Published in final edited form as:

Curr Protoc Mol Biol. 2019 April ; 126(1): e88. doi:10.1002/cpmb.88.

Transcriptome-wide mapping of m⁶A and m⁶Am at single-nucleotide resolution using miCLIP

Ben R. Hawley¹ and Samie R. Jaffrey^{1,2}

¹Department of Pharmacology, Weill Medical College, Cornell University, New York, NY 10065, USA

Abstract

The most prevalent modified base in mRNA, N⁶-methyladenosine (m⁶A), is found in several thousand transcripts, typically near the stop codon, but also throughout the coding sequence, 3'UTR, and 5'UTR (Desrosiers et al., 1974; Dominissini et al., 2012; Linder et al., 2015; Meyer et al., 2012; Perry et al., 1975). The highly similar nucleotide, N⁶,2'-O-dimethyladenosine is located at the first transcribed nucleotide of certain transcripts. These modifications have been implicated in numerous biological processes and as such, the precise mapping of m⁶A and m⁶Am is crucial to understand their function. We developed miCLIP, a method that maps both m⁶A and m⁶Am at individual nucleotide resolution. Here we describe this protocol, and slight improvements to both the experimental methodology and bioinformatics analysis.

Keywords

RNA modifications; epitranscriptomics; mapping; miCLIP; bioinformatics; m⁶A; m⁶Am

Introduction

Over the past several years, the new concept of epitranscriptomics has emerged in which the dynamic modification of mRNA heavily influences its fate in the cell. The most prevalent modified base is m⁶A, which can occur throughout the mRNA and has now been implicated in mRNA instability, splicing, and translation (Ke et al., 2015, 2017; Meyer et al., 2015; Sommer, Lavi, & Darnell, 1978; Wang et al., 2015; Xiao et al., 2016). In contrast, m⁶Am can only occur on the first transcribed nucleotide of an mRNA, and has been implicated in increased stability and translation (Boulas et al., 2018; Mauer et al., 2017).

Precise identification of m⁶A residues is challenging but necessary to understand its function. N⁶-methylated adenosine is found in a DRACH sequence consensus where D = A/G/U, R = A/G, H = A/C/U, and the methylated adenosine is always found directly upstream of a cytosine. However, not all DRACH are methylated *in vivo*.

Because adenosine, m⁶A, and m⁶Am have nearly identical chemical properties, there is currently no chemical method that results in selection modification that would allow these

²To whom correspondence should be addressed to S.R.J. (srj2003@med.cornell.edu).

nucleotides to be distinguished. Additionally, the presence of the N6 methylation does not appear to induce mutations or truncations during reverse transcription. Therefore, attempts to map m⁶A were not possible despite the clear evidence of the importance of this modification.

The first-generation m⁶A mapping approach, methyl-RNA immunoprecipitation and sequencing (MeRIP-Seq, also called m⁶A-Seq) (Dominissini et al., 2012; Meyer et al., 2012), allowed researchers, for the first time, to determine which transcripts contain m⁶A and identify 100 – 200 nt-wide regions in transcripts that are likely to contain m⁶A. m⁶A-binding antibodies are used to immunoprecipitate short RNA fragments which are then subject to high-throughput sequencing. Overlapping sequencing reads from m⁶A-containing fragments produce a “peak” whose summit reflects an underlying m⁶A residue (Meyer et al., 2012). However, the MeRIP-Seq mapping approach does not identify specific m⁶A residues and would instead attribute this to a DRACH motif near the point of highest read coverage (Schwartz et al., 2014).

However, this approach is complicated because m⁶A often appears in clusters, which can result in large peaks spanning several m⁶A residues (Meyer et al., 2012). Additionally, multiple DRACH motifs can be present underneath a peak, making it difficult to predict the specific methylated adenosine.

Since this method uses 6-methyladenine-specific antibodies, it binds both m⁶A and m⁶Am, since both contain this methylated nucleobase. Depending on the size of the mRNA fragments, m⁶Am-containing fragments are included in the sequencing library, and both m⁶A and m⁶Am can be detected. As a result, it is difficult to distinguish between m⁶A and m⁶Am, especially in mRNA 5'UTRs where m⁶Am is exclusively found, and m⁶A can be found as well.

We previously reported the miCLIP method, an adaptation of iCLIP whereby the UV crosslinking of anti-m⁶A antibodies to RNA induce specific mutational signatures that enable precise identification of m⁶A residues in RNA (Figure 1) (Linder et al., 2015). Following reverse transcription, a highly specific pattern of mutations or truncations are found at the crosslinked residues in the cDNA. These mutational signatures can then be inferred computationally to reveal precise positions of m⁶A residues. Through this method, m⁶A and m⁶Am residues can be mapped throughout the transcriptome at single-nucleotide resolution. Although m⁶Am is exclusively located at transcription-start site adenosines, the poor annotation of transcription-start site transcript isoforms can make it difficult to definitively assign a peak as m⁶A or m⁶Am. Methods to map m⁶A and m⁶Am, as well as criteria to facilitate the identification of the site as m⁶A or m⁶Am are described.

BASIC PROTOCOL 1: miCLIP library generation

To generate miCLIP libraries, poly(A)-selected RNA is first fragmented and incubated with anti-m⁶A antibodies, which is then crosslinked. This provides two benefits: first, stringent wash buffers can be used to reduce background interactions; and second, the resulting covalent interactions between antibody and nucleotides introduce specific mutagenic

signatures into cDNA when reverse transcribed. RNA fragments are ligated to adapters and these complexes are transferred to nitrocellulose membranes, where any remaining uncrosslinked RNA would not transfer. After elution from the membrane by digesting the antibody, RNA is reverse transcribed, and the cDNA purified and circularized. After re-linearization, the cDNA is PCR amplified to generate the final libraries.

NOTE: Use only molecular biology grade water and reagents (nuclease and nucleic acid free) for all solutions and buffers and filter purify.

Materials

TRIzol reagent (Invitrogen 15596018)

Oligo d(T)₂₅ magnetic beads (Thermo Scientific 61005 or NEB S1419S)

DNase I, RNase-free (Thermo Scientific EN0525)

RNA Clean & Concentrator kit (Zymo R1013)

5–20 µg of DNase-treated, poly(A)-selected RNA in 20 µl of water

RNA fragmentation reagent, including stop solution (Ambion AM9740)

Binding/low-salt buffer (see recipe)

Anti-m⁶A antibody (such as Abcam ab151230; see Critical Parameters)

Protein A/G magnetic beads (Thermo Scientific 88802)

High salt buffer (see recipe)

PNK wash buffer (see recipe)

5X PNK pH 6.5 buffer (see recipe)

T4 polynucleotide kinase (10 U/µl, NEB M0201S)

Ribonuclease inhibitor (Promega N2515 or Invitrogen 10777–019)

T4 RNA ligase 2, truncated K227Q (200 U/µl, NEB M0351S)

L3 linker oligo (20 µM)(see below)

PEG-8000

[γ-³²P]ATP, 3000 Ci/mmol (Perkin Elmer BLU002250UC) (see Critical Parameters)

