

## Video Article

# PAR-CLIP - A Method to Identify Transcriptome-wide the Binding Sites of RNA Binding Proteins

Markus Hafner<sup>1</sup>, Markus Landthaler<sup>2</sup>, Lukas Burger<sup>3</sup>, Mohsen Khorshid<sup>3</sup>, Jean Hausser<sup>4</sup>, Philipp Berninger<sup>4</sup>, Andrea Rothballer<sup>1</sup>, Manuel Ascano<sup>1</sup>, Anna-Carina Jungkamp<sup>2</sup>, Mathias Munschauer<sup>2</sup>, Alexander Ulrich<sup>1</sup>, Greg S. Wardle<sup>1</sup>, Scott Dewell<sup>5</sup>, Mihaela Zavolan<sup>3</sup>, Thomas Tuschl<sup>1</sup>

<sup>1</sup>Howard Hughes Medical Institute, Laboratory of RNA Molecular Biology, Rockefeller University

<sup>2</sup>Berlin Institute for Medical Systems Biology, Max-Delbrück-Center for Molecular Medicine

<sup>3</sup>Biozentrum der Universität Basel and Swiss Institute of Bioinformatics (SIB)

<sup>4</sup>Biozentrum der Universität Basel and Swiss Institute of Bioinformatics (SIB)

<sup>5</sup>Genomics Resource Center, Rockefeller University

URL: <http://www.jove.com/video/2034>

DOI: [doi:10.3791/2034](https://doi.org/10.3791/2034)

Keywords: Cellular Biology, Issue 41, UV crosslinking, RNA binding proteins, RNA binding motif, 4-thiouridine, 6-thioguanosine

Date Published: 7/2/2010

Citation: Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.C., Munschauer, M., Ulrich, A., Wardle, G.S., Dewell, S., Zavolan, M., Tuschl, T. PAR-CLIP - A Method to Identify Transcriptome-wide the Binding Sites of RNA Binding Proteins. *J. Vis. Exp.* (41), e2034, doi:10.3791/2034 (2010).

## Abstract

RNA transcripts are subjected to post-transcriptional gene regulation by interacting with hundreds of RNA-binding proteins (RBPs) and microRNA-containing ribonucleoprotein complexes (miRNPs) that are often expressed in a cell-type dependently. To understand how the interplay of these RNA-binding factors affects the regulation of individual transcripts, high resolution maps of *in vivo* protein-RNA interactions are necessary<sup>1</sup>.

A combination of genetic, biochemical and computational approaches are typically applied to identify RNA-RBP or RNA-RNP interactions. Microarray profiling of RNAs associated with immunopurified RBPs (RIP-Chip)<sup>2</sup> defines targets at a transcriptome level, but its application is limited to the characterization of kinetically stable interactions and only in rare cases<sup>3,4</sup> allows to identify the RBP recognition element (RRE) within the long target RNA. More direct RBP target site information is obtained by combining *in vivo* UV crosslinking<sup>5,6</sup> with immunoprecipitation<sup>7-9</sup> followed by the isolation of crosslinked RNA segments and cDNA sequencing (CLIP)<sup>10</sup>. CLIP was used to identify targets of a number of RBPs<sup>11-17</sup>. However, CLIP is limited by the low efficiency of UV 254 nm RNA-protein crosslinking, and the location of the crosslink is not readily identifiable within the sequenced crosslinked fragments, making it difficult to separate UV-crosslinked target RNA segments from background non-crosslinked RNA fragments also present in the sample.

We developed a powerful cell-based crosslinking approach to determine at high resolution and transcriptome-wide the binding sites of cellular RBPs and miRNPs that we term PAR-CLIP (Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation) (see Fig. 1A for an outline of the method). The method relies on the incorporation of photoreactive ribonucleoside analogs, such as 4-thiouridine (4-SU) and 6-thioguanosine (6-SG) into nascent RNA transcripts by living cells. Irradiation of the cells by UV light of 365 nm induces efficient crosslinking of photoreactive nucleoside-labeled cellular RNAs to interacting RBPs. Immunoprecipitation of the RBP of interest is followed by isolation of the crosslinked and coimmunoprecipitated RNA. The isolated RNA is converted into a cDNA library and deep sequenced using Solexa technology. One characteristic feature of cDNA libraries prepared by PAR-CLIP is that the precise position of crosslinking can be identified by mutations residing in the sequenced cDNA. When using 4-SU, crosslinked sequences thymidine to cytidine transition, whereas using 6-SG results in guanosine to adenosine mutations. The presence of the mutations in crosslinked sequences makes it possible to separate them from the background of sequences derived from abundant cellular RNAs.

