Illustrating the epitranscriptome at nucleotide resolution using methylation-iCLIP (miCLIP)

Harry George[1], Jernej Ule[2] and Shobbir Hussain[1*]

[1]Department of Biology and Biochemistry, University of Bath, Bath, BA2 7AY, UK; [2]Department of Molecular Neuroscience, University College London, London, WC1N 3BG, UK

Keywords: *epitranscriptome, epitranscriptomics, RNA methylation, methylation-iCLIP, miCLIP, NSun2*

***Correspondence: Shobbir Hussain, s.hussain@bath.ac.uk, +44 (0)1225 386293**

Running head: Methylation-iCLIP protocol

Summary

Next generation sequencing technologies have enabled the transcriptome to be profiled at a previously unprecedented speed and depth. This yielded insights into fundamental transcriptomic processes such as gene transcription, RNA processing, and mRNA splicing. Immunoprecipitation-based transcriptomic methods such as individual nucleotide resolution crosslinking immunoprecipitation (iCLIP) have also allowed high-resolution analysis of the RNA interactions of a protein of interest, thus revealing new regulatory mechanisms. We and others have recently modified this method in order to profile RNA methylation, and we refer to this customised technique as methylation-iCLIP (miCLIP). Variants of miCLIP have been used to map the methyl-5-cytosine (m5C) or methyl-6-adenosine (m6A) modification at nucleotide resolution in the human transcriptome. Here we describe the m5C-miCLIP protocol and discuss how it yields the nucleotide-resolution RNA modification maps. Finally, we describe insights into the influence of RNA sequence and structure on the addition of covalent RNA modifications that were gained from miCLIP studies, and how these have contributed to the new field of molecular genetics research coined ‘epitranscriptomics’.

1 Introduction

1.1 RNA Methylation and the development of transcriptomic techniques to detect their occurrence at nucleotide resolution

Covalent modifications to cellular RNA molecules can occur both on the sugar-phosphate RNA backbone as well as on nucleobases and can include those that occur at internal positions as well as those present on 5' or 3' termini. A few of these such as the m7G cap that is ubiquitously present on the 5' terminus of cellular mRNAs have been well characterised. However it is known that well over a hundred types of modified RNA nucleobases can exist, though the characterisation of these have been hampered largely due to a lack of suitable techniques to detect their sequence-specific occurrence. For example, it had for decades been suspected that internal methyl-6-adenosine (m6A) and methyl-5-cytosine (m5C) base methylations may be a common occurrence on mRNAs in mammalian cells ***(1)***, but it was only much more recently that the development of refined transcriptomic approaches allowed these modifications to be comprehensively mapped ***(2,3,4,5,6)*** thus greatly aiding efforts to investigate their mechanistic roles ***(6,7,8,9,10)***.

Previous important studies were also critical for setting the stage for the recent flurry of activity in the field. It had been shown that a multiprotein complex, which includes the METTL3 and METTL14 subunits, catalyses the formation of m6A in RNA ***(11)*** and that an additional enzyme encoded by the FTO gene can remove these methylation marks ***(12)***. The latter finding was particularly exciting as it suggested that these RNA methylation events could be dynamically controlled during the execution of important biological processes. Two seminal studies then appeared simultaneously in the literature describing the use of essentially the same immunoprecipitation-based method to map m6A modifications in the transcriptome ***(2,3)***. The approach consisted of fragmenting cellular RNA into small pieces and then using an antibody specific for the methylated base to immunoprecipitate m6A-containing fragments. Libraries were then prepared from the purified RNAs for high-throughput sequencing on a Next-Generation-Sequencing (NGS) platform. Although the m6A sites could in reality be located anywhere within the sequenced reads, computational algorithms were used to predict the likely methylation site. Importantly it was also found by both studies that most of the RNA fragments mapped to within close vicinity of the stop

codon of mRNAs, although the mechanistic significance of this finding remains to be determined.

An additional study published in 2012 also reported transcriptomic identification of m5C using bisulfite sequencing ***(4)***, which had been used previously to detect DNA m5C modifications. The authors optimised the protocol so that RNA could seemingly withstand the harsh chemical conditions of the bisulfite treatment. The analysis revealed widespread occurrence of the modification in several RNA biotypes, including mRNAs as previously predicted by Dubin and Taylor (1975) ***(1)***. The integrity of the RNA libraries following bisulfite treatment however could not be fully assessed and the method also does not enrich for modified RNA sites, as in the m6A-immunoproecptitaion protocol, meaning that methylation-site detection and calling was likely sub-optimal. Independent approaches were also clearly necessary and were provided by two studies again appearing fortuitously in quick succession in the literature which described distinct but related cross-linking immunoprecipitation (CLIP)-based techniques for detecting m5C RNA modifications ***(5,6)***.