4X LDS sample buffer (Invitrogen NP0008)

1M DTT (Sigma 646563)

NuPage Novex 4–12 % bis-tris protein gels (Invitrogen NP0321BOX)

MES SDS running buffer (Thermo Scientific NP0002)

Prestained protein standard (such as BioRad 161–0373)

Bis-Tris transfer buffer (Invitrogen NP0006)

Nitrocellulose membrane (0.45 µm pore such as BioRad 1620115)

Autoradiography film (such as Carestream Kodak BioMax MS films, Sigma Z363006)

SDS Proteinase K (PK) buffer (see recipe)

Proteinase K (20 mg/ml, Invitrogen 25530-049)

Phase-lock gel heavy tubes (5PRIME 2302830)

Acidic phenol:chloroform:isoamyl alcohol (25:24:1, pH 6.5)

Nucleic acid co-precipitant (such as GlycoBlue, Ambion AM9516)

2 M sodium chloride

100% ethanol

Barcoded reverse transcription (RT) primers (0.5 μ M; see below)

dNTP mix (10 mM each dNTP)

SuperScript III reverse transcriptase (Invitrogen Invitrogen 18080044)

3 M sodium acetate, pH 5.5

2X TBE-urea sample buffer (Invitrogen LC6876)

Novex 6% TBE-urea gel (Invitrogen EC6865BOX)

Low molecular weight DNA ladder (NEB N3233S)

1X TBE running buffer

SYBR Gold Nucleic Acid Stain (Invitrogen S11494)

Gel Breaker tubes (IST Engineering 3388-100)

1X Tris-EDTA (TE) buffer

CoStar Spin-X columns (0.22 μ m pore size, Corning CLS8160)

CircLigase II ssDNA ligase (Lucigen CL9025K)

FastDigest BamHI (Thermo Scientific FD0054)

Cut oligo (10 μ M; see below)

P3 and P5 PCR primers (10 μ M; see below)

AccuPrime SuperMix I (Invitrogen 12342010)

Novex 6% TBE gel (Invitrogen EC6265BOX)

AMPure XP magnetic beads (Beckman Coulter A63880)

Oligonucleotides

NOTE: All oligonucleotides can be purified using the supplier's standard method unless otherwise indicated.

L3 linker (HPLC-purified):

/5rApp/AGATCGGAAGAGCGGTTCAG/3ddC/

Reverse transcription primers (barcodes underlined):

NOTE: The barcodes will be reverse complemented in the raw sequencing data.

RT1: /5phos/NNaaccNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT2: /5phos/NNacaaNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT3: /5phos/NNattgNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT4: /5phos/NNaggtNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT5: /5phos/NNcgccNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT6: /5phos/NNccggNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT7: /5phos/NNctaaNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT8: /5phos/NNcattNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT9: /5phos/NNqccaNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT10: /5phos/NNgaccNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT11: /5phos/NNggttNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT12: /5phos/NNgtggNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT13: /5phos/NNtccgNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT14: /5phos/NNtgccNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT15: /5phos/NNtattNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

RT16: /5phos/NNttaaNNNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

Cut oligo (PAGE-purified):

GTTCAGGATCCACGACGCTCTTC/3ddC/

P3 Solexa PCR primer:

CAAGCAGAAGACGGCATAACGATCGGTCTCGGCATTCTGCTGAACCGCTCTCCGATCT

P5 Solexa PCR primer:

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

Collect RNA and enrich for mRNA

1. Grow cells following standard cell culture techniques. Following any treatments, aspirate cell media and extract total RNA by adding TRIzol reagent directly to the plate, following the manufacturer's protocol.

The quantity of cells required will depend on the chosen cell type, as the proportion of mRNA in total RNA can vary between 1–5%. We typically start with two to three 10 cm plates (or at least one 15 cm plate) per sample.

2. Enrich for mRNA by subjecting RNA to two rounds of poly(A) selection using oligo d(T)₂₅ magnetic beads according to the manufacturer's instructions. Briefly, total RNA is incubated with an appropriate volume of beads to allow poly(A) to hybridize. Beads are then separated by magnetism, allow non-poly(A) RNA to be washed away. Following the first elution, the beads can be washed and re-used for the second round. Beads must only be re-used for the same sample.
3. Digest any contaminating genomic DNA. Add DNase I buffer to 1X and 1 U of DNase I per 2 µg RNA, and incubate at 37°C for 30 minutes. Place on ice.
4. Purify enriched poly(A) RNA using either ethanol precipitation or an RNA Clean & Concentrator column following the manufacturer's instructions.

Fragment RNA

5. Resuspend or dilute samples to begin with 5–20 µg of DNase-treated, poly(A)-selected RNA in 20 µl water.
6. Add 2 µl of 10X Ambion fragmentation reagent to the RNA and mix. Incubate in a heat block at 75°C for 12 minutes.

The optimal RNA fragment lengths are between 30 and 130 nt; shorter than this will not produce useful sequencing libraries. We recommend determining the optimal time using the Ambion fragmentation reagent to obtain this size distribution of fragments, which should be around 12 minutes.

7. Place tube immediately on ice and add 2.2 µl of 10X Ambion stop solution and mix thoroughly. If sequencing of input is required, save a 1 µl aliquot of the fragmented RNA at –80°C.

Input libraries can be prepared using commercial small RNA library preparation kits (such as NEB #E7300).

Crosslink RNA-anti-m⁶A antibody complexes

8. Add binding/low-salt buffer to bring up the volume of the sample to 500 µl and add 10 µl of anti-m⁶A antibody (1 mg/ml stock) and 2 µl of RNase inhibitor.

For samples with 5–10 µg input RNA, using up to 25 µl of antibody at this step can help capture the greatest possible number of methylated RNA fragments.

9. Rotate the sample tube at in a cold room for 1.5 to 2 hours.
10. Transfer the sample to a single well of a 12-well tissue culture plate on ice in a shallow (~5 cm) tray.

Pre-chill the 12-well plate on ice prior to transferring sample.

11. Crosslink twice using 254 nm light with a dose of 150 mJ/cm² (such as in a Stratalinker 2400). Gently agitate the plate between irradiations. The plate lid must be removed during crosslinking.
12. Transfer the crosslinked sample to a new microcentrifuge tube and place on ice.
Safe stopping point: crosslinked samples may stored at –80°C and will remain stable for months.

Immunoprecipitate crosslinked complexes

13. Fully resuspend protein A/G magnetic beads and transfer 50 µl per sample to a new microcentrifuge tube. Wash the beads twice in 500 µl binding/low-salt buffer. Resuspend the washed beads in 100 µl binding/low-salt buffer. This can be performed just prior to crosslinking.
14. Add the beads to the crosslinked sample and rotate in a cold room for 1.5 h to overnight.
15. Wash the magnetic beads twice in 900 µl of high-salt buffer.
If desired, the supernatant containing the non-bound fraction can be saved.
To wash beads, remove from the magnet and resuspend in the new buffer. Return to the magnet to remove the wash buffer. To collect liquid, tubes can be briefly and gently centrifuged. This is particularly useful prior to the resuspension of beads in an enzymatic reaction (e.g. step 18), as any residual buffer will dilute the reaction and reduce its efficacy.
16. Wash the beads twice in 900 µl of binding/low-salt buffer.
17. Wash the beads twice in 500 µl of PNK wash buffer.