Application of the method to a number of diverse RNA binding proteins was reported in Hafner *et al.*<sup>18</sup>

## Video Link

The video component of this article can be found at <http://www.jove.com/video/2034/>

## Protocol

The protocol below describes the PAR-CLIP procedure for HEK293 cells expressing FLAG/HA-tagged IGF2BP1 upon induction with doxycycline. We will use an anti-FLAG antibody for immunoprecipitation.

PAR-CLIP will work with any cell line expressing detectable levels of the endogenous, untagged RNA binding protein (RBP) of interest if an efficient antibody for immunoprecipitation is available.

## Expanding Cells

1. Expand FlpIn-HEK293/TO/FLAG/HA-IGF2BP1 cells in growth medium. We recommend using between 100-400 x 10<sup>6</sup> cells (approx. 10-40 15 cm cell culture plates) as a starting point. Grow them to approximately 80% confluency.
2. 14 h before crosslinking add **a**) 4-thiouridine to a final concentration of 100 μM (1:1000 v/v of a 1 M 4-thiouridine stock solution) directly to the cell culture medium and **b**) induce expression of the FLAG/HA tagged IGF2BP1 by addition of 1 μg/ml of doxycycline (1:10,000 v/v of 10 mg/ml doxycycline stock solution). NOTE: instead of 4-thiouridine you can also use 100 μM of 6-thioguanosine.

## UV-Crosslinking

1. Wash cells once with 10 ml ice-cold PBS per plate and remove PBS completely.
2. Place plates on a tray with ice and irradiate uncovered with 0.15 J/cm<sup>2</sup> of 365 nm UV light in a Stratallinker 2400 (Stratagene) or similar device.
3. Scrape cells off with a rubber policeman in 1 ml PBS per plate, transfer to 50 ml centrifugation tubes and collect by centrifugation at 500 x g for 5 min at 4°C and discard the supernatant. 100 x 10<sup>6</sup> HEK293 cells (10 15 cm plates) will yield approx. 1 ml of wet cell pellet.
4. (optional) Unless you want to continue directly with cell lysis, shock freeze the cell pellet in liquid nitrogen and store at -80°C. Cell pellets can be stored for at least 12 months.

## Cell lysis and RNaseT1 digest

1. Take up cell pellet of crosslinked cells in 3 volumes of 1x NP40 lysis buffer and incubate on ice for 10 min.
2. Clear cell lysate by centrifugation at 13,000 x g for 15 min at 4°C.
3. Clear the lysate further by filtering it through a 0.2 μm membrane syringe filter (Pall Acrodisc or equivalent).
4. Add RNase T1 (Fermentas, 10,000 U/μl) to a final concentration of 1 U/μl and incubate in a water bath for 15 min at 22°C. Cool reaction subsequently for 5 min on ice before proceeding.

## Immunoprecipitation and recovery of crosslinked target RNA fragments

### Using the magnetic separator

Follow these guidelines throughout the sample preparation to prevent the magnetic beads from drying out.

1. Place the tube containing the beads on the magnetic stand for 1-2 minutes.
2. Add the buffer to the tube while the tube is on the magnetic separator.
3. Cap the tube, remove it from the magnetic separator, and resuspend the beads. You can resuspend the beads by flicking the tube with your finger or use a vortexer set at 5-6.
4. Centrifuge briefly to collect any beads that may remain in the tube cap.
5. Repeat steps 1 through 4 as required.

### Preparation of magnetic beads

1. Transfer 10 μl of Dynabeads Protein G magnetic particles (Invitrogen) per ml cell lysate (for a typical experiment it should be approx. 40-50 μl of beads) to a 1.5 ml microfuge tube. Wash beads twice with 1 ml of citrate-phosphate buffer.
2. Resuspend in twice the volume of citrate-phosphate buffer relative to the original volume of bead suspension.
3. Add 0.25 μg of anti-FLAG M2 monoclonal antibody (Sigma) per ml suspension and incubate on a rotating wheel for 40 min at room temperature.
4. Wash beads twice in 1 ml of citrate-phosphate buffer to remove unbound antibody.
5. Resuspend beads in twice the volume of citrate-phosphate buffer relative to the original volume of bead suspension.

