Published in final edited form as:

Curr Protoc Mol Biol. ; 116: 27.5.1–27.5.15. doi:10.1002/cpmb.21.

Protein Interaction Profile Sequencing (PIP-seq)

Shawn W. Foley^{1,2} and Brian D. Gregory^{1,2}

¹Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania

²Cell and Molecular Biology Graduate Program, University of Pennsylvania, Philadelphia, Pennsylvania

Abstract

Every eukaryotic RNA transcript undergoes extensive post-transcriptional processing from the moment of transcription up through degradation. This regulation is performed by a distinct cohort of RNA-binding proteins which recognize their target transcript by both its primary sequence and secondary structure. Here, we describe protein interaction profile sequencing (PIP-seq), a technique that uses ribonuclease-based footprinting followed by high-throughput sequencing to globally assess both protein-bound RNA sequences and RNA secondary structure. PIP-seq utilizes single- and double-stranded RNA-specific nucleases in the absence of proteins to infer RNA secondary structure. These libraries are also compared to samples that undergo nuclease digestion in the presence of proteins in order to find enriched protein-bound sequences. Combined, these four libraries provide a comprehensive, transcriptome-wide view of RNA secondary structure and RNA protein interaction sites from a single experimental technique.

Keywords

protein interaction profile sequencing (PIP-seq); RNA-protein interactions; RNA secondary structure; structure-specific ribonucleases

INTRODUCTION

Many biological processes are dependent on proper post-transcriptional regulation (Ebert and Sharp, 2012; Braunschweig et al., 2013). All of these post-transcriptional processes require a specific collection of RNA-binding proteins (RBPs). The interaction of RBPs with their target transcripts is regulated by both the primary sequence and secondary structure, or folding pattern, of the RNA molecule. Previous studies have primarily probed RNA-protein interactions via protein-centric approaches. Specifically, these studies utilized RBP-specific immunoprecipitation, followed by high throughput sequencing or microarray analyses on the co-purified RNA (Ule et al., 2003; Licatalosi et al., 2008; Hafner et al., 2010). While many insights have been made by these studies, they are limited to uncovering only the target transcripts or binding sites of a single protein.

We declare no competing financial interests.

Protein interaction profile sequencing (PIP-seq) is a genome-wide methodology to simultaneously identify protein bound RNA sequences as well as RNA secondary structure (Silverman et al., 2014; Gosai et al., 2015). PIP-seq utilizes a ribonuclease (RNase)-mediated protein footprinting approach to identify RBP-bound RNA sequences and RNA secondary structure on a global scale. Briefly, whole tissue or tissue culture cells are first treated with formaldehyde in order to cross-link RNA-protein interactions (Basic Protocol 1). The cross-linked material is then lysed and separated into four samples, two footprinting samples, and two structure-only samples. In the structure-only samples, proteins are first denatured and digested by treatment with SDS and proteinase K. These samples are then treated with an RNase that specifically digests single-stranded RNA (ss-RNase) or double-stranded RNA (dsRNase). Complementarily, the footprinting samples are first treated with the structure-specific RNases prior to protein denaturation and digestion followed by RNA extraction (Basic Protocol 2). Lastly, the four samples undergo strand-specific high-throughput sequencing library preparation (Basic Protocol 3; Figure 27.5.1). After sequencing, RNA secondary structure can be inferred by comparing the coverage of each structure-only sample. Additionally, protein-bound sequences can be identified by searching for sequences enriched in the footprinting samples as compared to the structure-only samples. Together, these data produce a simultaneous view of the global landscapes of both RNA secondary structure and RNA-protein interactions.

BASIC PROTOCOL 1: FORMALDEHYDE CROSS-LINKING OF TISSUE

In order to identify protein bound RNA sequences, the RNA-protein interactions must be preserved after cell lysis. To accomplish this, the tissue is treated with a 1% formaldehyde solution under a vacuum to cross-link these *in vivo* interactions. The cross-linking reaction is then quenched by vacuum infiltrating 125 mM glycine into the cells, followed by washing with distilled, deionized water, prior to performing protein interaction profiling. Cross-linking of culture cells can be accomplished as previously described (Silverman and Gregory, 2015).