Dephosphorylate 3' ends of the RNA fragments:

18. Resuspend the magnetic beads in 20 µl of dephosphorylation master mix:
15 µl water
4 µl 5X PNK pH 6.5 buffer
0.5 µl T4 polynucleotide kinase (10 U/µl)
0.5 µl RNase inhibitor
19. Incubate the sample in a thermal mixer at 37°C for 20 min, shaking constantly at 1,100 RPM.
20. Wash the beads once in 500 µl of PNK wash buffer
21. Wash the beads once in 500 µl of high-salt buffer.
22. Wash the beads twice in 500 µl of PNK wash buffer.

Ligate the 3' linker oligo

23. Resuspend the magnetic beads in 20 µl of ligation master mix:

5 µl water
2 µl 10X T4 RNA ligase buffer
1 µl T4 RNA ligase 2, truncated K227Q (200 U/µl)
0.5 µl RNase inhibitor
1.5 µl L3 linker (20 µM)
10 µl 50% PEG-8000

24. Incubate the sample in a thermal mixer at 25°C for 3 hours, shaking intermittently at 1,100 RPM (30 s shaking followed by 4.5 minutes without shaking).
25. Wash the beads once in 500 µl of PNK wash buffer.
26. Wash the beads twice in 500 µl of high-salt buffer.
For each wash, rotate the sample tube at 4°C for 5 min.
27. Wash the beads twice in 500 µl of PNK wash buffer.

Radiolabel RNA fragments—NOTE: The following steps use a radioisotope. Consult Critical Parameters section for safe use of radioisotopes.

28. Resuspend the magnetic beads in 40 µl of radiolabeling master mix:
33 µl water
4 µl 10X T4 polynucleotide kinase buffer
1 µl [γ -³²P]ATP
2 µl T4 polynucleotide kinase (10 U/µl)
29. Incubate the sample in a thermal mixer at 37°C for 20 min, shaking constantly at 1,100 RPM.
30. Wash beads twice in 500 µl of binding/low-salt buffer.
31. Wash beads twice in 500 µl of PNK wash buffer.

Purify RNA-antibody complexes using SDS-PAGE and membrane transfer

32. Resuspend the magnetic beads in 20 µl of sample buffer:
14 µl water
5 µl 4X NuPage LDS sample buffer
1 µl 1 M DTT
33. Incubate the sample in a thermal mixer at 70°C for 15 min, shaking constantly at 1,100 RPM.
34. Place the sample tube in a magnetic rack and allow the magnetic beads to pellet completely for 2 min.

Safe stopping point: supernatant can be transferred to a fresh tube and stored at -80°C . To prevent loss of radioactive signal, proceed within a few days.

35. Prepare a NuPage Novex 4–12% Bis-Tris protein gel using MES buffer. Load the supernatant, keeping at least one empty lane between samples. In a separate lane, load 10 μl of a prestained protein standard.
36. Run the gel at $4-8^{\circ}\text{C}$ in a cold room at 200 V for 35 min.
37. Set up transfer of the RNA-antibody complexes from the gel to a nitrocellulose membrane in an electrophoretic transfer unit according to the manufacturer's instructions. Transfer in a cold room at 400 mA for 30 minutes.

Discard the wells and running front of the gel (and the used running buffer) in radioactive waste. Successful transfer of the radiolabeled RNA-antibody complexes can be verified using a Geiger counter.

38. Wrap the membrane in plastic wrap and secure in an autoradiography cassette. Expose to autoradiography film with an intensifying screen at -80°C .

The gel can be sealed in plastic wrapping and exposed alongside the membrane to validate that the transfer was complete.

Exposure times will vary depending on the amount of input RNA, antibody binding and immunoprecipitation efficiency, and activity of the $[\gamma-^{32}\text{P}]\text{ATP}$. An initial exposure of 1 hour should be sufficient; this can be used to determine if a longer exposure is required.

Elution of RNA by antibody digestion

39. Excise the sections of the developed film that contain signal (Figure 2). Align to the plastic-wrapped membrane and use as a template to mark the regions containing the RNA-antibody complexes with a standard lab marker.
40. Excise the marked regions of the membrane using a sterile razor blade for each sample and discard the excised piece of plastic wrap. Cut the membrane into small fragments and transfer to a microcentrifuge tube.

Use sterile forceps to handle plastic wrap and membrane pieces. Sterilize forceps between each use. Cut just outside of the markings to ensure all sample is excised.
41. Add 200 μl of SDS Proteinase K buffer to the sample tube. If the membrane pieces are not submerged, add up to an additional 200 μl SDS Proteinase K buffer.
42. Add 10 μl Proteinase K (20 mg/ml) and incubate at 50°C in a thermal mixer for 60 min, shaking constantly at 1,100 RPM.
43. Briefly centrifuge the membrane fragments and transfer the supernatant to a phase-lock gel heavy tube. Add an equal volume of acidic phenol:chloroform:isoamyl alcohol.

Pre-centrifuge the phase-lock gel tube for 2 min at 20,000 RCF prior to adding sample.

44. Incubate the sample in a thermal mixer at 30°C for 5 min, shaking constantly at 1,100 RPM.
45. Centrifuge the sample tube at 20,000 RCF for 5 min, then transfer the supernatant to a new microcentrifuge tube.

Avoid touching or disrupting the gel interphase when using phase-lock gel tubes. Used phase-lock gel tubes must be disposed appropriately (as phenol waste).

46. Add the following to the supernatant:

Please note the use of sodium chloride as the salt used in the precipitation at this step. The typical salt, sodium acetate, does not remove the SDS found in the proteinase K buffer.

1 µl GlycoBlue

0.1 volume of 2 M sodium chloride

2.5 volumes 100% ethanol

47. Precipitate the RNA fragments on ice for 1 hour or at –20°C overnight.

Reverse transcribe RNA fragments

48. Centrifuge the precipitated RNA fragments at 20,000 RCF for 30 min at 4°C. Discard the supernatant.
49. Add 800 µl of 75% ethanol to the sample tube and vortex. Centrifuge at 20,000 RCF for 5 min at 4°C. Discard the supernatant and allow the RNA pellet to air dry for 5 minutes.

A brief centrifugation following removal of the ethanol wash can be performed to collect the excess at the bottom, which can then be easily removed to reduce the time required to air dry the pellet. When air drying RNA pellets, avoid over drying as this can drastically reduce the ability to resuspend the RNA. This is especially important when resuspension is to occur in an enzymatic reaction buffer.

50. Resuspend the RNA pellet in 7 µl of denaturing master mix:

5 µl water

1 µl RT primer (0.5 µM)

NOTE: Use an RT primer with a unique experimental barcode for each sample.