### Immunoprecipitation (IP), second RNase T1 digestion, and dephosphorylation

1. Add 20 μl of freshly prepared antibody-conjugated magnetic beads per ml of partial RNase T1 treated cell lysate and incubate in 15 ml centrifugation tubes on a rotating wheel for 1 h at 4°C.
2. Collect magnetic beads on a magnetic particle collector for 15 and 50 ml centrifugation tubes (Invitrogen) and transfer to 1.5 ml microfuge tubes.
3. Wash beads 3 times in 1 ml of IP wash buffer.
4. Add RNaseT1 (Fermentas, 10,000 U/μl) to a final concentration of 100 U/μl and incubate the bead suspension in a water bath for 15 min at 22°C. Cool subsequently on ice for 5 min.
5. Wash beads 3 times in 1 ml of high-salt wash buffer.
6. Resuspend beads in 1 volume of dephosphorylation buffer
7. Add calf intestinal alkaline phosphatase (CIAP, NEB) to a final concentration of 0.5 U/μl, and incubate the suspension for 10 min at 37°C.
8. Wash beads twice in 1 ml of phosphatase wash buffer
9. Wash beads twice in polynucleotide kinase (PNK) buffer without DTT (the DTT concentration necessary for the enzymatic reaction is high enough to damage the magnetic beads).
10. Resuspend beads in one original bead volume of PNK buffer

### Radiolabeling of RNA segments crosslinked to immunoprecipitated proteins

1. To the bead suspension described above, add  $Y\text{-}^{32}\text{P}\text{-ATP}$  to a final concentration of  $0.5\ \mu\text{Ci}/\mu\text{l}$  and T4 PNK (NEB) to  $1\ \text{U}/\mu\text{l}$  in one original bead volume. Incubate the suspension for 30 min at  $37^\circ\text{C}$ .
2. Add non-radioactive ATP to obtain a final concentration of  $100\ \mu\text{M}$  and incubate for another 5 min at  $37^\circ\text{C}$ .
3. Wash the magnetic beads 5 times with  $800\ \mu\text{l}$  of PNK buffer without DTT.
4. Resuspend the beads in  $70\ \mu\text{l}$  of SDS-PAGE loading buffer.

### SDS-PAGE and electroelution of crosslinked RNA-protein complexes from gel slices

1. Incubate the radiolabeled suspension for 5 min in a heat block at  $95^\circ\text{C}$  to denature and release the immunoprecipitated RBP with crosslinked RNA and vortex.
2. Remove the magnetic beads on the separator and transfer the supernatant to a clean  $1.5\ \text{ml}$  microfuge tube.
3. Load  $40\ \mu\text{l}$  of the supernatant per well of a Novex Bis-Tris 4-12% (Invitrogen) precast polyacrylamide gel and run the gel for 55 min at 200 V.
4. Disassemble the gel chamber and dismantle the gel, leaving it mounted on one plate. To facilitate the alignment of the gel to the phosphorimager paper printout, we recommend implanting three tiny radioactive gel pieces asymmetrically at three of the four corners of the gel. Radioactive gel pieces can be collected from gels that were previously used to purify radiolabeled synthetic oligonucleotides. Wrap the gel in plastic film (e.g. Saran wrap) to avoid contamination.
5. Expose the gel to a blanked phosphorimager screen for 1 h and visualize on a phosphorimager.
6. Align the gel on top of the phosphorimager printout using the implanted gel pieces for orientation. Cut out the bands that correspond to the expected size of RBP (IGF2BP1, approx.  $75\ \text{kDa}$ ) and transfer to a D-Tube Dialyzer Midi Tube and add  $800\ \mu\text{l}$   $1\times$  SDS running buffer.
7. Electroelute the crosslinked RNA-RBP complex in  $1\times$  SDS running buffer at 100 V for 2 h. excised from the gel and electroeluted in a D-Tube Dialyzer Midi (Novagen) in  $800\ \mu\text{l}$  SDS running buffer according to the instructions of the manufacturer.