Materials

- Tissue sample
- Cross-linking buffer (see recipe)
- 1 M Glycine (Sigma, cat. no. G8790)
- Distilled water
- Liquid N₂
- 50 ml conical tubes
- End-over-end rotator
- Vacuum desiccator (Nalgene, cat. no. 5310-0250)
- Vacuum line
- Paper towels

1. Transfer the tissue to a 50-ml conical tube containing 30 ml of cross-linking buffer and let rotate end-over-end for 1 min.

We have previously used 1 to 3 g of plant tissue in 30 ml of cross-linking buffer. Be sure that the buffer is in excess, and the formaldehyde solution can properly penetrate the tissue.
2. Remove the lid and place the conical tube in the bell-shaped vacuum container connected to a vacuum line, and let incubate under vacuum for 10 min.
3. Quench the reaction by adding 4.25 ml of 1 M Glycine to a final concentration of 125 mM.
4. Mix the quenched solution thoroughly by inversion, remove the cap, and place under vacuum for 5 min.
5. Wash the tissue 5 times, each time with ~30 ml distilled, deionized water.
6. Pat the tissue dry with a paper towel, then transfer the tissue into a new 50-ml tube, and snap freeze by submersion in liquid N₂.
7. Store the tissue up to 6 months at –80°C.

BASIC PROTOCOL 2: PROTEIN INTERACTION PROFILING

Formaldehyde cross-linked tissue (Basic Protocol 1) is then homogenized, lysed, and treated with both RNases and proteases during protein interaction profiling. The two structure only libraries first have proteins denatured and digested with an SDS and proteinase K solution, followed by digestion with either an ssRNase or dsRNase. The coverage across these two libraries will later be compared to infer secondary structure. Additionally, two footprinting samples will be prepared. These samples first undergo RNase digestion prior to protease treatment. These proteins will inhibit RNA digestion creating protein protected sites (PPSs), sequences that are enriched in the footprinting sample when compared to the structure only samples.

Materials

Formaldehyde cross-linked tissue (Basic Protocol 1)
Liquid N₂
RIP buffer (see recipe)
DNase solution (see recipe)
3 M NaOAc (pH = 5.5) (Life Technologies, cat. no. AM9740)
100% EtOH (Decon Labs, cat. no. 2716)
70% EtOH
RNase ONE buffer (Promega, cat. no. M4261)
RNase ONE (Promega, cat. no. M4261)

RNase stop buffer, 16× (see recipe)
QIAzol (Qiagen, cat. no. 79306)
10× RNA Structure Buffer (Life Technologies, cat. no. AM2275)
RNase V1 (Life Technologies, cat. no. AM2275)
Mortar and pestle
1.7-ml tubes
Plastic pestle
1000- and 200-μl pipet tips
2.0-ml tubes
Centrifuge
37°C heat block
65°C heat block
Vortex mixer

Homogenize and DNase treat tissue prior to library preparation

1. Precool mortar and pestle with liquid nitrogen, then homogenize the tissue with a mortar and pestle.
2. Resuspend the tissue in 850 μl RIP buffer and transfer to a precooled 1.7-ml tube
3. Grind with plastic pestle, and then pass through 1000-μl and 200-μl pipet tips ~20 times each.
4. Add 160 μl DNase solution and incubate for 30 min at room temperature.
5. Transfer 250 μl of lysate and 600 μl RIP into two 2.0-ml tubes for structure only library samples and two 1.7-ml tubes for footprinting library samples.

To prepare ssRNase-treated structure only library

- 6 To one of the 850 μl samples in 2.0-ml tubes, add 75 μl of 16× RNase stop buffer and incubate for 15 min at room temperature.
- 7 Precipitate RNA by adding 90 μl of 3 M NaOAc and 1 ml of 100% EtOH, and then store for 1 to 2 hr at −80°C.
- 8 Centrifuge for 45 min at 20,000 × *g*, 4°C.
- 9 Remove the supernatant and add 700 μl of 70% EtOH to wash.
- 10 Centrifuge for 5 min 20,000 × *g*, 4°C.
- 11 Remove the supernatant and let the pellet air dry for 10 min.
- 12 Resuspend in 850 μl RIP buffer.