1 µl dNTP mix (10 mM)

51. Transfer the resuspended RNA to a new 0.2 ml PCR tube. Incubate in a thermocycler:

70°C for 5 min

Pause at 25°C

52. Add 13 µl RT master mix to the sample:

7 µl water

4 µl 5X First strand buffer

1 µl 100 mM DTT

0.5 µl RNase inhibitor

0.5 µl SuperScript III

53. Incubate the sample in a thermocycler:

25°C for 5 min

42°C for 20 min

50°C for 40 min

80°C for 5 min

Pause at 4°C

54. Transfer the sample to a microcentrifuge tube and add:

230 µl water

1 µl GlycoBlue

25 µl 3 M sodium acetate, pH 5.5

1000 µl 100% ethanol

55. Precipitate the cDNA on ice for 1 hour or at –20°C overnight.

Purify the cDNA fragments

56. Centrifuge the precipitated RNA fragments at 20,000 RCF for 30 min at 4°C. Discard the supernatant.

57. Add 800 µl of 75% ethanol to the sample tube and vortex. Centrifuge at 20,000 RCF for 5 min at 4°C. Discard the supernatant and allow the RNA pellet to air dry for 5 minutes.

58. Resuspend the cDNA pellet in 12 µl 1X TBE-urea sample buffer and denature at 65°C for 5 min. Place tube immediately on ice.

59. Load the sample on a Novex 6% TBE-urea gel according to the manufacturer's instructions. In a separate lane, load 0.5 µl of a low molecular weight DNA ladder.

Leave at least one empty lane between samples and ladder.

60. Run the gel at 200 V for 35 min, or until the dark blue dye of the sample buffer has migrated approximately three-quarters of the way through the gel.

61. Using a sterile razor blade for each sample, remove the lane of the gel carrying the ladder and stain this portion of the gel using 1X SYBR gold in gel running buffer for 5 min.

62. Place the gel and the stained ladder lane on clean plastic wrap on a UV transilluminator. Using the stained ladder and dye markers as a guide, cut three bands with a sterile razor blade: 120–200 nt (high), 85–120 nt (medium), and 70–85 nt (low).

A piece of paper placed underneath the gel containing sample can prevent unwanted exposure of these to UV light, while allowing the ladder to be visualized.

We strongly recommend that at least the high and medium gel slices be collected separately for each individual sample. Do not extract gel slices lower than 70 nt to avoid inclusion of unincorporated RT primers and L3 linkers.

63. Transfer each of the gel slices into a gel breaker tube placed into a new microcentrifuge tube. Centrifuge each sample at 20,000 RCF for 2 min to crush the gel slices. Discard the gel breaker tubes.

As an alternative to commercial gel breaker tubes, a 19G needle can be used to pierce 2–3 holes in the bottom of 0.5 ml tubes.

64. To elute cDNA, add 400 µl 1X TE buffer to the crushed gel. Freeze the sample in a –80°C freezer for 5 min, then shake constantly at 1,100 RPM at 4°C overnight.

65. Transfer the sample to a CoStar SpinX column and centrifuge at 20,000 RCF for 2 min. Transfer the eluate to a phase-lock gel heavy tube.

66. Add 400 µl of acidic phenol:chloroform:isoamyl alcohol to the tube and in a thermal mixer at 30°C for 5 min, shaking constantly at 1,100 RPM.

67. Centrifuge at 20,000 RCF for 5 min and transfer the supernatant to a new tube.

68. Add the following:

1 µl GlycoBlue

40 µl 3 M sodium acetate, pH 5.5

1000 µl 100% ethanol

69. Precipitate the cDNA on ice for 1 hour or at –20°C overnight.

Circularize and cut the cDNA fragments

70. Centrifuge the precipitated RNA fragments at 20,000 RCF for 30 min at 4°C. Discard the supernatant.

71. Add 800 µl of 75% ethanol to the sample tube and vortex. Centrifuge at 20,000 RCF for 5 min at 4°C. Discard the supernatant and allow the RNA pellet to air dry for 5 minutes.

72. Resuspend the cDNA pellet in 8 µl of circularization master mix:

6.5 µl water
0.8 µl 10X CircLigase buffer
0.4 µl 50mM MnCl₂
0.3 µl CircLigase II

- 73.** Transfer the sample to a 0.2 ml PCR tube and incubate in a thermocycler using the following settings:

60°C for 90 min

Pause at 4°C

- 74.** Add 30 µl of oligo annealing master mix to the sample and mix gently:

25 µl water

4 µl 10X FastDigest buffer

1 µl Cut oligo (10 µM stock)

- 75.** Allow the cut oligo to anneal to the circularized cDNA using the following settings:

95°C for 2 min

Decrease temperature by 1°C every 20 sec until 25°C is reached

Pause at 25°C

- 76.** Add 2 µl FastDigest BamHI to the sample and linearize the cDNA using the following settings:

37°C for 30 min

80°C for 5 min

- 77.** Transfer the cDNA to a new tube and add the following:

210 µl water

1 µl GlycoBlue

25 µl 3 M sodium acetate, pH 5.5

1000 µl 100% ethanol

- 78.** Precipitate the cDNA on ice for 1 hour or at –20°C overnight.

PCR amplify the cDNA libraries

- 79.** Centrifuge the precipitated RNA fragments at 20,000 RCF for 30 min at 4°C. Discard the supernatant.

- 80.** Add 800 µl of 75% ethanol to the sample tube and vortex. Centrifuge at 20,000 RCF for 5 min at 4°C. Discard the supernatant and allow the RNA pellet to air dry for 5 minutes

- 81.** Resuspend the cDNA pellet in 21 μ l of water.
- The cDNA can now be stored at -20°C , however we recommend avoiding freeze-thaw cycles.
- 82.** Prepare three test PCR reactions for each sample:
- 3.75 μ l water
 - 1 μ l cDNA
 - 0.25 μ l P5 and P3 Solexa PCR primer mix (10 μ M stock)
 - 5 μ l AccuPrime SuperMix I
- 83.** Incubate the reactions in a thermocycler using the following settings:
- 94 $^{\circ}\text{C}$ for 2 min
 - Then, 15, 20, or 25 cycles of:
 - 94 $^{\circ}\text{C}$ for 15 sec
 - 65 $^{\circ}\text{C}$ for 30 sec
 - 68 $^{\circ}\text{C}$ for 30 sec
 - Pause at 4 $^{\circ}\text{C}$
- 84.** Add 2.5 μ l of 5X TBE sample buffer to each sample and load on a Novex 6% TBE gel according to the manufacturer's instructions. In a separate lane, load 0.5 μ l of a low molecular weight DNA ladder.
- 85.** Stain the gel with 1X SYBR Gold in gel running buffer for 5 minutes. Visualize the stained gel on a UV transilluminator or BioRad ChemiDoc imager (Figure 3).
- 86.** 86 Determine the appropriate number of PCR cycles for the preparative PCR reaction.
- We recommend to determine the optimum number of PCR cycles for each sample using the outcome of the three test PCR reactions. In general, the optimum number of PCR cycles is the number required to faintly see the PCR products on the gel. However, because the amount of cDNA template used for the preparative PCR is greater than the amount used for the test PCR, we recommend that the number of cycles used for the preparative PCR be at least one less. Typically, we prepare the medium sized library for less than 15 cycles.
- 87.** Assemble the preparative PCR reaction:
- 11 μ l water
 - 8 μ l cDNA
 - 1 μ l of P5 and P3 Solexa primer mix (10 μ M stock)
 - 20 μ l AccuPrime SuperMix I

88. Amplify the cDNA in a thermocycler for the determined number of cycles using the settings in step 83.
89. Purify the PCR products using 1.8 volumes of AMPure XP magnetic beads (Beckman Coulter) using the manufacturer's instructions. Elute the purified PCR products in 30 μ l to 50 μ l of water.