### Proteinase K digestion

1. Add an equal volume of  $2\times$  Proteinase K Buffer with respect to the electroeluate, followed by the addition of Proteinase K (Roche) to a final concentration of  $1.2\ \text{mg}/\text{ml}$ . Incubate for 30 min at  $55^\circ\text{C}$ .
2. Recover the RNA by acidic phenol/chloroform/IAA extraction (25:24:1, pH 4.0) followed by a chloroform extraction. Add  $1\ \mu\text{l}$  of glycogen ( $10\ \text{mg}/\text{ml}$  stock) precipitate the RNA by adding 3 volumes of ethanol. Dissolve the pellet in  $10.5\ \mu\text{l}$  water.

### cDNA library preparation and deep sequencing

Carry the recovered RNA through a standard cDNA library preparation protocol originally described for cloning of small regulatory RNAs<sup>19</sup>. The first step, 3' adapter ligation, was carried out as described on a  $20\ \mu\text{l}$  scale using  $10.5\ \mu\text{l}$  of the recovered RNA. Use the Solexa sequencing adapter sets described. Depending on the amount of RNA recovered, 5'-adapter-3'-adapter products without inserts may be detected after amplification of the cDNA as additional PCR bands. In such cases, excise the longer PCR product of expected size from a 3% NuSieve low-melting point agarose gel, elute the PCR product from the gel pieces using the GelElute kit (Qiagen) and sequence using the Solexa technology. One Solexa sequencing run usually affords between 6 and 10 million sequence reads that are enough for a transcriptome wide coverage of the binding sites of RNA binding proteins.

### Bioinformatic analysis

Careful bioinformatic analysis of the sequence reads needs to be done to obtain meaningful insights into the RNA binding sites for the examined RBP, such as the RNA recognition element, the preferred binding regions the RBP has (exonic vs. intronic, coding sequence vs. untranslated sequence). The sequence reads need to be aligned against the genome and EST databases. We usually use reads mapping uniquely to the genome with up to one mismatch, insertion or deletion to build clusters of sequence reads that can then be further analyzed. The frequency of characteristic mutations in the clustered sequenced reads, T to C transitions when using 4-SU and G to A transitions when using 6-SG, are indicative of successfully crosslinked sequences. In our experience uncrosslinked RNAs labeled with 4-SU show a background mutation rate of approximately 20%. This rate increases to approx. 50-80% upon crosslinking.

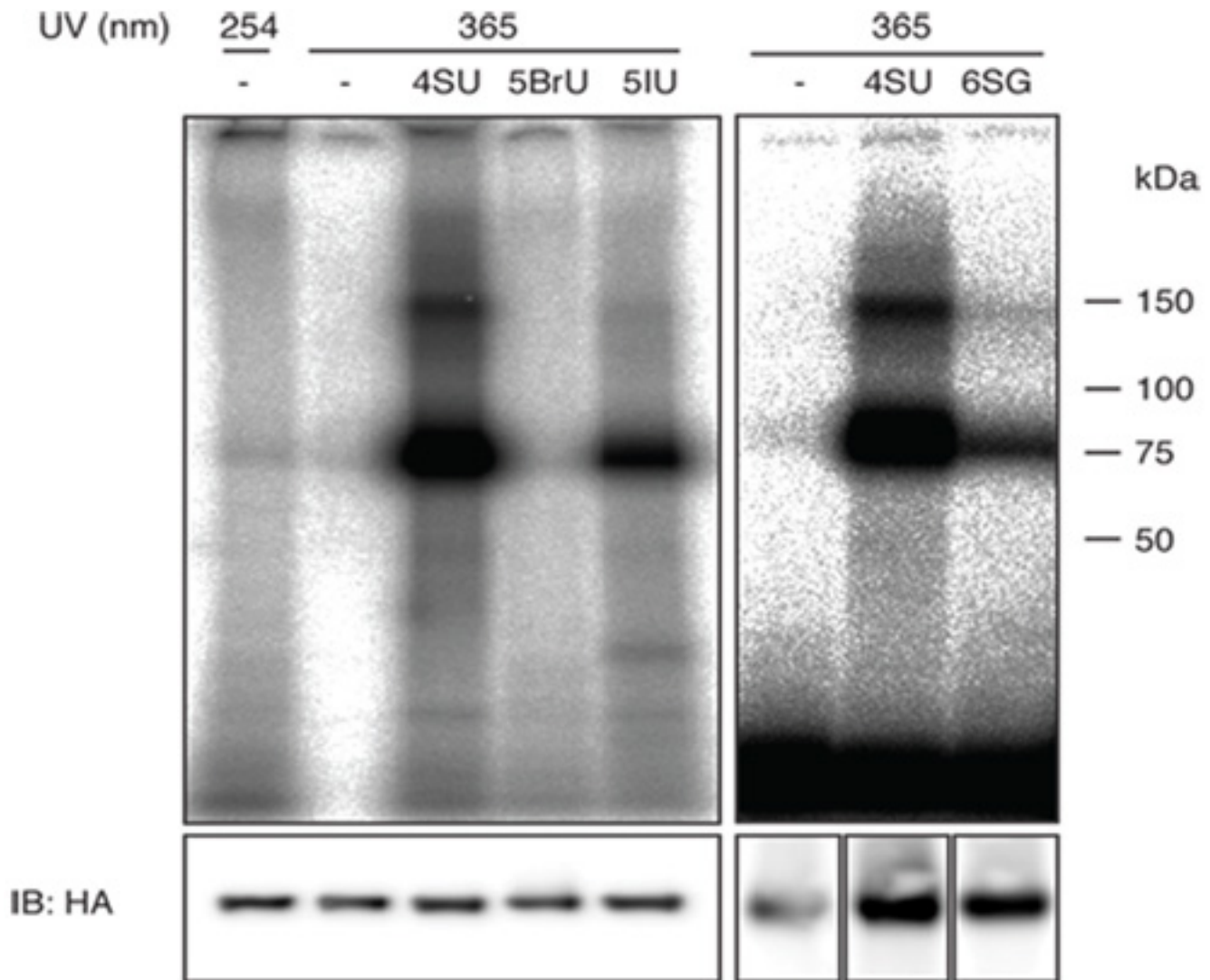
A detailed description of the bioinformatic analysis can be found in the Supplementary material of the publication by Hafner *et al.*<sup>18</sup>

