCs grow in dense colonies, plates should never be allowed to reach 100% confluency as observed for cell lines that grow in monolayers (*e.g.* HEK293 or 3T3 cells). An mESC plate should be passaged when ~50% confluent and definitely before the medium turns yellow.
   1. To detach cells, wash plates with dense mESC cultures once in 1x phosphate buffered saline (PBS) supplemented with 2 mM EDTA, then add 0.05% trypsin dissolved in DMEM and return plate to incubator for 5 min (37 °C, 5% CO_2, and >95% humidity).
   2. Remove plate from incubator and resuspend cells with equal volume of DMEM by pipetting up and down vigorously several times.
   3. Count cells using hemocytometer and ensure that cells form a single-cell suspension (*i.e.* no clusters), then centrifuge cell mixture 500 x g for 5 min at room temperature. Remove supernatant and resuspend cell pellet in 1 - 2 mL DMEM.
   4. Inoculate cells at a density of 10,000 cells/cm^2 in mESC medium (*i.e.* ~800,000 cells per 10-cm plate) using the gelatin-coated plates prepared in step 1.3.1.

2. Crosslink of Protein–RNA Interactions in Live Cells

NOTE: RNA-protein crosslinking is mediated by the photo-activatable ribonucleoside analog 4-thiouridine (4SU). 4SU has a longer absorbance maximum than endogenous nucleotides and can only be incorporated into RNA; therefore intermediate-wavelength UVB can be used to selectively crosslink RNA to proteins[25,26]. UVB treatment of 4SU-treated cells leads to covalent crosslinks between 4SU-containing RNA and amino acids, with a reported preference for Tyr, Trp, Met, Lys, and Cys[27].

1. Expand mESCs to the required number of 10-cm plates as described in step 1.
   NOTE: Enough plates should be prepared for 3 - 5 biological replicates for the +4SU sample and the -4SU control. One 10-cm plate (*i.e.* 5 - 10 million cells) is sufficient for each biological replicate.
2. Before the plates reach confluency, add 4SU directly into medium to a final concentration of 500 µM. Leave -4SU control dishes untreated.
3. Shake plates gently to distribute the 4SU homogeneously throughout the medium. Return them to the same tissue culture incubator for 2 h (37 °C, 5% CO_2, and >95% humidity).
4. Transfer plates to an ice bucket and discard medium.
5. Rinse the plate gently with 5 mL of ice-cold PBS.
   NOTE: Rinse gently or cells might detach.
6. Add 2 mL of ice-cold PBS and place plates without lids in a crosslinker equipped with UVB-emitting bulbs (312 nm). Irradiate plates at an energy setting of 1 J/cm^2.
   NOTE: It is critical that plastic lids are removed or they might shield the cells from UV light and prevent crosslinking.
7. Collect cells using cell lifters and transfer to ice-cold PBS in conical tubes. Centrifuge at 2,000 x g for 5 min, 4 °C.
8. Remove the supernatant and resuspend cell pellet in 5 mL ice-cold PBS with 2 mM EDTA and 0.2 mM phenylmethylsulfonyl fluoride (PMSF).
9. Count cells using a hemocytometer. Expect 5 - 10 and 10 - 20 million cells from 10-cm and 15-cm plates respectively.
10. Centrifuge at 2,000 x g for 5 min, 4 °C.

Copyright © 2017 Journal of Visualized Experiments

NOTE: Pelleted cells can be immediately used for step 3 or they can be flash-frozen in liquid nitrogen and stored at -80 °C indefinitely. This is a potential stopping point.

3. Isolation of Nuclei

NOTE: Nuclei are isolated to remove cytoplasmic proteins and increase coverage of nuclear proteins. This step can be replaced with other forms of cellular fractionation to study RBPs in different cellular compartments.