The sample barcodes within the reverse transcription primer allows samples to be pooled. If the appropriate number of preparative PCR cycles was chosen for each sample to produce equally concentrated libraries, these can be pooled prior to AMPure XP purification. We recommend 2 rounds of bead purification to ensure removal of all primers and primer dimers.

90. Submit an appropriate quantity of the purified library for high-throughput sequencing.

We recommend sequencing miCLIP libraries in paired-end mode with a read length of 50 base pairs. For libraries of up to 6 unique samples, we find that sequencing the samples on a single lane of a high-throughput flowcell is sufficient for adequate sequencing coverage (~30 million reads a sample).

However, for a greater number of unique samples, or if higher sequencing depth is desired, additional lanes should be used.

BASIC PROTOCOL 2: miCLIP bioinformatics analysis

A robust bioinformatics pipeline is crucial to reliably identify m^6A and m^6Am nucleotides. The most common goal with a miCLIP experiment is to identify these modifications throughout the transcriptome in a particular cell or tissue type, or following a treatment.

This walkthrough describes our pipeline for this analysis. Starting with raw sequencing reads in fastq format, we remove adapter sequences and demultiplex the pooled samples using the barcodes found in the RT primers. The read sequences and the five random nucleotides also found in the barcode are then used to filter out PCR duplicates. The pre-processed reads are then aligned to a reference genome.

We then utilize components of the CTK package (a series of scripts developed to analyze crosslink experiments such as iCLIP) to identify substitution and truncation events as a result of antibody crosslinking at m^6A . As m^6A is found in a DRACH consensus (D = A/G/U, R = A/G, H = A/C/U), we utilize C $\rightarrow$ T transitions to identify m^6A , which is by far the most common substitution we observe (Boulias et al., 2018). Significant sites are then filtered further to identify those in the DRACH consensus, thus building the list of high-confidence m^6A sites.

More recently, there has been increased interest in the similar modification m^6Am which can occur at the first transcribed nucleotide of mRNA (Boulias et al., 2018; Mauer et al., 2017). As anti- m^6A also enriches for m^6Am , a researcher can identify m^6Am sites from a miCLIP dataset. We offer two approaches to do so. With m^6Am , peaks often appear asymmetrical due to the pileup of reads beginning at the start of the transcript. One method is to identify

the beginning of these peaks. The alternative utilizes A→T transitions which occur at a modified adenosine. With m⁶Am, these then need to be filtered for a transcription start site (TSS). While one can use reference annotations, more appropriate would be the use of a TSS-mapping dataset in the particular cell type and condition (such as cap analysis gene expression (CAGE) data). With the recent identification of the sole m⁶Am writer enzyme PCIF1 (Boulias et al., 2018), a miCLIP with this depleted can be used to identify regions of m⁶Am, and then the previous methods used to pinpoint the exact m⁶Am nucleotide.

Materials

See Table 1 for a list of software and files required for miCLIP analysis.

Processing of raw sequencing reads

1. Trim 3' adaptor sequences using flexbar (Dodt, Roehr, Ahmed, & Dieterich, 2012):

Orphan reads generated at this step (due to a mate read not passing quality filters) can be processed as in all subsequent steps and concatenated into the final read file in step 7.

```
flexbar -r forward_reads.fastq -p reverse_reads.fastq -f 11.8 -
as AGATCGGAAGAGCGGTTCAG --pre-trim-phred 30 -s -t sample
```

2. Demultiplex reads using the RT primer barcodes used. We use the pyCRAC suite (Webb, Hector, Kudla, & Granneman, 2014) because it preserves the random portion of the barcode when demultiplexing paired-end reads, by placing the first three nucleotides of the random barcode into the header of each read.

The “barcodes.txt” file should contain one line for each barcode in the format NNNXXXX, where NNN is the first part of the random barcode and XXXX is the reverse complement of the experiment barcode. For example, barcode 1 is NNNGGTT. If this script generates orphan reads, they can be processed like those in step 1.

```
pyBarcodeFilter.py -f sample_1.fastq -r sample_2.fastq -b
pyCRAC_barcodes.txt
```

3. Move the remainder of the random barcode (the two nucleotides following the experimental barcode) to the read headers:

The “getline” function used here is specific to the “mawk” interpretation of awk; if this is not the version of awk used in your system it should be called explicitly as in the example given.

```
mawk -F "###" '{sub(/.../, "###"$2, $2); getline($3); $4 =
substr($3,1,2); $5 = substr($3,3); print $1 $2 $4"\n"$5}'
sample_1_NNNGGTT_cond1.fastq > sample_1_NNNGGTTNN_cond1.fastq
```

4. Collapse PCR duplicates using pyCRAC (Webb et al., 2014):


```
pyFastqDuplicateRemover.py -f sample_1_NNNGGTTNN_cond1.fastq -r
sample_2_NNNGGTT_cond1.fastq -o dedup_reads
```

5. Transform the header in the forward read file to be compatible with the downstream CTK analysis:

```
mawk -F '[_/]' '/^>/{print $1_"$2_"$3"/"$4#"3#"2;
getline($9); print $9}' dedup_reads_1.fasta >
dedup_reads_1.cims.fasta
```

6. Use the header information of the forward reads to create a CTK-compatible header for the reverse reads, and reverse-complement these reads. This strategy reduces the noise of mutations introduced by sequencing errors and increases the confidence in a crosslink-induced substitution. We had previously developed a perl script for this purpose that can be found at https://github.com/jaffreylab/miCLIP_MiMB_2016:

```
perl match_barcodes_and_reverse_complement.pl
dedup_reads_1.cims.fasta dedup_reads_2.fasta
```

7. Concatenate the two reads files.

Previously generated orphan reads can also be concatenated into this file.

```
cat dedup_reads_1.cims.fasta dedup_reads_2.cims.rc.fasta >
reads.cims.cat.fasta
```

8. Align the concatenated reads to an index of the reference genome. If using bwa:

```
bwa aln -t 4 -n 0.06 -q 20 /path/to/index/index.fa
reads.cims.cat.fasta > reads.cims.cat.sai
```

```
bwa samse /path/to/index/index.fa reads.cims.cat.sai
reads.cims.cat.fasta > reads.cims.cat.sam
```

As this workflow makes heavy use of the CTK package, we now recommend using the CTK authors' preferred alignment tool bwa. Any aligner that produces SAM files with MD tags can be used (as these tags are utilized by the CTK `parseAlignment.pl` script). Options include bowtie and bowtie2; other tools should be usable however this has not been tested in our hands.