The pellet will be difficult to resuspend, pipet and vortex thoroughly.

- 13 Add 100 μ l RNase ONE buffer and 10 μ l RNase ONE.
- 14 Incubate for 1 hr and 15 min at 37°C, mixing ever 15 min.
- 15 Add 75 μ l of 16 $\times$ RNase stop buffer and mix thoroughly.
- 16 Incubate for 10 min at room temperature.
- 17 Transfer the samples to 65°C for 2 hr, mixing every 30 min.
- 18 Divide the sample into two 2.0-ml tubes with ~500 μ l each.
- 19 Add 700 μ l QIAzol to each tube, vortex thoroughly for 5 sec, and store at –20°C.

The samples can remain overnight at –20°C.

To prepare dsRNase-treated structure only library

- 20 To the other 850 μ l sample in 2.0-ml tubes, add 75 μ l of 16 $\times$ RNase stop buffer and incubate for 15 min at room temperature.
- 21 Precipitate RNA by adding 90 μ l of 3 M NaOAc, 1 ml 100% EtOH, and store for 1 to 2 hr at –80°C.
- 22 Centrifuge for 45 min at 20,000 $\times g$, 4°C.
- 23 Remove the supernatant and add 700 μ l of 70% EtOH to wash.
- 24 Centrifuge for 5 min at 20,000 $\times g$, 4°C.
- 25 Remove the supernatant and let the pellet air dry for 10 min.
- 26 Resuspend the pellet in 850 μ l RIP buffer.

The pellet will be difficult to resuspend, pipet, and vortex thoroughly.

- 27 Add 100 μ l of 10 $\times$ RNA Structure Buffer and 25 μ l RNase V1.
- 28 Incubate at 1 hr 15 min at room temperature, mixing ever 15 min.
- 29 Add 75 μ l of 16 $\times$ RNase stop buffer and mix throughly.
- 30 Incubate for 10 min at room temperature.
- 31 Transfer samples to 65°C for 2 hr, mixing every 30 min.
- 32 Divide the sample into two 2.0-ml tubes with ~500 μ l each.
- 33 Add 700 μ l QIAzol to each tube, vortex thoroughly for 5 sec, and store at –20°C.

The samples can remain overnight at –20°C.

To prepare ssRNase-treated footprinting library

- 34 To one of the 850 μ l samples in 1.7-ml tubes, add 100 μ l RNase ONE buffer and 10 μ l RNase ONE.
- 35 Incubate for 1 hr at 37°C, mixing ever 15 min.

- 36 Add 75 μ l of 16 $\times$ RNase stop buffer and mix thoroughly.
- 37 Incubate for 10 min at room temperature.
- 38 Transfer samples to 65°C for 2 hr, mixing every 30 min.
- 39 Divide the sample into two 2.0-ml tubes with $\sim$ 500 μ l each.
- 40 Add 700 μ l QIAzol to each tube, vortex thoroughly for 5 sec, and store at -20°C .

The samples can remain overnight at -20°C .

To prepare dsRNase treated structure-only library

- 41 To the other 850 μ l sample in 1.7-ml tubes, add 100 μ l of 10 $\times$ RNA Structure Buffer and 25 μ l RNase V1.
- 42 Incubate for 1 hr at room temperature, mixing every 15 min.
- 43 Add 75 μ l of 16 $\times$ RNase stop buffer and mix thoroughly.
- 44 Incubate for 10 min at room temperature.
- 45 Transfer the samples to 65°C for 2 hr, mixing every 30 min.
- 46 Divide the sample into two 2.0-ml tubes with $\sim$ 500 μ l each.
- 47 Add 700 μ l QIAzol to each tube, vortex thoroughly for 5 sec, and store at -20°C .

The samples can remain overnight at -20°C .