1. **Prepare ice cold Buffer A (10 mM Tris-HCl pH 7.9, 4 °C, 1.5 mM $MgCl_2$, 10 mM KCl). Use at least 2.5 mL per 10 million cells (as counted in step 2.9) for extraction.**
   NOTE: Stocks of Buffer A (e.g. a 500 mL bottle) can be prepared in advance and stored at 4 °C for 6 months.
   1. Prior to use, add 0.5 mM dithiothreitol (DTT) and 0.2 mM PMSF to the desired amount (2.5 mL per 10 million cells) of ice-cold buffer A.
2. Wash cells using 2 mL Buffer A per 10 million cells. Spin at 2,500 x g for 5 min at 4 °C.
3. Discard supernatant, add Buffer A supplemented with 0.2% of a nonionic, non-denaturing detergent, such as octyl-phenoxy-polyethoxy-ethanol (see the **Table of Materials**), 500 µL per 10 million cells. Rotate for 5 min at 4 °C.
4. Centrifuge at 2,500 x g for 5 min.
5. Remove the supernatant; the pellet comprises intact nuclei devoid of cytoplasm.
   NOTE: Nuclei can be used immediately for step 4 or flash-frozen in liquid nitrogen and stored at -80 °C indefinitely. This is a potential stopping point.

4. Lysis of Nuclei

NOTE: Crosslinked nuclei are lysed in a mass spectrometry-compatible buffer to release proteins and protein-RNA complexes.

1. Add lysis buffer (9 M urea, 100 mM Tris-HCl pH 8.0, 24 °C), 200 µL per 10 million cells. Sonicate to shear chromatin using a 1/8" probe at 50% amplitude setting for 5 - 10 s.
2. Spin at 12,000 x g for 5 min and collect 175 µL of supernatant to avoid the pellet.
   NOTE: 100 µL will be used for the next RBR-ID steps and the remaining 75 µL of lysate can be stored at -80 °C for later use and/or western blot analysis.
3. Measure protein concentration via Bradford assay or similar technique[28]. Dilute protein concentration in the different samples to 1 mg/mL or the lowest sample concentration using appropriate amounts of lysis buffer.
   NOTE: From the nuclei of 10 million mESCs expect 100 - 200 µg of protein.

5. Trypsin Digestion

NOTE: Proteins are digested to generate peptides suitable for bottom-up mass spectrometry (MS) analysis.

1. Add DTT solution to nuclear lysate (5 mM final); incubate for 1 h at RT.
2. Add iodoacetamide from fresh stock (14 mM final); incubate for 30 min at RT in the dark.
3. Dilute lysate with 5 times volume of 50 mM NH_4HCO_3
4. Add trypsin (µg protein : µg trypsin = 200 : 1); incubate at 37 °C overnight.

6. Desalting of peptides

1. **Prepare one custom-made stage tip per sample.**
   1. 6.1.1. Place a disk of solid phase extraction C18 material into the bottom of a 200 µL pipette tip.
   2. Add 50 µL of oligo R3 reversed phase resin on top of the C18 disk. Place stage tip inside centrifugation adaptor and the place tip with adaptor inside a 1.5 mL microfuge tube.
   3. Add 100 µL 100% acetonitrile to tips to wash. Centrifuge at 1000 x g for 1 min to remove the solvent.
   4. Equilibrate with 50 µL 0.1% trifluoroacetic acid (TFA). Centrifuge at 1000 x g for 1 min.
2. Add 100% formic acid to digested protein sample to decrease the pH to 2 - 3.
   NOTE: This should require ~5 µL per 1 mL of diluted peptide sample, but pH should be measured and more formic acid added if necessary.
3. Load sample onto custom-made stage tip, centrifuge at 1000 x g for 1 min until all of sample goes through the tip.
4. Wash bound peptides with 50 µL 0.1% TFA. Centrifuge at 1000 x g for 1 min.
5. Elute peptides by adding 50 µL elution buffer (75% acetonitrile, 0.025% TFA) to stage tip and centrifuge at 1000 x g for 1 min to force the elution buffer out of the tip.
6. Dry the sample using a vacuum centrifuge set at 150 x g and 1 - 3 kPa for 1 h.
   NOTE: Dried peptides can be stored at -80 °C indefinitely. This is a potential stopping point.

7. Removal of Crosslinked RNA

NOTE: Treat peptides with nuclease to remove crosslinked RNA.