Initially described for the use of detecting protein-RNA interactions, efficient CLIP-based techniques employed UV covalent crosslinking to enable the RNA interaction partners of a protein of interest to be identified with high stringency and at a remarkably high signal to noise ratio ***(13,14)***. The methylation of m5C RNA is catalysed via a covalent catalytic intermediate between the RNA methylase enzyme and RNA substrate: if these catalytic intermediates could be frozen or 'trapped', they could be immunoprecipitated to detect the RNA methylation sites. Khoddami and Cairns (2013) described their 5-azacytidine-mediated RNA immunoprecipitation (Aza-IP) approach where they employed the cytidine nucleoside analogue 5-azacytidine ***(5)***, which becomes incorporated into DNA and RNA during replication and transcription respectively. m5C DNA and RNA methylases will methylate cytosine residues as well as 5-azacytosine residues via a covalent catalytic intermediate, but when they methylate the latter, the resolution of the covalent adduct cannot take place due to the altered chemistry of the nucleobases. Thus cells treated with 5-Azacytidine enabled CLIP-based experiments to be performed where an antibody against an epitope-tagged m5C RNA methylase was used to immunoprecipitate the frozen catalytic intermediates, enabling identification of the methylation substrates. In CLIP experiments, during the cDNA synthesis step of the library preparation, mutations in the sequence are incorporated as the reverse

transcriptase traverses the crosslink site; these crosslink-induced mutation sites (CIMS) are estimated to occur at a frequency of ~8-20% of reverse transcription events, and their presence can be used to pinpoint the crosslink/interaction site ***(15)***. Similarly, in the Aza-IP libraries, it was observed that the sequence reads often also contained CIMS at the crosslink/methylation site and hence these could be used to infer the methylation site at nucleotide resolution ***(5)***.

Aza-IP involves altering the physiology of cells by requiring a 12 hour pre-incubation with the cytotoxic agent 5-azacytidine, which is known to lead to hypomethylation of DNA even at low doses, and its effect of the nucleoside analogue following incorporation into different RNA species are not fully characterised but may interfere with RNA translation and metabolism. Further, more than 80% of reverse transcription events are estimated to stall at the protein-RNA crosslink site during cDNA synthesis, rather than reading through the crosslink site ***(16)***; in CLIP-based techniques, these truncated cDNAs are lost during NGS library preparation as they lack the terminal adapter sequence required for PCR amplification. This is overcome by the individual-nucleotide resolution crosslinking immunoprecipitation (iCLIP) method, which was developed to enable amplification of such truncated cDNAs ***(17)***. In iCLIP, the last nucleotide of the truncated cDNAs can be used to identify the crosslink site with increased efficiency ***(16,17)***. The methylation-iCLIP (miCLIP) technique developed in Hussain et al. 2013a describes a customised version of iCLIP ***(6)*** that takes advantage of a mutation within the key cysteine residue of m5C RNA methylases that is required for resolution of catalytic intermediates ***(18,19,20)*** (Figure 1). Substitution of this catalytic cysteine to alanine led to effective trapping of catalytic intermediates, which enabled to crosslink the methylases to its endogenous RNA targets without the need of photocrosslinking. These crosslinked epitope tagged enzyme-substrate covalent adducts are then immunoprecipitated, the cDNA library is prepared using the iCLIP protocol, and the positions of cDNA truncations identified the m5C methylation sites with nucleotide-resolution.

The miCLIP study demonstrated that identified m5C substrates were likely of biological relevance. For example, miCLIP for the NSun2 m5C RNA methylase detected methylation sites at various nucleotide positions within the Vault ncRNAs; these were then confirmed by performing bisulfite sequencing of Vault ncRNA extracted from cells obtained from human

patients with NSun2-deficiency syndromes ***(6)***. The study also showed that the methylation events mechanistically impacted on the processing of Vault hairpin RNAs into small RNAs with microRNA-like function. Importantly there are six other members of the NSun family of m5C RNA methylases which should also be amenable to miCLIP (Figure 1).

Recently, a further iCLIP-based technique was developed in order to identify m6A sites at nucleotide resolution in transcriptomes ***(21)***. The technique involves incubation of a m6A-antibody with purified cellular RNA followed by *in vitro* UV-crosslinking. These crosslinked antibody-RNA complexes are then immunoprecipitated, the cDNA library is prepared using the iCLIP protocol, and the positions of cDNA truncations identified the m6A methylation sites with nucleotide-resolution. The procedure was named m6A-iCLIP and abbreviated to miCLIP. To allow a broader use of this term, we suggest that miCLIP is described as methylation-iCLIP, which can then be applied both for cases where protein-RNA complexes are crosslinked within intact cells ***(6)*** or purified RNA ***(21)*** in order to identify sites of RNA methylation.

1.2 An overview of m5C-iCLIP

The aforementioned work ***(2,3,4,5,6,21)*** as well as important studies that have also described novel transcriptomic techniques to detect other types of RNA modifications ***(22,23,21)*** have set the stage for 'epitranscriptomics'. This area of research concerns the sequence-context specific characterisation and investigation of RNA modifications in complex transcriptomes ***(24,25,26)***. Methylation-iCLIP will continue to play an important role in this fledgling field, and in this chapter we describe the m5C-iCLIP protocol. In classic iCLIP experiments, UV crosslinking is required to introduce covalent crosslinks between proteins and interacting RNAs; this step is omitted from the miCLIP protocol. Instead expression of a point-mutant form of an RNA methylase of interest is required which will lead to RNA substrate trapping; we have previously and are currently performing this using point mutant forms of the NSun family of RNA m5C methylases. Experiments can be performed by transfecting suitable expression vectors into cells which also offers the advantage that an antibody against an epitope tag can be used for immunoprecipitation. Genome editing methods however, may also potentially be used in order to achieve

endogenous expression of the point mutant. Following expression, cells are lysed and the enzyme-substrate covalent adducts are recovered by immunoprecipitation; the use of a good, highly specific antibody is key to the experiment and we generally recommend using an antibody against an epitope tag. A multi-step library preparation procedure is then employed which is designed to enable methylation sites to be determined at nucleotide resolution following sequencing of prepared libraries on an NGS platform. Bioinformatics analysis of recovered data can be performed by using the online PIPE-CLIP tool ***(27)***. As in iCLIP, the primers used for cDNA synthesis during library prep contain a randomised barcode (often referred to as unique molecular identifiers or UMIs), which allow individual cDNAs to be counted during analyses, thus yielding a reliable quantitative estimate of crosslinking efficiencies at each nucleotide.