Call m⁶A sites with the CTK pipeline

9. Parse alignment output to generate a bed file of read coordinates and a separate file of mutation coordinates. If using bwa (or any tool that produces SAM files with MD tags):

```
perl parseAlignment.pl -v --map-qual 1 --min-len 18 --mutation-
filereads.cims.mutation.txt reads.cims.cat.sam reads.cims.tag.bed
```

10. Collapse PCR duplicates based on read coordinates and barcode identities:

```
perl tag2collapse.pl -v --random-barcode -EM 30 --seq-error-
model alignment --weight-in-name --keep-max-score --keep-tag-
name reads.cims.tag.bed reads.cims.tag.uniq.bed
```

At this step, it is essential that the random barcode in the header of the reverse reads was reverse complemented (step 6). Otherwise, read pairs will be collapsed as PCR duplicates.

11. Remove mutations found in PCR duplicated reads:

```
python joinWrapper.py reads.cims.mutation.txt
reads.cims.tag.uniq.bed 4 4 N reads.cims.tag.uniq.mutation.txt
```

12. Extract C→T transitions and generate a .bed file of their coordinates:

```
awk '{if($6=="+" && $8=="C" && $9==">" && $10=="T" || $6=="-"
&&$8=="G" && $9==">" && $10=="A") {print $0}}'
reads.cims.tag.uniq.mutation.txt | cut -f 1-6 >
reads.cims.tag.uniq.C2T.bed
```

13. Evaluate the reproducibility of mutations by permutation using CIMS.pl:

```
perl CIMS.pl -big -n 5 -c ./reads_cache_C2T --keep-cache -p -
outp reads.C2T.stats.txt -v reads.cims.tag.uniq.bed
reads.cims.tag.uniq.C2T.bed reads.cims.tag.uniq.C2T.CIMS.txt
```

14. Call m⁶A sites located in DRACH motifs.

Our lab previously created scripts to process the output from CIMS.pl and CITS.pl, which can be found at https://github.com/jaffreylab/miCLIP_MiMB_2016. This script filters for number of transitions ($m \geq 2$) and transition frequency (1% $m/k \leq 50\%$) for each site. Prior to running these scripts, the output files can be filtered for more significant FDR. We recommend an initial run-through with an FDR of 1. Depending on sequencing depth, levels of PCR duplication, and efficacy of the particular antibody lot, more or less m⁶A sites can be identified. We typically observe around 10,000 m⁶A sites per sample, with particularly successful experiments identifying around 60,000.

```
kmer.annotate.cims.sh reads.cims.tag.uniq.C2T.CIMS.txt genome.fasta
```

15. Following identification of m⁶A sites, a researcher may wish to see if the metagene of these is typical of m⁶A metagene profiles, indicating a successful miCLIP experiment. Our lab previously published a robust pipeline for this purpose called MetaPlotR (Olarerin-George & Jaffrey, 2017). The scripts and a walkthrough of this can be found at <https://github.com/olarerin/metaPlotR>. The output from CIMS.pl, CITS.pl, and their processed files (as well as called m⁶Am sites) are all suitable for this pipeline.

Calling m⁶A sites with CITS—For calling truncations, it is recommended to only use the forward reads. After step 5, skip to alignment in step 8 and proceed through to step 11 to generate the unique reads and mutations files.

16. Obtain deletions from the de-duplicated mutation file in 11:

```
perl getMutationType.pl -t del reads.cims.tag.uniq.mutation.txt
reads.cims.tag.uniq.del.bed
```

17. Use the CITS pipeline to identify truncation events:

```
perl CITS.pl -big -p 0.001 -gap 25 -v reads.cims.tag.uniq.bed
reads.cims.tag.uniq.del.bed reads.tag.uniq.CITS.bed
```

18. The output from CITS.pl can be annotated in a similar to manner to the CIMS method in step 14 using this script found at https://github.com/jaffreylab/miCLIP_MiMB_2016:

```
kmer_annotate_cits.sh reads.tag.uniq.CITS.bed genome.fasta
```

Identifying m⁶Am sites using read drop-off/pileup method—This method attempts to capture the 5'-most nucleotide of the asymmetrical peaks found at the transcription start site (TSS), typical of m⁶Am. A→T transitions can also be used (see below). *These m⁶Am calling methods aim to identify 6mA peaks at the TSS. To be sure this is truly m⁶Am and not a 5'UTR m⁶A (or is an m⁶Am at a rare TSS isoform) an additional identification method should be used. This could be an mRNA 5' end mapping method (e.g. CAGE) or a sample where the m⁶Am writer PCIF1 (Boulias et al., 2018) is depleted to identify regions of m⁶Am.*

19. Take the first nucleotide of each read (do not forget to account for strand):

```
awk 'BEGIN{OFS="\t"}{if ($6 == "+")print $1, $2, $2+1, $4, $5,
$6; else print $1, $3-1, $3, $4, $5, $6}' reads.cims.tag.uniq.bed >
reads.firstnt.bed
```

20. Pileup first nucleotides using tag2clusters.p from CTK and filter low coverage clusters:

```
perl tag2cluster.pl -s -v -maxgap "-1" reads.firstnt.bed
reads.firstnt.clusters.bed
```

```
awk '$5 >= 5' reads.firstnt.clusters.bed | sort -k1,1 -k2,2n >
reads.firstnt.clusters.filtered.bed
```

21. Merge 0-distance clusters (those that are immediately adjacent), remove all non-merged clusters (i.e. orphan pileups; these are most likely noise), then take the first nucleotide of these clusters:

```
bedtools merge -s -c 4,5,6 -o count,max,first -I
reads.firstnt.clusters.filtered.bed | awk 'BEGIN{OFS="\t"}{print
$1,
$2, $3, $5, $6, $4}' | awk '$3-$2 >1' | awk 'BEGIN{OFS="\t"}{if
($6 ==
"+")print $1, $2, $2+1, $4, $5, $6; else print $1, $3-1, $3, $4,
$5,
$6}' | sort -k1,1 -k2,2n > reads.firstnt.onlymergedclusters.bed
```

22. Find those clusters at or within a few nucleotides of an annotated A-starting TSS (or ideally, CAGE/5' end maps within the same cells and conditions).

Identifying m⁶Am sites using A→T transitions—While C→T transitions are the most common crosslink-induced modification at m⁶A sites, m⁶Am can be followed by any nucleotide and thus the only constant is the modified adenosine. This is frequently transitioned to a T and this signature can be used to identify the precise m⁶Am.

23. Extract all A→T transitions:

In our experience this does not always extract all of the A→T transitions. An alternative approach is to use samtools mpileup and extract the substitutions in this way.

```
awk '{if($6=="+" && $8=="A" && $9==">" && $10=="T" || $6=="-"
&& $8=="T" && $9==">" && $10=="A") {print $0}}'
reads.cims.tag.uniq.mutation.txt | cut -f 1-6 >
reads.cims.tag.uniq.A2T.bed
```

24. Filter for transitions above a certain ratio. We chose a 10% frequency (Boulias et al., 2018); however this should be determined by observing the A→T frequency at a few known m⁶Am genes such as *ATF4*. Then find those at or close to a TSS/ CAGE data as in step 22.