1. Resuspend the pellet in 50 µL 2x nuclease buffer (100 mM NH_4HCO_3, 4 mM $MgCl_2$).
2. Measure peptide concentration using absorbance at 280 nm assuming 1 mg/mL of peptides for an A_{280} of 1.
   NOTE: The expected concentration is 1 mg/mL.

3. Adjust all samples to 1 mg/mL or to the lowest concentration using 2x nuclease buffer.
4. Prepare a master mix of 50 µL H_2O + 1 µL high purity nuclease (≥250 U) (refer to the **Table of Materials**) per sample.
5. Take 50 µL of peptide sample and add 50 µL of H_2O + nuclease master mix.
6. Incubate at 37 °C for 1 h. Proceed to perform mass spectrometry.
   NOTE: The peptides are now ready for mass spectrometry. They can be analyzed immediately or stored at -80 °C indefinitely. This is a potential stopping point.

8. Nano Liquid Chromatography, Mass Spectrometry, and Raw Data Processing

NOTE: Because 4SU-crosslinking changes the mass of the peptide, their ions do not count toward the intensity of the non-crosslinked peptide during LC-MS/MS, which therefore appears to be decreased by the crosslinking. The degree and consistency of this decrease reflects the degree of protein-RNA crosslinking for each peptide[24].

1. Prepare HPLC buffers as follows: 0.1% formic acid in HPLC-grade water (buffer A); 0.1% formic acid in HPLC-grade acetonitrile (Buffer B).
2. If a nano-HPLC is available, connect a nano-column with the following specifications: internal diameter 75 µm; packed with C18-AQ particles; length 20 - 25 cm.
   NOTE: If the HPLC available runs at higher flow rates use a proper column size for the indicated flow rate. Higher flow chromatography decreases sensitivity.
3. Program the HPLC method as follows: flow rate 250 - 300 nL/min; gradient from 2 to 28% buffer B in 120 min, from 28 to 80% B for the next 5 min and isocratic 80% B for 10 min.
   NOTE: If the HPLC does not have an automated column equilibration prior sample loading, start the program with 10 min at 0% B before loading the sample.
4. Program the MS acquisition method to perform data-dependent acquisition (DDA) as standard shotgun proteomics experiments. The instrument duty cycle should alternate full MS scans with tandem MS scans (or MS/MS) of the most intense signals. Use the optimal settings for proteomics MS runs.
   NOTE: For confident results, use instruments that can perform at least the full MS scan in high resolution mode (*e.g.* orbitrap, time-of-flight). For accurate quantification, make sure that intervals between full MS scans are no longer than 3 s. Longer duty cycles might prevent accurate definition of extracted ion chromatograms.
5. Load 1 - 2 µg of sample onto the HPLC column and run the HPLC-MS/MS method as programmed. For each biological replicate perform at least two independent runs and treat them as technical replicates.
6. After MS, import the raw files into a proteomics software package for MS/MS spectra identification and peptide quantification via extracted ion chromatography. We recommend software packages that can perform chromatographic alignment of multiple MS runs (see the **Table of Materials**)[29,30].
7. If the option is available in the software suite, enable "Match Between Runs" prior to starting data processing.
8. Once the analysis is finished, export the peptide list together with the quantification values.

Representative Results

Figure 1 depicts the RBR-ID workflow. Due to the relatively low crosslinking efficiency of this technique, it is very important to consider both the depletion level and consistency of the observed effect (*P*-value) across biological replicates. **Figure 2** shows a volcano plot of RBR-ID result. Peptides that overlapped RNA recognition motif (RRM) domain show highly consistent depletion level. RRM domains can be used as a positive control for RBR-ID analysis.

RBR-ID can be used to map RBRs in live cells. An RBR-ID score can be calculated for each peptide to estimate RNA-binding potential: RBR-ID score = $-\log_2$(normalized +4SU intensity/normalized -4SU intensity) x $(\log_{10}(P\text{-value}))^2$. **Figure 3** shows a heatmap of RBR-ID scores projected on the surface of the spliceosomal subunit U1-70k. The bright red color in vicinity of the RNA contact, as determined from the crystal structure, indicates correct identification of the protein-RNA interaction.