- 48 Perform Qiagen miRNeasy RNA extraction following the manufacturer's protocol.

BASIC PROTOCOL 3: STRAND-SPECIFIC HIGH-THROUGHPUT SEQUENCING LIBRARY PREPARATION

This protocol describes the process of producing a strand-specific high-throughput sequencing library using the RNA isolated at the end of Basic Protocol 2. This is a modified version of the Illumina strand-specific high-throughput library preparation protocol, which includes several size selection steps that are necessary for removal of adapter duplexes. In order to remove highly abundant RNA sequences such as rRNA and tRNA this protocol utilizes a Duplex Specific Nuclease (DSN) treatment. This treatment first denatures the cloned DNA at 98°C, and then allows the DNA molecules to slowly reanneal for 5 hr at 68°C. The high temperature allows slow reannealing, resulting in the most highly abundant transcripts reannealing more quickly than those that are less abundant. The double-stranded DNA sequences are then digested by the double-stranded nuclease (DSN), dramatically reducing the rRNA and tRNA content.

Materials

Protein interaction profile RNA (Basic Protocol 2)

10× Fragmentation Reagent (Life Technologies, cat. no. AM8740)
Fragmentation Stop Solution (Life Technologies, cat. no. AM8740)
DEPC-treated H₂O
5 mg/ml Glycogen, ultrapure (Life Technologies, cat. no. AM9510)
3 M NaOAc (pH = 5.5) (Life Technologies, cat. no. AM9740)
100% EtOH (Decon Labs, cat. no. 2716)
80% EtOH
Ice
T4 DNA ligase buffer (NEB, cat. no. B0202S)
T4 polynucleotide kinase (NEB, cat. no. M0201S)
10 mM ATP (Life Technologies, cat. no. AM8110G)
10× TBE (Bio-Rad, 161-0733)
Gel loading buffer II (Life Technologies, cat. no. AM8546G)
10-bp DNA ladder (Life Technologies, cat. no. 10821-015)
10 mg/ml Ethidium Bromide (Life Technologies, cat. no. 15585-011)
0.3 M NaCl
5 μM 3' Adapter (RA3) (TGGAATTCTCGGGTGCCAAGG)
RNA Ligase Buffer (NEB, cat. no. B0216L)
RNaseOUT (Life Technologies, cat. no. 10777019)
200 U/μl Epicenter T4 RNA Ligase 2, truncated (NEB, cat. no. M0242S)
25 μM 5' Adapter (RA5) (GUUCAGAGUUCUACAGUCCGACGAUC)
T4 RNA Ligase 1 (NEB, cat. no. M0204S)
RNA RT Primer (RTP) (GCCTTGGCACCCGAGAATTCCA)
5× First-Strand Buffer (Life Technologies, cat. no. 18064-014)
50 mM dNTPs
100 mM DTT (Life Technologies, cat. no. 18064-014)
SuperScript II Reverse Transcriptase (Life Technologies, cat. no. 18064-014)
2× Phusion Mix (NEB, cat. no. M0531S)
5 mM Betaine (MP Biomedicals, cat. no. 215046180)
10 μM RNA PCR Primer (RTP) (GCCTTGGCACCCGAGAATTCCA)
10 μM RNA PCR Primer Index
Nuclease-free water

25-bp DNA ladder (Life Technologies, cat. no. 10597-011)
DSN Hybridization buffer (see recipe)
10× DSN Master Mix (Evrogen, cat. no. EA001)
DSN Enzyme (Evrogen, cat. no. EA001)
DSN STOP solution (Evrogen, cat. no. EA001)
1.7-ml tubes
70°C heat block
Centrifuge
37°C heat block
1.0-mm, 10-well 15% TBE-Urea gels (Life Technologies, cat. no. EC6885BOX)
Gel box for running prepoured gels (Life, Technologies, cat. no. EI0001)
18-G needles
Razor blades
Gel Breaker Tubes (IST Engineering, cat. no. 388-100)
2-ml tubes
End-over-end rotator
Spin-X columns (Costar, cat. no. 8160)
200- μ l PCR tubes
Thermal cycler
6% TBE gels 1.0 mm, 10 well (Invitrogen, cat. no. EC6265BOX)