### Optional Steps

#### Determination of incorporation levels of 4-thiouridine into total RNA

Isolate total RNA from the cell line stably expressing the RBP of interest after growing in medium supplemented with  $100\ \mu\text{M}$  4SU 16 h prior to harvest. As a control, harvest cells grown without 4SU addition. Isolate total RNA by addition of 3 volumes of Trizol reagent (Sigma) to the washed cell pellets following the manufacturer's instructions. Further purify total RNA using Qiagen RNeasy according to the manufacturer's protocol. To prevent oxidation of 4SU during RNA isolation and analysis, add  $0.1\ \text{mM}$  dithiothreitol (DTT) to the wash buffers and subsequent enzymatic steps. Digest and dephosphorylated total RNA to single nucleosides for HPLC analysis as described before<sup>20</sup>. Briefly, in a  $30\ \mu\text{l}$  volume, incubate  $40\ \mu\text{g}$  of purified total RNA were for 16 h at  $37^\circ\text{C}$  with  $0.4\ \text{U}$  bacterial alkaline phosphatase (Worthington Biochemical) and  $0.09\ \text{U}$  snake venom phosphodiesterase (Worthington Biochemical). As a reference standard, use a synthetic 4SU-labeled RNA, (we standardly use CGUACGCGGAAUACUUCGA(4SU)U) and also subject it to complete enzymatic digestion. Separate the resulting mixtures of ribonucleosides by HPLC on a Supelco Discovery C18 (bonded phase silica  $5\ \mu\text{M}$  particle,  $250 \times 4.6\ \text{mm}$ ) reverse phase column (Bellefonte PA, USA). HPLC buffers are  $0.1\ \text{M}$  TEAA in 3% acetonitrile (A) and 90% acetonitrile in water (B). Use an isocratic gradient: 0% B for 15 min, 0 to 10 % B for 20 min, 10 to 100% B for 30 min. Apply a 5 min 100 % B wash applied between runs to clean the HPLC column.

## Representative Results



**Figure 1** (right panel) shows a representative result of a PAR-CLIP performed with cell lines expressing FLAG/HA-tagged IGF2BP1 with 4-SU and 6-SG. Note that the crosslinking efficiency of 6-SG for IGF2BP1 is lower than the crosslinking efficiency for 4-SU. The lower crosslinking efficiency will result in a higher background of sequences derived from fragments of abundant cellular RNAs and therefore you should consider scaling up the experiment when using less efficient photoreactive nucleosides.