Upon identification of novel RBPs and their RBRs, it is highly recommended that their RNA-binding activity be confirmed by an independent method. In our previous study[24], we identified TET2 as a novel RBP and mapped the RNA-binding activity to a C-terminal region, as shown in **Figure 4A** by plotting the RBR-ID score along the primary sequence of the protein. **Figure 4B** shows that the requirement for this novel RBR could be verified by performing PAR-CLIP[26] using the WT sequence and comparing the signal to that of a mutant lacking the predicted RBR. Additional control and more detailed explanation of this validation experiment are available in He *et al.* 2016[24].

Copyright © 2017 Journal of Visualized Experiments

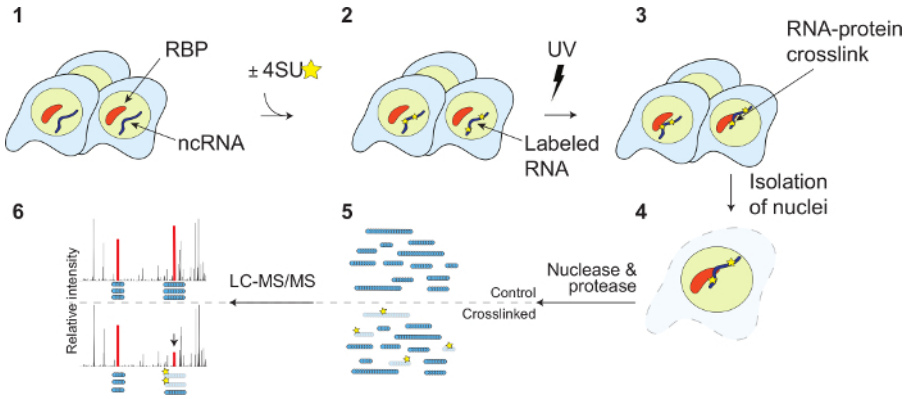

Figure 1: RBR-ID overview. Mouse ESCs are treated with 4SU or not (1-2) and irradiated with 312 nm UV (3). Nuclei are isolated (4) and extracts digested with protease and nuclease, yielding both crosslinked and uncrosslinked peptides (5). Covalent RNA adducts at crosslink sites alter the peptide mass, leading to a corresponding decrease in intensity of the uncrosslinked peptide's mass spectrum (6, arrow). Please click here to view a larger version of this figure.

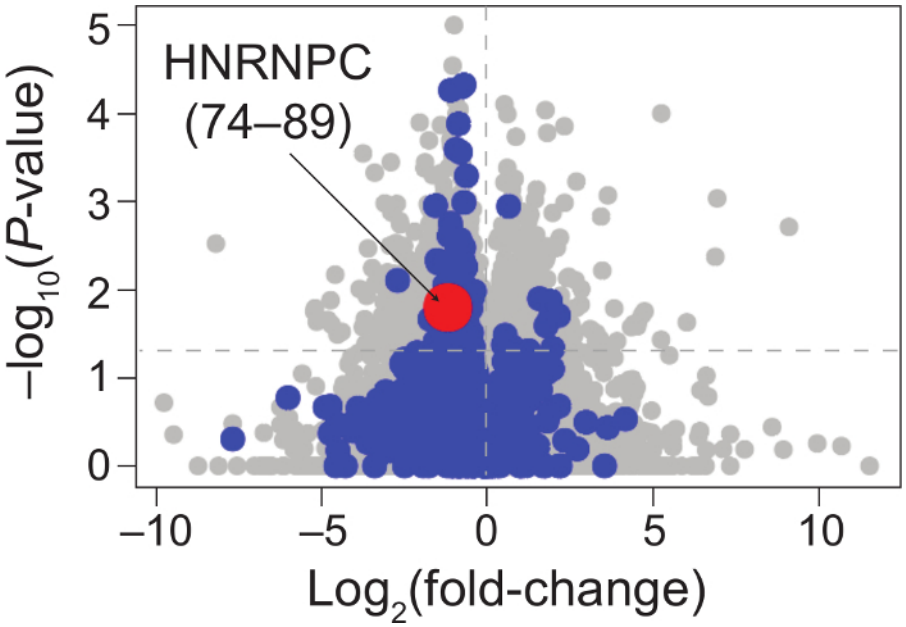

Figure 2: Proteome-wide RBR-ID in mESCs. Volcano plots showing log-fold changes in peptide intensities on the *x* axis and Student's *t*-test *P*-values on the *y* axis for ± 4SU treatments (312 nm). Peptides overlapping annotated RRM domains are in blue. An RNA-binding peptide from HNRNPC is highlighted in red. Previously published in He *et al.* 2016[24]. Please click here to view a larger version of this figure.