2 Materials

2.1 Cell culture and transfection

1. HEK293 cells.

2. Growth medium: Dulbecco's Modified Eagles Medium (DMEM), 10% Fetal Bovine Serum, Penicillin-Streptomycin.

3. Trypsin-EDTA: 1x solution

4. Phosphate Buffered Saline (PBS): 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO4, 1.8 mM KH_2PO_4

5. Lipofectamine 2000

2.2 Immunoprecipitation and linker ligation

1. Protein G or A magnetic Dynabeads

2. RNase I: 100 U/µL

3. High RNase: 1:50 dilution of RNaseI

4. Low RNase: 1:1000 dilution of RNaseI

5. Turbo DNase: 2 U/µL

6. T4 PNK (with 3' phosphatase activity) plus 10x PNK buffer

7. RNasin: 20-40 U/µL

8. T4 RNA Ligase I: 30 U/µL

9. PEG400

2.3 SDS PAGE analysis and nitrocellulose transfer

1. ATP [γ-^{32}P], 10mM; Activity: 3000Ci/mmol

2. NuPAGE loading buffer

3. 4-12% NuPAGE Bis-Tris gel

4. NuPAGE MOPS SDS running buffer

5. Pre-stained protein marker

6. Nitrocellulose membrane

7. Autoradiography cassette

8. X-ray film

2.4 Library preparation

1. Proteinase K: 20 mg/mL

2. Phenol-chloroform (24:23 v/v)

3. Phase-lock gel heavy tubes

5. Glycoblue: 15mg/ml

6. Sodium acetate solution: 3 M in water, pH 5.5

7. Superscript III reverse transcriptase and 5 x RT buffer (Life Technologies)

8. dNTP mix: 10 mM dATP, 10 mM dTTP, 10 mM dCTP, 10 mM dGTP

9. DTT solution: 0.1 M in water

10. RNase I: 100 U/µL

11. TBE-Urea loading buffer, 2x (Thermofisher Scientific)

12. 6% TBE-Urea polyacrylamide gel

13. Costar SpinX columns

14. Glass pre-filters

15. CircLigase II enzyme and 10 x CircLigase II buffer (Epicentre)

16. FastDigest BamHI enzyme and 10 x FastDigest Buffer

17. P5/P3 solexa primer mix: 10 µM each primer

18. Accuprime Supermix (Life Technologies)

19. 6% TBE polyacrylamide gel

20. SYBR green I, 10,000x (Thermofisher Scientific)

2.5 Buffers

1. Lysis buffer: 50 mM Tris–HCl, pH 7.4, 100 mM NaCl, 1 % Igepal CA-630, 0.1 % SDS.

2. High-salt wash buffer: 50 mM Tris–HCl, pH 7.4, 1 M NaCl, 1 mM EDTA, 1 % Igepal CA-630, 0.1 % SDS.

3. PNK buffer: 20 mM Tris–HCl, pH 7.4, 10 mM $MgCl_2$, 0.2 % Tween-20.

4. 5× PNK pH 6.5 buffer: 350 mM Tris–HCl, pH 6.5, 50 mM $MgCl_2$, 25 mM dithiothreitol (DTT).

5. 4× ligation buffer: 200 mM Tris–HCl, pH 7.8, 40 mM $MgCl_2$, 40 mM dithiothreitol (DTT).

6. PK buffer: 100 mM Tris–HCl, pH 7.4, 50 mM NaCl, 10 mM EDTA.

7. PK buffer-7M Urea buffer: 100 mM Tris–HCl, pH 7.4, 50 mM NaCl, 10 mM EDTA, 7 M urea

8. TE Buffer: 10 mM Tris, 1mM EDTA, pH8

2.6 Oligonucleotides

1. Pre-adenylated linker (L3-App): rApp/AGATCGGAAGAGCGGTTCAG/ddC, 20μM; rApp = adenylation, ddC = dideoxycytidine (Note 1).

2. RT primers (Note 2): 10 μM each

Rt1clip/5Phos/NNAACCNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt2clip/5Phos/NNACAANNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt3clip/5Phos/NNATTGNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt4clip/5Phos/NNAGGTNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt6clip/5Phos/NNCCGGNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt7clip/5Phos/NNCTAANNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt8clip/5Phos/NNCATTNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt9clip/5Phos/NNGCCANNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt11clip/5Phos/NNGGTTNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt12clip/5Phos/NNGTGGNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt13clip/5Phos/NNTCCGNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt14clip/5Phos/NNTGCCNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt15clip/5Phos/NNTATTNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Rt16clip/5Phos/NNTTAANNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

N = random nucleotides.