Perform RNA fragmentation

1. Add 1 μ l of 10× Fragmentation Reagent to 9 μ l RNA in a 1.7-ml tube.
2. Incubate for 3 min at 70°C.
3. Immediately add 1 μ l Fragmentation Stop Solution
4. Bring volume up to 100 μ l with DEPC-treated H₂O.
5. Precipitate RNA by adding 3 μ l Glycogen, 1/10 volume 3 M NaOAc (pH = 5.5), and 3× volume 100% EtOH.
6. Store at –80°C for at least 2 hr.
7. Centrifuge the samples for 80 min at 20,000 $\times g$, 4°C.
8. Wash samples with 700 μ l 80% EtOH.
9. Remove the supernatant and let the samples air dry for 10 min.
10. Resuspend the samples in 16 μ l DEPC-treated H₂O.

11. Incubate for 25 min on ice.

Perform T4 polynucleotide kinase treatment

12. Add 2 μ l NEB T4 DNA ligase buffer, 1 μ l T4 polynucleotide kinase ligase, 1 μ l of 10 mM ATP to 16 μ l RNA.
13. Incubate for 1 hr at 37°C for 1 hr.
14. Add 80 μ l DEPC treated H₂O to bring volume to 100 μ l.
15. Precipitate RNA (Basic Protocol 3, steps 5 to 9)
16. Resuspend precipitant in 10 μ l DEPC-treated H₂O.
17. Incubate for 25 min on ice.

Perform size selection no. 1

18. Assemble 15% TBE-Urea polyacrylamide gel in a gel box with ~500 ml 1 $\times$ TBE.
19. Preclear the wells with an 18-G needle.
20. Prerun the gel for 25 min at 155 V.
21. Add 10 μ l Gel Loading Buffer II to the RNA samples.
22. Heat the samples for 5 min at 70°C, and then snap cool on ice for 3 min.
23. Prepare 1.5 μ g 10-bp DNA ladder in 10 μ l DEPC-treated H₂O and 10 μ l Gel Loading Buffer II.
24. Preclear the wells again with 18-G needle.
25. Load the samples into the gel and run for 90 min at 155 V or until bromphenol blue has traveled 80% of the gel.
26. Stain the gel in 100 ml of 1 $\times$ TBE and 14 μ l of 10 mg/ml Ethidium Bromide for 10 min at room temperature.
27. Using an ethanol-cleaned razor blade, excise the gel region from 15 to 150 nucleotides for each sample on the gel.

Make sure to use clean razor blades for each sample to avoid contamination.
28. Place the gel slices in Gel Breaker Tube inside of a 2.0-ml tube, and centrifuge the samples for 2 min at 20,000 $\times g$, room temperature.
29. Add 300 μ l of 0.3 M NaCl and rotate end-over-end for 4 hr.
30. Transfer the entire sample to a Spin-X column and centrifuge the samples for 2 min at 20,000 $\times g$, room temperature.
31. Transfer the flowthrough to a new 1.7-ml tube and precipitate RNA (see steps 5 to 9).

- 32 Resuspend the RNA in 5 μ l DEPC-treated H₂O
- 33 Allow the samples to resuspend on ice for 25 min.