The left panel of Figure 1 shows a comparison of using different photoreactive uridine analogs that could be potentially used for PAR-CLIP compared to traditional UV 254 nm crosslinking.

The intensity of the radioactive band of the correct length gives you a good idea whether the PAR-CLIP experiment has worked and you have isolated sufficient RNA to carry through a small RNA sequencing protocol (step-by-step description for cDNA library preparation of small RNAs sequencing can be found in <sup>19</sup>). The frequency of characteristic mutations in the sequenced reads, T to C transitions when using 4-SU and G to A transitions when using 6-SG, are indicative of successfully crosslinked sequences. In our experience uncrosslinked RNAs labeled with 4-SU show a background mutation rate of approximately 20%. This rate increases to approx. 50-80% upon crosslinking.

## Disclosures

T.T. is a cofounder and scientific advisor to Alnylam Pharmaceuticals and an advisor to Regulus Therapeutics.

## Acknowledgements

We thank members of the Tuschl laboratory for helpful discussions. M.H. is supported by the Deutscher Akademischer Austauschdienst (DAAD). This work was supported by the Swiss National Fund Grant #3100A0-114001 to M.Z.; T.T. is an HHMI investigator, and work in his laboratory was supported by NIH grants GM073047 and MH08442 and the Starr Foundation.

## References

1. Keene, J. D., RNA regulons: coordination of post-transcriptional events. *Nat. Rev. Genet* 8 (7), 533 (2007).
2. Tenenbaum, S. A. *et al.*, Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. *Proc. Nat. Acad. Sci.* 97 (26), 14085 (2000).
3. Gerber, A. P. *et al.*, Genome-wide identification of mRNAs associated with the translational regulator PUMILIO in *Drosophila melanogaster*. *Proc. Nat. Acad. Sci.* 103 (12), 4487 (2006).
4. Lopez de Silanes, I. *et al.*, Identification of a target RNA motif for RNA-binding protein HuR. *Proc. Nat. Acad. Sci.* 101 (9), 2987 (2004).
5. Greenberg, J. R., Ultraviolet light-induced crosslinking of mRNA to proteins. *Nucl. Acids Res.* 6 (2), 715 (1979).
6. Wagenmakers, A. J. *et al.*, Cross-linking of mRNA to proteins by irradiation of intact cells with ultraviolet light. *Eur. J. Biochem.* 112 (2), 323 (1980).
7. Mayrand, S. *et al.*, Structure of nuclear ribonucleoprotein: identification of proteins in contact with poly(A)<sup>+</sup> heterogeneous nuclear RNA in living HeLa cells. *The Journal of Cell Biology* 90 (2), 380 (1981).
8. Dreyfuss, G. *et al.*, Characterization of heterogeneous nuclear RNA-protein complexes *in vivo* with monoclonal antibodies. *Mol. Cell. Biol.* 4 (6), 1104 (1984).
9. Adam, S. A. and Dreyfuss, G., Adenovirus proteins associated with mRNA and hnRNA in infected HeLa cells. *J. Virol.* 61 (10), 3276 (1987).
10. Ule, J. *et al.*, CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302 (5648), 1212 (2003).
11. Licatalosi, D. D. *et al.*, HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 456 (7221), 464 (2008).
12. Yeo, G. W. *et al.*, An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. *Nat. Struct. Mol. Biol.* 16 (2), 130 (2009).
13. Sanford, J. R. *et al.*, Splicing factor SFRS1 recognizes a functionally diverse landscape of RNA transcripts. *Genome Res.* 19 (3), 381 (2009).
14. Granneman, S. *et al.*, Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc. Nat. Acad. Sci.* (2009).
15. Guil, S. and Caceres, J. F., The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat. Struct. Mol. Biol.* 14 (7), 591 (2007).
16. Chi, S. W. *et al.*, Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 460 (7254), 479 (2009).
17. Zisoulis, D. G. *et al.*, Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nat. Struct. Mol. Biol.* advance online publication (2010).
18. Hafner, M. *et al.*, Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* manuscript accepted for publication (2010).
19. Hafner, M. *et al.*, Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* 44 (1), 3 (2008).
20. Andrus, A. and Kuimelis, R. G., Base composition analysis of nucleosides using HPLC. *Current Protocols in Nucleic Acid Chemistry* Chapter 10, Unit 10 6 (2001).