3. BamHI-cut-oligo (GTTCAGGATCCACGACGCTCTTCaaaa): 10 μM

4. Illumina PCR primers: 10 μM each

P5Solexa: AATGATACGGCGACCACCGAGATCTACACT CTTTCCCTACACGACGCTCTTCCGATCT

P3Solexa:
CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT

3. Methods

3.1 Cell culture and transfection

1. In a 10 cm dish, grow HEK293 cells in growth medium in a humidified incubator with 5% CO_2 until they reach 50-70% confluence.

2. Use Lipofectamine 2000 (Life Technologies) to transfect cells with mammalian expression vector encoding the RNA methylase of interest harbouring the substrate trapping point mutation (Note 3) and incubate cells for 24 hours.

3. Following incubation, wash cells twice in PBS. Disperse cells by incubating cells for 1 minute in Trypsin-EDTA and then add 30 mL of growth medium.

4. Collect cells by centrifugation and then wash twice more in PBS via centrifugation and resuspension of cell pellet.

5. Transfer cells in 1 mL PBS to microcentrifuge tubes, and subject to centrifugation to pellet cells for a final time. Cell pellets may be used immediately or frozen at -80^{o}C for later use.

3.2 Immunoprecipitation

All subsequent steps in the protocol should be performed on wet ice unless otherwise stated.

3.2.1 Dynabead preparation

1. Transfer 100 μL of resuspended magnetic dynabeads (Note 4) to two 1.5 mL microtubes and label one ‘High RNase and the other ‘Low RNase’. Also repeat this for a no-antibody control sample.

2. Wash the beads twice with 1 mL of lysis buffer.

3. After the last wash, resuspend the beads in 500 µL of lysis buffer and then add 5-10 µg of antibody.

4. Rotate tubes for 30 min at room temperature.

5. Wash beads 5 times with 1 mL lysis buffer and leave in last wash until ready to proceed to step 3.2.3.

3.2.2 Lysis and partial RNA digestion

1. Add 1mL of cold lysis buffer to cell pellet, and then pipette to resuspend cell pellet. Incubate on ice for 30 mins and syringe lysate through a 19 gauge needle every 10 mins during the incubation.

2. Add 10 µL of high or low RNase (Note 5) to the corresponding tube and 2 µL Turbo DNase to the cell lysate and then place the samples into a thermomixer set at 37 °C for 3 min, shaking at 1100 rpm. Following the 3 min incubation immediately transfer the microtube onto wet ice and leave for at least 3 min.

3. Centrifuge for 20 min at 20,000x*g* at 4 °C and then transfer the supernatant to a 1.5mL microtube. The remaining pellet can be discarded.

3.2.3. Immunoprecipitation

1. Add the clarified lysate to the beads from step 3.2.1 and then rotate for 2-3 hours at 4^{o}C.

2. Use a magnetic rack to collect the dynabeads, discard the supernatant, and then wash the beads twice with 1 mL high salt wash buffer.

3. Wash 3 more times with 1 mL PNK buffer for each wash and discard the supernatant from the final wash.

3.3 RNA 3’ dephosphorylation

1. Resuspend beads to homogeneity in 20 µL 3’ dephosphorylation mix containing 4 µL 5× PNK pH 6.5 buffer, 0.5 µL PNK (with 3’ phosphatase activity), 0.5 µL RNasin, 15 µL water.

2. Incubate for 20 min at 37 °C in a thermomixer at 1100 rpm.

3. Remove from thermomixer and add 1 mL of PNK buffer to wash beads. Perform 4 further washes in 1 mL PNK buffer.

3.4 Preadenylated L3 adapter ligation

1. Remove the supernatant and resuspend the beads in 20 µL adapter ligation mix containing 8 µL water, 5 µL 4× ligation buffer, 1 µL T4 RNA ligase, 0.5 µL RNasin, 1.5 µL pre-adenylated linker L3-App, 4 µL PEG400.

2. Incubate overnight at 16 °C in a thermomixer at 1100 rpm.

3. Remove from thermomixer and add 1 mL of PNK buffer to wash beads. Perform 4 further washes in 1 mL PNK buffer, and then leave in 1 mL of PNK buffer.

3.5 RNA 5' end labelling

1. Transfer 200 µL of the resuspended beads (the remaining 800 µL will be used in step 4) to a new RNAse free microtube and remove the supernatant.

2. Resuspend beads in 10 µL of radioactive PNK mix containing 0.5 µL PNK, 1 µL 32P-γ-ATP, 1 µL 10× PNK buffer, 7.5 µL water.

3. Incubate for 5 min at 37^{o}C and then discard the supernatant as radioactive waste.

4. Resuspend beads in 20 µL 1 x NuPAGE loading buffer and add this to the remaining cold beads from section 3.4, step 3, following removal of the supernatant from them.

6. Incubate for 5 min at 80^{o}C.

7. Place back on the magnet and collect the supernatant which will be loaded directly onto the gel in the next step.

3.6 SDS–PAGE and nitrocellulose transfer

1. Load the samples on a 4–12% NuPAGE Bis-Tris gel. Also load 5 µL of a pre-stained protein size marker. Use 0.5 L 1x NuPAGE MOPS SDS running buffer.

2. Run the gel at 180 V for 50 min and once complete, cut off the foot of the gel containing the dye front and discard it as solid radioactive waste.