Ligate 3' and 5' adapters

- 34 Transfer the 5 μ l RNA samples to a 200 μ l PCR tube and add 1 μ l of 5 μ M 3' Adapter (RA3).
- 35 Mix thoroughly, and incubate for 2 min at 70°C for 2 min, then for 2 min 4°C.
- 36 Add 4 μ l of the following ligation reagents and vortex for 5 sec:
- 2 μ l RNA Ligase Buffer
 - 1 μ l RNaseOUT
 - 1 μ l 200 U/ μ l Epicenter T4 RNA Ligase 2, truncated.
- 37 Incubate for 1 hr 15 min at 28°C.
- 38 Near the end of the incubation, transfer 1 μ l of 25 μ M 5' adapter (RA5) to 70°C for 2 min and snap cool on ice 2 min.
- The total amount of 5' adapter needed for all ligations can be denatured in a single tube
- 39 Add 1 μ l of 10 mM ATP and 1 μ l T4 RNA Ligase 1 to the tube containing the cooled 5' adapter.
- This can be made into a master mix with the denatured 5' adapter for all samples to avoid pipetting errors.
- 40 Transfer 3 μ l of the 5' adapter mix to each RNA sample and incubate for 1 hr at 28°C.
- Store the samples up to 1 week at -20°C, or proceed to Size Selection no. 2.

Perform size selection no. 2

- 41 Prepare 15% TBE-Urea polyacrylamide gel (see steps 18 to 20)
- 42 Add 10 μ l Gel Loading Buffer II to the RNA ligase reaction and incubate for 5 min at 70°C, and then snap cool on ice for 3 min.
- 43 Run the samples on the 15% TBE-Urea polyacrylamide gel and stain with ethidium bromide (see steps 23 to 26).
- 44 Using an ethanol-cleaned razor for each sample on the gel, excise the gel from 65 to 300 nucleotides.
- Make sure to use clean razor blades for each sample to avoid contamination.
- 45 Elute and precipitate the RNA from the gel slice (see steps 27 to 31).
- 46 Resuspend the RNA in 6 μ l DEPC-treated H₂O.

- 47 Incubate for 25 min on ice.

Perform reverse transcription

- 48 Preheat a thermal cycler to 70°C.
- 49 Transfer 6 µl of the RNA samples to separate wells in a strip 200-µl PCR tubes.
- 50 Add 1 µl of 100 µM RT Primer (RTP) and vortex for 5 sec.
- 51 Incubate the samples for 2 min at 70°C, and then for 2 min 4°C.
- 52 Add 5.5 µl of the following RT reagents and mix thoroughly:
- 2 µl 5 × First-Strand Buffer
 - 0.5 µl 50 mM dNTP mix
 - 1 µl 100 mM DTT
 - 1 µl RNaseOUT
 - 1 µl Superscript II Reverse Transcriptase.
- 53 Incubate the RT reaction for 1 hr at 50°C.
- 54 Proceed directly to PCR Amplification.

Perform PCR amplification

- 55 Prepare the following PCR master mix:
- 50 µl Phusion Mix 2×
 - 33.5 µl 5 mM Betaine
 - 2 µl RNA PCR Primer (RP1).
- 56 Add 85.5 µl PCR master mix to RT reaction
- 57 Add 2 µl of RNA PCR Primer Index to each sample.
- Each sample should have a different RNA PCR Primer Index allowing multiplexing during sequencing.
- 58 Divide each 100 µl sample between four 200-µl PCR tubes, ~25 µl per tube.
- 59 Incubate the samples on a thermal cycler as follows:
1 cycle:	30 sec	98°C (initial denaturation)
12 cycles:	10 sec	98°C (denaturation)
	30 sec	60°C (annealing)
	30 sec	72°C (extension)
1 cycle:	5 min	72°C (final extension)
Final step:	Indefinite	4°C (hold).
- 60 Recombine the four PCR tubes.

- 61 Precipitate DNA by adding 3 μ l Glycogen, 1/10 volume 3 M NaOAc (pH = 5.5), and 3 $\times$ volume of 100% EtOH.
- 62 Store for at least 2 hr at -80°C .
- 63 Centrifuge the samples for 45 min at $20,000 \times g$, 4°C .
- 64 Wash the samples with 700 μ l of 70% EtOH.
- 65 Centrifuge the samples for 5 min at $20,000 \times g$, 4°C .
- 66 Remove the supernatant and let the samples air dry for 10 min.
- 67 Resuspend in 10 μ l nuclease-free H_2O .
- 68 Incubate for 25 min on ice.