Reagents and Solutions

Binding/low-salt buffer:

50 mM Tris HCl pH 7.4

150 mM NaCl

0.5 % NP-40

Store indefinitely at room temperature

High-salt buffer:

50 mM Tris HCl pH 7.4

1 M NaCl

1 % NP-40

0.1 % SDS

Store indefinitely at room temperature

PNK wash buffer:

20 mM Tris HCl pH 7.4

10 mM MgCl₂

0.2 % Tween-20

Store indefinitely at room temperature

PNK pH 6.5 buffer, 5X:

350 mM Tris HCl pH 6.5

50 mM MgCl₂

5 mM DTT

Store single-use aliquots indefinitely at –20°C

SDS Proteinase K (PK) buffer:

100 mM Tris pH 7.5

50 mM NaCl

1 mM EDTA

0.2% SDS

Store indefinitely at room temperature

COMMENTARY

Background Information

The miCLIP protocol is currently the only way to obtain nucleotide-level resolution of m⁶A and m⁶Am throughout the transcriptome. miCLIP offers several advantages over other less precise methodologies such as MeRIP-Seq/m⁶A-Seq. Crosslinking of the anti-m⁶A antibody means stringent washes can be used to drastically reduce non-specific RNA inclusion into libraries, and thus less noise in the sequencing data. The ability to pinpoint precise modified nucleotides means adjacent m⁶A in clusters can be properly identified, where such a cluster would appear as a single peak in other methods. This is also important in the identification of m⁶Am, as it is possible to distinguish between an N6-methylated adenosine at the start nucleotide (m⁶Am) and a 5'UTR m⁶A. A disadvantage with miCLIP is the requirement for large quantities of starting material (at least 5 µg poly(A)-selected RNA). While we have produced usable libraries with as little as 1 µg mRNA, this is not recommended due to the increased chance of failure.

Critical Parameters and Troubleshooting

Cleanliness guidelines for miCLIP—Researchers performing miCLIP must work in a clean environment that is sterile and free of RNase and contaminating DNA. A source of

highly detrimental contamination is previously prepared libraries, as these sequences can be amplified and sequenced, potentially ruining an entire experiment. General RNA cleanliness guidelines can be found here: <https://www.neb.com/tools-and-resources/usage-guidelines/avoiding-ribonuclease-contamination>. Unless PCR laminar flow cabinets are available, the bench must be sterilized and kept in this way for the duration of the experiment. We routinely clean benches, pipettes and tube trays with RNase-Zap (Ambion), DNA-Zap, then 70% ethanol. After cleaning, and when not in use, cover the bench and tube trays with paper towels or an absorbent pad. Finally, use a fresh bag of tubes and pipette filter tips for the particular miCLIP only, and change gloves regularly.

Use of radioisotopes—Radioactive ATP is used from step 28. All steps involving use of radioactive materials require use of protective equipment, including gloves, wrist guards, a lab coat, and eye protection, and must be carried out behind Plexiglas shielding of 3/8-inch thickness or greater. Radioactive materials must be used and disposed of in strict accordance with federal and institutional regulations and guidelines.

Use Geiger counters to track radioactive material during these steps (steps 28–47). This is useful to appreciate loss of material during bead washes and the gel transfer in step 37 (the gel can also be sealed in plastic wrapping and exposed along with the membrane to validate that transfer was complete). Some residual radioactive signal will remain on beads (after elution) and in the gel (after transfer). This is due to the vast excess of radioactive ATP, and residual magnetic beads.

Choice of anti-m⁶A antibody—We routinely use rabbit polyclonal anti-m⁶A antibodies from Abcam (ab151230) due to its high reproducibility. Several other antibodies are available and can be tested for use in miCLIP; however, a researcher utilizing a different antibody should compare this directly with ab151230 to ensure a similar substitution and truncation profile is observed.

Using magnetic beads—Advice for using magnetic beads is included as notes at relevant steps in the protocol. In general, fully resuspend all beads in wash buffers using a pipette and aspirate slowly to ensure no beads get trapped in the tip. If this does occur, additional wash buffer can be used.

Place tubes containing beads back onto the magnet and allow separation to occur for 30 to 60 seconds, until the solution is clear.

Following resuspension of beads in the final wash, briefly spin to collect all liquid in the bottom of the tube and take care to remove all wash buffer (including any small bubbles that may stick to the side of the tube).

Loss of beads and dilution of enzymatic reaction due to residual wash buffer can lead to impaired library generation.

General tips to improve success—In general, increasing the starting material will increase the success of miCLIP. We regularly use 10 µg of 2X poly(A)-selected and DNA-free RNA.

When preparing and using buffers and reagents that contain DTT and dNTPs, we recommend making single-use aliquots before freezing.

Use low-bind/low-retention microcentrifuge tubes where possible (such as Eppendorf 022431021). Use low-retention, PCR-grade, filtered pipette tips.

Understanding the Results

If it is possible to start with 10 µg of poly(A)-selected RNA, we find that miCLIP is very successful, and we can amplify libraries with around 13 PCR cycles thus having low PCR duplication. Depending on the antibody lot, we find varying rates of substitutions and thus called m⁶A sites. However, we reliably identify around 10,000 m⁶A sites, with several notable experiments identifying over 60,000.

Time Considerations

If all safe stopping points are utilized and precipitations carried out overnight, the miCLIP protocol will take over a week. It is possible to complete this in only a few days by proceeding through to step 34 on the first day, then step 64 on the second day. For a researcher experienced with bioinformatics analyses, the analysis pipeline takes around 1–2 days and is most limited by processing power. As many of the scripts used are single-threaded, we recommend running multiple samples in parallel rather than looping one after another to speed up this process.

Acknowledgments

This work was adapted from Grozhik A.V., Linder B., Olarerin-George A.O., Jaffrey S.R. Mapping m⁶A at Individual-Nucleotide Resolution Using Crosslinking and Immunoprecipitation (miCLIP). *Methods Molecular Biology*, 1562:55–78, 2017. We thank members of the Jaffrey lab for their helpful comments and support. This work was supported by NIH grants CA186702 and DA037150 (S.R.J.).