3. Transfer the protein–RNA complexes from the gel to a nitrocellulose membrane at 30V for 1 hour.

4. Once the transfer is complete, wrap the nitrocellulose membrane in saran wrap and fix it into an autoradiography cassette which contains a fluorescent marker.

5. Expose the membrane to an X-ray film and then transfer the cassette into a −80 °C freezer. Exposures can be performed for 30 min, 1 h and overnight.

3.7 Library preparation

3.7.1 RNA isolation

1. Use the fixed fluorescent marker in the autoradiography cassette to re-align the processed film and then mark the area of the nitrocellulose membrane that contains the protein-RNA complex smear in the low RNase condition (Note 6). Using a sharp blade, carefully cut out this region of nitrocellulose, dissect further into small pieces, and then transfer them to a RNase free 1.5 mL microtube.
2. Add 10 µL proteinase K in 200 µL PK buffer to the nitrocellulose fragments so that they are all submerged. Transfer tubes to a thermomixer and incubate for 20 min shaking at 1100 rpm at 37 °C.

3. Add 200 µL of PK buffer-7 M urea and incubate for further 20 min at 37 °C and 1100 rpm.

4. Collect the solution and add it to a 2 mL Phase Lock Gel Heavy tube. Add 400 µL phenol/chloroform and then incubate in a thermomixer for 5 min at 30 °C shaking at 1100 rpm.

5. Separate the phases by spinning for 5 min at 20,000xg and room temperature.

6. Transfer the aqueous layer into a non-stick RNase free 1.5 mL microfuge tube and add 1 µL glycoblue and 40 µL sodium acetate solution. Mix and then add 1 mL 100% ethanol. Mix again and then place at −20 °C overnight.

7. Centrifuge at 20,000g at 4 °C for 20 min and then remove the supernatant. Wash the pellet with 1 mL 80% ethanol and spin again for 5 min. Resuspend the pellet thoroughly in 5 µL H_2O and transfer to an RNase free PCR tube.

3.7.2 Reverse transcription

1. Add 1 µL of RT primer (Note 7) and 1 µL dNTP mix to the resuspended pellet and then incubate in a thermocycler set to the following programme: 70 °C for 5 min, then hold at 25 °C until the RT mix is added next.

2. Add 13 µL of the RT mix containing 7 µL H_2O, 4 µL 5× RT buffer, 1 µL DTT, 0.5 µL RNasin, 0.5 µL Superscript III reverse transcriptase, transfer tube to a thermal cycler and perform the reverse transcription with the following program: 25 °C for 5 min, 42 °C for 20 min, 50 °C for 40 min, 80 °C for 5 min, then hold at 4 °C.

3. Once the RT program is complete, add 350 µL TE buffer, 1 µL GlycoBlue and 40 µL sodium acetate pH 5.5. Mix, and then add 1 mL 100 % ethanol. Mix again and precipitate overnight at −20 °C.

3.7.3 Gel purification (Note 8)

1. Centrifuge for 20 min at 20,000xg at 4 °C. Remove the supernatant and add 1 mL 80% ethanol. Repeat centrifugation, remove supernatant and then resuspend the pellet in 6 µL H_2O.

2. Add 6 µL 2 × TBE-urea loading buffer to the samples and also to 6 µL DNA size marker. Heat samples at 70 °C for 2 min immediately before loading onto a 6% TBE-urea gel (Life Technologies).

3. Run gel in 1X TBE buffer for 40 min at 180 V until the lower dye is close to the bottom.

4. Cut off the last lane containing the size marker and incubate it in 20 mL 1X TBE buffer with 2 µL SYBR green II for 10 min. Wash the gel twice with 1XTBE and visualize by UV transillumination. Print the result with 100% scale, and use it as a mask to guide the excision of cDNA bands from the rest of the gel.

5. Together with the preadenylated L3 adapter sequence, the primer sequence accounts for 52 nt of the cDNA. Cut three bands, at 70–80 nt, 80–150 nt and 150-200 nt and transfer them each to a non-stick RNase free 1.5 mL microtube (Note 9).

6. Add 400 µL TE buffer to the gel piece and crush it into tiny pieces with a 1 mL pipette tip.

7. Transfer to a thermomixer and incubate with shaking at 1100 rpm for 1h at 37 °C. Then place on dry ice for 5 min, and then place back for 1 h at 1100 rpm at 37 °C in the thermomixer.

8. Transfer the liquid portion of the supernatant into a Costar SpinX column into which you have placed two 1 cm glass pre-filters.

9. Centrifuge at 20,000xg for 1 min at room temperature and transfer the flow-through to a non-stick RNase free 1.5mL microtube.

10. Add 1 µL glycoblue and 40 µL sodium acetate solution and mix. Add 1 mL 100% ethanol, mix again and precipitate at −20 °C overnight.

3.7.4 cDNA circularisation (Note 10)

1. Centrifuge for 20 min at 20,000xg at 4 °C. Remove the supernatant and add 1 mL 80% ethanol. Repeat centrifugation, remove supernatant and then resuspend the pellet in 8 µL ligation mix containing 6.5 µL water, 0.8 µL 10× CircLigase Buffer II, 0.4 µL 50 mM $MnCl_2$, 0.3 µL CircLigase II. Transfer the mix to nuclease-free PCR tubes and incubate for 1 h at 60 °C.