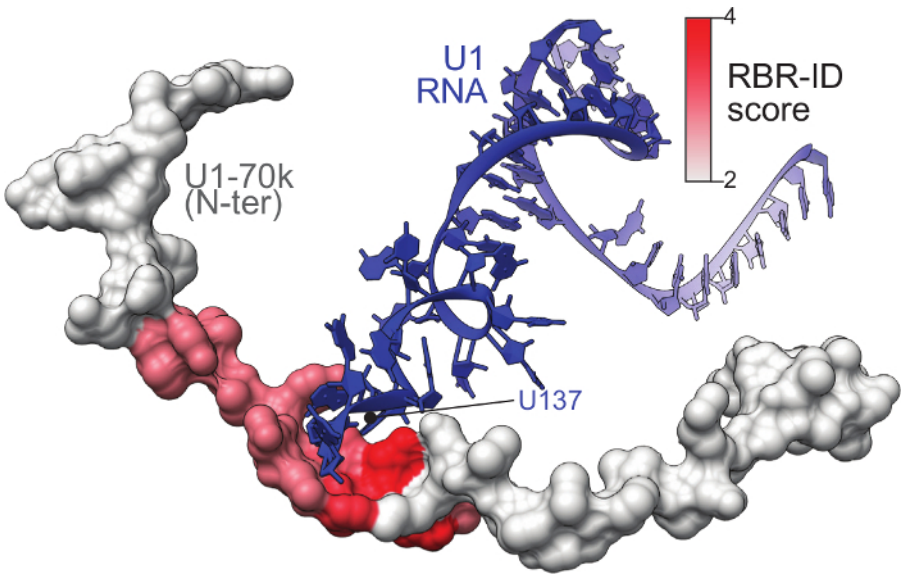

Figure 3: RBR-ID maps the sites of protein–RNA interactions. Zoomed-in regions of the crystal structure of U1 snRNP (PDB ID: 4PJO[31]) showing protein surfaces color-coded according to their RBR-ID score and interacting RNAs for U1-70K. Previously published in He *et al.* 2016[24]. Please click here to view a larger version of this figure.