Literature Cited

- Boulias K, Toczydlowska-Socha D, Hawley BR, Liberman-Isakov N, Takashima K, Zaccara S, ... Greer EL (2018). Identification of the m⁶Am methyltransferase PCIF1 reveals the location and functions of m⁶Am in the transcriptome. *BioRxiv*, 485862 10.1101/485862
- Desrosiers R, Friderici K, & Rottman F (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proceedings of the National Academy of Sciences of the United States of America*, 71(10), 3971–3975. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/4372599> [PubMed: 4372599]
- Dotz M, Roehr J, Ahmed R, & Dieterich C (2012). FLEXBAR—Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. *Biology*, 1(3), 895–905. 10.3390/biology1030895 [PubMed: 24832523]
- Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, ... Rechavi G (2012). Topology of the human and mouse m⁶A RNA methylomes revealed by m⁶A-seq. *Nature*, 485, 201 Retrieved from 10.1038/nature11112 [PubMed: 22575960]
- Ke S, Alemu EA, Mertens C, Gantman EC, Fak JJ, Mele A, ... Darnell RB (2015). A majority of m⁶A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes & Development*, 29(19), 2037–2053. 10.1101/gad.269415.115 [PubMed: 26404942]
- Ke S, Pandya-Jones A, Saito Y, Fak JJ, Vågbo CB, Geula S, ... Darnell RB (2017). m⁶A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify

cytoplasmic turnover. *Genes & Development*, 31(10), 990–1006. 10.1101/gad.301036.117 [PubMed: 28637692]

- Linder B, Grozhik AV, Olarerin-George AO, Meydan C, Mason CE, & Jaffrey SR (2015). Single-nucleotide-resolution mapping of m⁶A and m⁶Am throughout the transcriptome. *Nature Methods*, 12(8), 767–772. 10.1038/nmeth.3453 [PubMed: 26121403]
- Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, ... Jaffrey SR (2017). Reversible methylation of m⁶Amin the 5' cap controls mRNA stability. *Nature*, 541(7637), 371–375. 10.1038/nature21022 [PubMed: 28002401]
- Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, ... Jaffrey SR (2015). 5' UTR m⁶A Promotes Cap-Independent Translation. *Cell*, 163(4), 999–1010. 10.1016/j.cell.2015.10.012 [PubMed: 26593424]
- Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, & Jaffrey SR (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell*, 149(7), 1635–1646. 10.1016/j.cell.2012.05.003 [PubMed: 22608085]
- Olarerin-George AO, & Jaffrey SR (2017). MetaPlotR: a Perl/R pipeline for plotting metagenes of nucleotide modifications and other transcriptomic sites. *Bioinformatics*, 33(10), 1563–1564. Retrieved from 10.1093/bioinformatics/btx002 [PubMed: 28158328]
- Perry RP, Kelley DE, Friderici K, & Rottman F (1975). The methylated constituents of L cell messenger RNA: Evidence for an unusual cluster at the 5' terminus. *Cell*, 4(4), 387–394. [PubMed: 1168101]
- Schwartz S, Mumbach MR, Jovanovic M, Wang T, Maciag K, Bushkin GG, ... Regev A (2014). Perturbation of m⁶A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Reports*, 8(1), 284–296. 10.1016/j.celrep.2014.05.048 [PubMed: 24981863]
- Sommer S, Lavi U, & Darnell JE (1978). The absolute frequency of labeled N-6-methyladenosine in HeLa cell messenger RNA decreases with label time. *Journal of Molecular Biology*, 124(3), 487–499. [PubMed: 712844]
- Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, ... He C (2015). N6-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell*, 161(6), 1388–1399. 10.1016/j.cell.2015.05.014 [PubMed: 26046440]
- Webb S, Hector RD, Kudla G, & Granneman S (2014). PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. *Genome Biology*, 15(1), R8 10.1186/gb-2014-15-1-r8 [PubMed: 24393166]
- Xiao W, Adhikari S, Dahal U, Chen Y-S, Hao Y-J, Sun B-F, ... Yang Y-G (2016). Nuclear m⁶A Reader YTHDC1 Regulates mRNA Splicing. *Molecular Cell*, 61(4), 507–519. 10.1016/j.molcel.2016.01.012 [PubMed: 26876937]

Figure 1.

Schematic of the miCLIP protocol. Capped (grey circle) and polyadenylated cellular RNA containing both m⁶Am (red triangle) and m⁶A (red circle) is fragmented and incubated with an anti-m⁶A antibody. Following UV-crosslinking, the antibody-RNA complexes are recovered using protein A/G-affinity beads and a 3'-adapter is ligated to the RNA. Antibody-RNA complexes are then purified by transferring to a nitrocellulose membrane and eluted using proteinase K, leaving only a small peptide fragment crosslinked at the m⁶A/m⁶Am site. RNA fragments are reverse transcribed, which results in mutations or

truncations at the crosslink site in the resulting cDNA. Finally, the cDNA is circularized, re-linearized and amplified by PCR to generate final libraries for sequencing.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2. Sample miCLIP autoradiogram. Here, fragmented mRNA was crosslinked to an anti-m⁶A antibody and radiolabeled. The crosslinked RNA-protein complexes were then reduced and purified with SDS-PAGE and nitrocellulose membrane transfer. Crosslinked RNA fragments are found as smears extending from the light and heavy chains of the reduced antibody, at 25 kD and 50 kD, respectively. This autoradiogram is then used as a template to cut out the pieces of membrane containing the RNA fragments (indicated by the red dashed lines) for elution.

Figure 3.

Example of the test PCR reactions to determine the number of cycles used in the preparative PCR reaction. Here, the high and medium fractions of a cDNA library was amplified for 15, 20, or 25 cycles, run on a 6% TBE gel, and visualized with SYBR Gold on a BioRad ChemiDock. Here 15 cycles produced sufficient libraries, while 25 cycles lead to overamplified products. In this experiment, 13 cycles was chosen to amplify the final library using 8 μ l cDNA.

Table 1.

List of software packages and files needed for miCLIP data processing. Version numbers given are the earliest versions for each package we have utilized or are currently utilizing.

Program/package	Version	Use	Availability
Flexbar	2.5	Trimming of adaptors and low-quality nucleotides (both forward and reverse reads).	https://github.com/seqan/flexbar
pyCRAC	1.1.3	For sample demultiplexing and fast, initial PCR duplicate removal. pyCRAC moves the random barcode into the read header, which is useful for our downstream purposes.	https://bitbucket.org/sgrann/pycrac
bwa	0.7.17	Alignment to the genome. We originally used novoalign however bwa is now recommended. Other aligners such as bowtie can also be used.	https://github.com/lh3/bwa
Samtools	1.8	Requirement for many programs; for processing of alignment (SAM/BAM) files.	http://www.htslib.org/
Bedtools	2.27.1	Requirement for many programs; for processing of read/feature coordinate (BED) files.	https://bedtools.readthedocs.io/en/latest/
CTK	1.1.3	Parsing of alignment files to extract C→T transitions and determining significant sites. Further filtering of PCR duplicates.	https://zhanglab.c2b2.columbia.edu/index.php/CTK_Documentation
miCLIP tools	1.0	Script required for step 6, to reverse complement the second read of the pair.	https://github.com/jaffreylab/miCLIP_MiMB_2016
MetaplotR	1.0	Takes single-nucleotide feature of interest (e.g. determined m ⁶ A sites) to generate metagenes.	https://github.com/olarerin/metaPlotR
Reference genome		FASTA format reference genome (and index in .fai format) for alignment.	http://hgdownload.cse.ucsc.edu/downloads.html
Transcript annotation		GTF/BED format annotations such as GENCODE to identify putative m ⁶ Am sites at the transcript start sites (TSS)	https://www.gencodegenes.org/
CAGE data		To further confirm an m ⁶ Am is truly at a TSS. Ideally, data should be collected from the cell type and conditions used in miCLIP.	http://fantom.gsc.riken.jp/data/