2. Add 30 µL oligonucleotide annealing mix: 26 µL H_2O, 3 µL FastDigest Buffer, 1 µL BamHI Cut_oligo and anneal in a thermal cycler using the following program: 95 °C for 2 min; then successive cycles of 20 s, starting from 95 °C and decreasing the temperature by 1 °C each cycle down to 25 °C; then hold at 25 °C.

3. Add 2 µL FastDigest BamHI enzyme and incubate for 30 min at 37 °C, followed by an incubation at 80 °C for 5 min.

4. Add 350 µL TE, 1 µL GlycoBlue, 40 µL sodium acetate solution and mix. Then add 1 mL 100 % ethanol, mix again and precipitate overnight at −20 °C.

3.8 PCR amplification and library submission

1. Centrifuge for 20 min at 20,000xg at 4 °C. Remove the supernatant and add 1 mL 80% ethanol. Repeat centrifugation, remove supernatant and then resuspend the pellet in 20 µL

H_20 and transfer to PCR tube. 10 µL of this will be used in the PCR in the next step and the remainder can be stored at −20 °C for further future use if necessary.

2. Prepare the PCR mix as follows: 10 µL cDNA, 1 µL primer mix of P5Solexa and P3Solexa, 20 µL Accuprime Supermix, 9 µL H_2O.

3. Place tube in thermal cycler and run the following program: 94 °C for 2 min, followed by 25 cycles of: 94 °C for 15 s, 65 °C for 30 s, 68 °C for 30 s; complete with 68 °C for 3 min, then hold at 25 °C.

4. Once the program is complete, mix 8 µL PCR product with 2 µL 5x TBE loading buffer, load on a 6% TBE polyacrylamide gel and run at 180 V and 120 mA for 30 min. Stain and visualize with SYBR green I (Note 11).

5. PCR amplified miCLIP libraries can then be quantified using a Qubit fluorometer (Life Technologies).

6. Dilute libraries to 10 nM, or to a sequencing facility-specified concentration, and sequence using 50nt single or paired-end runs on an Illumina sequencer. We have routinely performed sequencing on the GAII and also the MiSeq platform to generate sufficient RNAseq reads.

7. Bioinformatics analysis may be performed using the online PIPE-CLIP tool ***(27)***.

Notes

1. Pre-adenylation (rApp) of the DNA linker is required for efficient ligation to the 3'-OH group of RNAs. The dideoxycytidine group (ddC) is added to the 3' end of the adenylated linker in order to prevent circularisation.
2. The 5'end of each primer contains a 9-nucleotide barcode consisting of 2 random nucleotides (N) followed by a 4-nucleotide fixed barcode and finished off with 3 more random nucleotides. The fixed barcode allows multiplexing of replicates or samples, which is important to reduce the costs of the downstream next generation sequencing. The random barcode allows for tracing and counting of individual cDNA

synthesis events, thus avoiding the overrepresentation of abundant RNAs which would otherwise be encountered by analysing exponentially amplified PCR product. The last 9 nucleotides of the RT primers have sequence complementarity to the final 9 nucleotides of the L3-App linker.

3. Enzymes that catalyse RNA modification via a two-step catalytic process as described in Figure 1 will be amenable to methylation-iCLIP. Several enzymes have been described which are known to catalyse RNA modifications via a covalent intermediate ***(28)***, however the biochemical mechanisms via which these adducts are resolved are often poorly defined. It is possible that amino acid residues within the catalytic domain are critical for resolution, as is observed in the NSun family of RNA m5C methylases, but these will require identification for each particular enzyme before methylation-iCLIP can be performed.
4. Generally, protein G for mouse antibodies, and protein A for rabbit antibodies is used.
5. RNA digestion conditions will need to be pre-optimised for each protein-RNA complex under study. We initially recommend using a 1:50 dilution of RNaseI for 'high RNase' and a 1:1000 dilution for 'low RNase'. High RNase treatment should result in a sharp band around the expected molecular weight of your RNA methylase of interest during the SDS-PAGE analysis (step 3.6), whereas the low RNase treatment should result in a smear above the expected molecular weight of the protein.
6. The high RNAse sample is only used in the protocol in order to check for the presence of the desired protein-RNA complex following immunoprecipitation, but is not used to prepare libraries as the RNAs would be too short in length to map uniquely to the genome. The RNAs present in the higher molecular weight smear in the low RNase condition are thus used for library preparation only.
7. A distinct RT primer (one of Rt1clip to Rt16clip) can be used for each of several replicates or samples if desired. The barcode on these RT primers (see note 2) will then allow for multiplexing of samples thus enabling the cost of next-generation-sequencing to be managed.
8. Gel purification is necessary to remove residual RT primer and also for size fractionation of cDNAs (see note 9).