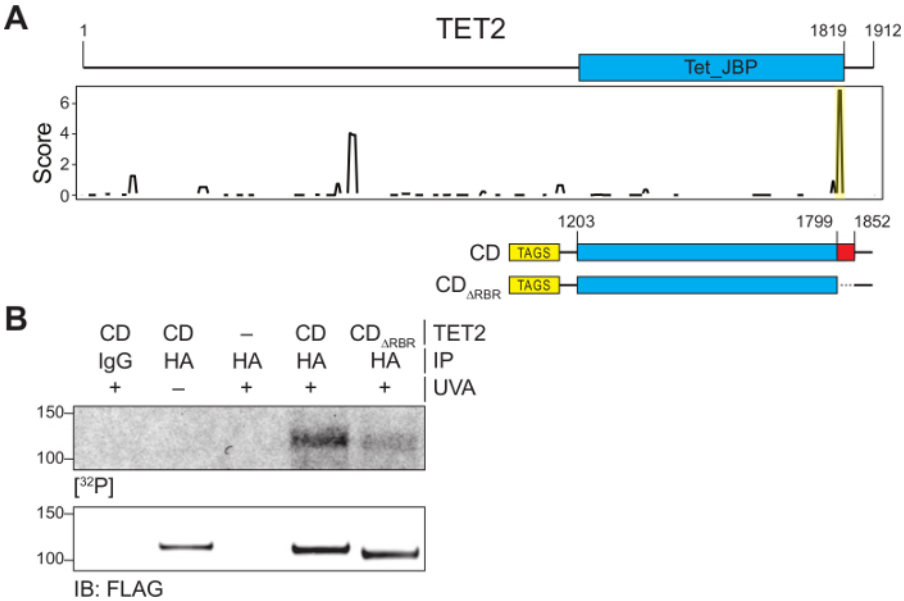

Figure 4: Validation of the RBR of TET2. (**A**) Primary sequence and known domains for TET2; smoothed residue-level RBR-ID score plotted along the primary sequence (middle); and scheme of epitope-tagged catalytic domain fragment (CD) and RBR-deleted (CDΔRBR) constructs used for validation (bottom). (**B**) PAR-CLIP of transiently expressed TET2 CD and $\Delta CD_{\Delta RBR}$ in HEK293 cells. Autoradiography for ^{32}P-labeled RNA (top) and control western blot (bottom). Previously published in He *et al.* 2016[24]. Please click here to view a larger version of this figure.

Discussion

We describe a detailed experimental protocol to perform RBR-ID in mESCs and, with appropriate modifications, in any cell that can incorporate 4SU into RNA. Other cell types may require optimization of the approach to ensure a sufficient signal to noise ratio. Additionally, while the protocol described herein focuses on the examination of nuclear RBPs, the RBR-ID technology should be easily adapted to different cellular compartments, such as the cytosol or specific organelles, by use of different fractionation strategies. Parameters that may require optimization include 4SU concentration and incorporation time as well as the energy of UV crosslinking. These parameters are important to ensure efficient formation of RNA-protein crosslinks and yield a satisfactory signal to noise ratio. We have shown that 312 nm UVB light is much more efficient at crosslinking 4SU-containing RNAs to proteins compared to the 365 nm UVA typically used in PAR-CLIP[32]. Direct comparison of RBR-ID performed with 312 nm and 365 nm demonstrated that 312 nm UVB resulted in the identification of a much larger portion of known RBPs[24].

Two other methods are available to perform proteome-wide RBP identification. One, called RBDmap, is similar to RBR-ID in that it utilizes UV crosslinking and MS. RBDmap was used to identify RBPs and map their RNA-binding activities in HeLa cells[23]. This method relies on positive identification of peptides adjacent to the RNA-binding sites and relies on two sequential oligo-dT pull-downs, suggesting that it might have a

lower false positive rate than RBR-ID. However, the pull-downs allow only the identification of RBPs that bind to polyA+ RNA and require large amounts of input material, up to 10 - 100 times more than RBR-ID. RBR-ID can be performed with as little as 5 µg of protein per biological replicate (~2 µg per technical replicate), which can be collected from as few as 250,000 cells.

The second method, termed SONAR, relies on data mining of large protein-protein interaction databases to detect proteins that frequently co-purify with known RBPs and whose interaction with these RBPs might therefore be mediated by RNA[18]. This clever approach is very powerful and allows for the identification of novel RBPs without performing any wet lab experiment, provided that suitable databases are available; however, it cannot reveal the interaction sites and therefore is complementary but not alternative to RBR-ID.

In summary, our RBR-ID technique can identify novel RBPs and map their RBRs with limited input sample amounts and without requiring any biochemical purification of the protein-RNA crosslinked complexes. The technique makes no assumption on the type of RNA bound by the identified RBPs and therefore can map the RNA binding activity of proteins that bind to non-polyadenylated, many of which are noncoding, suggesting that RBR-ID might prove a useful technique to study the regulatory roles of noncoding RNAs.

Disclosures

The authors have nothing to disclose.

Acknowledgements

R.B. was supported by the Searle Scholars Program, the W.W. Smith Foundation (C1404), and the March of Dimes Foundation (1-FY-15-344). B.A.G acknowledges support from NIH grants R01GM110174 and NIH R01AI118891, as well as DOD grant BC123187P1. R.W.-T. was supported by NIH training grant T32GM008216.

References

1. Bonasio, R., & Shiekhattar, R. Regulation of transcription by long noncoding RNAs. *Annu Rev Genet.* **48** 433-455 (2014).
2. Bonasio, R. Emerging topics in epigenetics: ants, brains, and noncoding RNAs. *Ann N Y Acad Sci.* **1260** 14-23 (2012).
3. Rinn, J. L., & Chang, H. Y. Genome regulation by long noncoding RNAs. *Annu Rev Biochem.* **81** 145-166 (2012).
4. Goff, L. A., & Rinn, J. L. Linking RNA biology to lncRNAs. *Genome Res.* **25** (10), 1456-1465 (2015).
5. Holoch, D., & Moazed, D. RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet.* **16** (2), 71-84 (2015).
6. Chu, C. *et al.* Systematic discovery of Xist RNA binding proteins. *Cell.* **161** (2), 404-416 (2015).
7. McHugh, C. A. *et al.* The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature.* **521** (7551), 232-236 (2015).
8. Minajigi, A. *et al.* Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science.* **349** (6245) (2015).
9. Lunde, B. M., Moore, C., & Varani, G. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol.* **8** (6), 479-490 (2007).
10. Cabili, M. N. *et al.* Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **25** (18), 1915-1927, (2011).
11. Castello, A. *et al.* Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell.* **149** (6), 1393-1406 (2012).
12. Baltz, A. G. *et al.* The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell.* **46** (5), 674-690 (2012).
13. Kwon, S. C. *et al.* The RNA-binding protein repertoire of embryonic stem cells. *Nat Struct Mol Biol.* **20** (9), 1122-1130 (2013).
14. Beckmann, B. M. *et al.* The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat Commun.* **6** 10127 (2015).
15. Conrad, T. *et al.* Serial interactome capture of the human cell nucleus. *Nat Commun.* **7** 11212 (2016).
16. Wilusz, J. E. *et al.* A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev.* **26** (21), 2392-2407 (2012).
17. Kim, T. K., & Shiekhattar, R. Architectural and Functional Commonalities between Enhancers and Promoters. *Cell.* **162** (5), 948-959 (2015).
18. Brannan, K. W. *et al.* SONAR Discovers RNA-Binding Proteins from Analysis of Large-Scale Protein-Protein Interactomes. *Mol Cell.* **64** (2), 282-293 (2016).
19. Bonasio, R. *et al.* Interactions with RNA direct the Polycomb group protein SCML2 to chromatin where it represses target genes. *Elife.* **3** e02637 (2014).
20. Kaneko, S. *et al.* Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. *Mol Cell.* **53** (2), 290-300 (2014).
21. Kaneko, S. *et al.* Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. *Genes Dev.* **24** (23), 2615-2620 (2010).
22. Saldana-Meyer, R. *et al.* CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev.* **28** (7), 723-734 (2014).
23. Castello, A. *et al.* Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol Cell.* **63** (4), 696-710 (2016).
24. He, C. *et al.* High-Resolution Mapping of RNA-Binding Regions in the Nuclear Proteome of Embryonic Stem Cells. *Mol Cell.* **64** (2), 416-430 (2016).
25. Favre, A. *et al.* 4-Thiouridine photosensitized RNA-protein crosslinking in mammalian cells. *Biochem Biophys Res Commun.* **141** (2), 847-854 (1986).
26. Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell.* **141** (1), 129-141 (2010).
27. Favre, A., Saintome, C., Fourrey, J. L., Clivio, P., & Laugaa, P. Thionucleobases as intrinsic photoaffinity probes of nucleic acid structure and nucleic acid-protein interactions. *J Photochem Photobiol B.* **42** (2), 109-124 (1998).

Copyright © 2017 Journal of Visualized Experiments

28. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* **72** 248-254 (1976).
29. Tyanova, S., Temu, T., & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc.* **11** (12), 2301-2319 (2016).
30. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics.* **26** (7), 966-968 (2010).
31. Kondo, Y., Oubridge, C., van Roon, A. M., & Nagai, K. Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition. *Elife.* **4** (2015).
32. Hafner, M. *et al.* PAR-CliP--a method to identify transcriptome-wide the binding sites of RNA binding proteins. *J Vis Exp.* (41) (2010).

Copyright © 2017 Journal of Visualized Experiments