9. The band cut at 70-80nt will contain a cDNA insert size of 18-28nt; this band usually consists of a high cDNA complexity, but due to the small size of inserts they do not always map to the genome. If the RNA methylase of interest is suspected to methylate microRNAs, then this band should be included for sequencing. The band cut at 80-150nt will contain insert sizes of 29-98nt with a high cDNA complexity; these cDNAs will also map uniquely to the genome and are the best cDNAs. The band cut at 150nt-200nt will contain insert sizes of 98-148nt but this fraction tends to consist of a limited cDNA complexity.
10. Most cDNA synthesis events will truncate at the crosslink site thus allowing the methylation site to be subsequently determined. In order to achieve PCR amplification of these cDNAs however, ligation of a 5' terminal adapter is required. This is achieved by performing a cDNA circularisation followed by a BamHI digestion of a restriction site originally located within the 3'terminal adapter (step 3). This results in a portion of the 3' adapter now being ligated to the 5'end, thus serving as a PCR priming site.
11. The size of the cDNA insert will be the size of PCR product minus the combined length of the P3/P5Solexa primers and the barcode (128 nt). Thus for example, a band cut at 70–80 nt on the cDNA gel (with 18-28nt cDNA + 52 nt primer) will generate 146–156 nt PCR products.

Figure legend

Figure 1

(A) Sequence alignment of a region of the catalytic domain of the seven known members of the human family of m5C RNA methylases. Two key conserved cysteine residues (Cys1 & Cys2) play critical roles in the methylation process. Identical residues in the alignment are indicated by asterix. **(B)** The conserved cysteine residue Cys1 forms a transient covalent bond (purple line) with the cytosine being methylated during the methylation process. A

second cysteine Cys2 is then critical for the resolution of this covalent bond to complete the catalytic process (left). In methylation-iCLIP, Cys2 is mutated to alanine (right) which results in the capture of the catalytic intermediate where the methylase is covalently bound to its substrate. This complex can then be immunoprecipitated using an antibody against the methylase or an epitope tag, and the trapped RNAs are then amenable to purification under highly stringent conditions. Libraries prepared from recovered RNA substrates are then subjected to high-throughput sequencing.

Acknowledgments

We wish to acknowledge Dr Julian Konig who co-developed the original iCLIP protocol, and Dr Yoichiro Sugimoto for helpful feedback and discussions during the development of methylation-iCLIP.

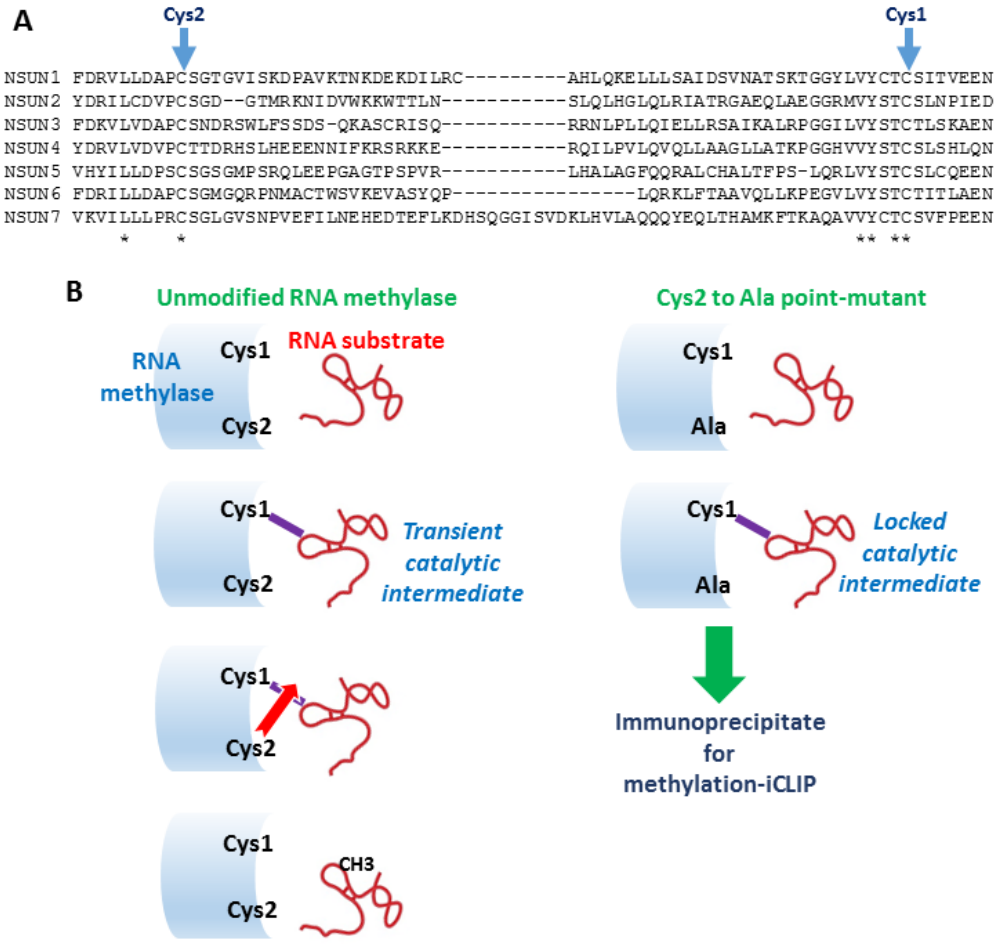
References

1. Dubin DT, Taylor RH. (1975). The methylation state of poly A-containing messenger RNA from cultured hamster cells. Nucleic Acids Res. 2:1653-68.
2. Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m^6A RNA methylomes revealed by m^6A-seq. Nature 485, 201–206.
3. Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell 149, 1635–1646.
4. Squires, J. E., Patel, H. R., Nousch, M., Sibbritt, T., Humphreys, D. T., Parker, B. J., et al. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res. 40, 5023–5033.
5. Khoddami, V., and Cairns, B. R. (2013). Identification of direct targets and modified bases of RNA cytosine methyltransferases. Nat. Biotechnol.31, 458–464.
6. Hussain, S., Sajini, A. A., Blanco, S., Dietmann, S., Lombard, P., Sugimot, Y., et al. (2013a). NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. Cell Rep. 4, 255–261.

7. Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N^6-methyladenosine-dependent regulation of messenger RNA stability.Nature 505, 117–120.
8. Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., et al. (2015). N(6)-methyladenosine modulates messenger RNA translation efficiency. Cell 161, 1388–1399.
9. Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, Pestova TV, Qian SB, Jaffrey SR. (2015). 5' UTR m(6)A Promotes Cap-Independent Translation. Cell. 163: 999-1010.
10. Hussain S, Bashir ZI. (2015) The epitranscriptome in modulating spatiotemporal RNA translation in neuronal post-synaptic function. Front Cell Neurosci. 9:420.
11. Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., and Rottman, F. M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. RNA 3, 1233–1247.
12. Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N^6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat. Chem. Biol. 7, 885–887.
13. Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB. (2003) CLIP identifies Nova-regulated RNA networks in the brain. Science. 302:1212-5.
14. Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, Darnell JC, Darnell RB. (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature. 456:464-9.
15. Zhang C, Darnell RB. (2011) Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. Nat Biotechnol. 29:607-14.
16. Sugimoto Y., Konig J., Hussain S., Zupan B., Curk T., Frye M., Ule J. (2012) Genome Biol. 13:R67.
17. König J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol. 17:909-15.
18. King, M.Y., and Redman, K.L. (2002). RNA methyltransferases utilize two cysteine residues in the formation of 5-methylcytosine. Biochemistry 41, 11218–11225.
19. Redman, K.L. (2006). Assembly of protein-RNA complexes using natural RNA and mutant forms of an RNA cytosine methyltransferase. Biomacromolecules 7, 3321–3326.
20. Hussain, S., Benavente, S.B., Nascimento, E., Dragoni, I., Kurowski, A., Gillich, A., Humphreys, P., and Frye, M. (2009). The nucleolar RNA methyltransferase Misu (NSun2) is required for mitotic spindle stability. J. Cell Biol. 186, 27–40

21. Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat. Methods 12, 767–772.
22. Carlile, T. M., Rojas-Duran, M. F., Zinshteyn, B., Shin, H., Bartoli, K. M., and Gilbert, W. V. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature 515, 143–146.
23. Schwartz, S., Bernstein, D. A., Mumbach, M. R., Jovanovic, M., Herbst, R. H, León-Ricardo, B. X., et al. (2014). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell 159, 148–162.
24. Saletore, Y., Meyer, K., Korlach, J., Vilfan, I. D., Jaffrey, S., and Mason, C. E. (2012). The birth of the epitranscriptome: deciphering the function of RNA modifications. Genome Biol. 13, 175.
25. Hussain, S., Aleksic, J., Blanco, S., Dietmann, S., and Frye, M. (2013b). Characterizing 5-methylcytosine in the mammalian epitranscriptome. Genome Biol. 14, 215.
26. Rinn, J. L., and Ule, J. (2014). Oming in on RNA–protein interactions. Genome Biol. 15, 401.
27. Chen B, Yun J, Kim MS, Mendell JT, Xie Y. (2014) PIPE-CLIP: a comprehensive online tool for CLIP-seq data analysis. Genome Biol. 15:R18.
28. Guelorget A and Golinelli-Pimpaneau B. (2011) Mechanism-based strategies for trapping and crystallizing complexes of RNA-modifying enzymes. Structure. 19:282–291.

Figure 1


A
Cys2
Cys1
NSUN1 FDRVLLDAPCSGTGVISKDPAVKTNKDEKDILRC---------AHLQKELLLSAIDSVNATSKTGGYLVYCTCSITVEEN
NSUN2 YDRILCDVPCSGD--GTMRKNIDVWKKWTTLN----------SLQLHGLQLRIATRGAEQLAEGGRMVYSTCSLNPIED
NSUN3 FDKVLVDAPCSNDRSWLFSSDS-QKASCRISQ----------RRNLPLLQIELLRSAIKALRPGGILVYSTCTLSKAEN
NSUN4 YDRVLVDVPCTTDRHSLHEEENNIFKRSRKKE----------RQILPVLQVQLLAAGLLATKPGGHVVYSTCSLSHLQN
NSUN5 VHYILLDPSCSGSGMPSRQLEEPGAGTPSPVR----------LHALAGFQQRALCHALTFPS-LQRLVYSTCSLCQEEN
NSUN6 FDRILLDAPCSGMGQRPNMACTWSVKEVASYQP----------LQRKLFTAAVQLLKPEGVLVYSTCTITLAEN
NSUN7 VKVILLLPRCSGLGVSNPVEFILNEHEDTEFLKDHSQGGISVDKLHVLAQQQYEQLTHAMKFTKAQAVVYCTCSVFPEEN
B
Unmodified RNA methylase
Cys2 to Ala point-mutant
RNA methylase
Cys1
Cys2
RNA substrate
Ala
Transient catalytic intermediate
Locked catalytic intermediate
CH3
Immunoprecipitate for methylation-iCLIP