Perform size selection no. 3

- 69 Assemble 6% TBE-Urea polyacrylamide gel in a gel box with ~ 500 ml $1\times$ TBE.
- 70 Add 10 μ l Gel Loading Buffer II to the DNA samples.
- 71 Prepare 1.5 μ g 25-bp DNA ladder in 8.5 μ l DEPC-treated H_2O and 10 μ l Gel Loading Buffer II
- 72 Preclear the wells with an 18-G needle.
- 73 Load the samples into the gel and run for 30 min at 155 V or until bromphenol blue has traveled 80% of the gel.
- 74 Stain the gel in 100 ml of $1\times$ TBE and 14 μ l of 10 mg/ml Ethidium Bromide for 10 min at room temperature.
- 75 Using an ethanol-cleaned razor for each sample on the gel, excise gel from 135 bp to the end of the signal, being careful to avoid the adapter-adapter artifacts at 120 bp.

Make sure to use clean razor blades for each sample to avoid contamination.
- 76 Place the gel slices in the Gel Breaker Tube inside of a 2.0-ml tube, and centrifuge the samples for 2 min at $20,000 \times g$, room temperature.
- 77 Add 300 μ l of $1\times$ NEB Buffer 2 and rotate end-over-end for 2 hr.
- 78 Transfer the entire sample to a Spin-X column and centrifuge the samples for 2 min at $20,000 \times g$, room temperature.
- 79 Transfer flow-through to a new tube and precipitate DNA (see steps 61 to 66)
- 80 Resuspend the precipitated DNA in 15.5 μ l nuclease-free H_2O
- 81 Allow the samples to resuspend for 25 min on ice.

Duplex-specific nuclease (DSN) treatment

- 82** Put a total of 100 ng of sample DNA into 13.5 µl nuclease-free H₂O in a 200-µl PCR tube.

Store the remaining library up to 1 year at –80°C. The DSN treatment can be repeated if more library is necessary for sequencing.

- 83** Add 4.5 µl DSN Hybridization buffer, and vortex for 5 sec.

- 84** Incubate the samples in the thermal cycler as follows:

2 min 98°C
5 hr 68°C

The 5-hr incubation has been found to reduce rRNA and tRNA in Arabidopsis. For other species, a range of incubation times should be tested.

- 85** After 4.5 hr, dilute 4 µl of 10× DSN Master buffer in 16 µl nuclease-free H₂O to make 2× DSN Master buffer.

- 86** Incubate 2× DSN Master buffer at 68°C for the remainder of the 5 hr incubation.

- 87** Add 20 µl 2× DSN Master buffer to each sample, and mix thoroughly by pipetting.

Do not allow samples to cool down; mix by pipetting and briefly centrifuge.

- 88** Incubate the samples for another 10 min at 68°C.

- 89** Add 2 µl DSN enzyme to the DNA samples, and mix thoroughly.

Do not allow samples to cool down; mix by pipetting and briefly centrifuge.

- 90** Incubate the samples for another 25 min at 68°C.

- 91** Add 40 µl DSN STOP solution, and mix thoroughly.

- 92** Transfer the samples to 1.7-ml tubes and add 20 µl nuclease-free H₂O, bringing the volume up to 100 µl.

- 93** Precipitate the DNA (see steps 61 to 66)

- 94** Resuspend the DNA in 6 µl nuclease-free H₂O

- 95** Allow the samples to resuspend for 25 min on ice.

PCR amplification no. 2

- 96** Prepare the following PCR master mix:

50 µl Phusion Mix 2×
40 µl 5 mM Betaine
2 µl RNA PCR Primer.

- 97** Add 92 µl PCR master mix to the PCR reaction

- 98** Add 2 µl of RNA PCR Primer Index to each sample.

Be sure that this is the same PCR Primer Index used in step 57.

99 Divide each 100 µl sample between four 200-µl PCR tubes, ~25 µl per tube.

100 Incubate the samples on the thermal cycler as follows:

1 cycle:	30 sec	98°C (initial denaturation)
12 cycles:	10 sec	98°C (denaturation)
	30 sec	60°C (annealing)
	15 sec	72°C (extension)
1 cycle:	10 min	72°C (final extension)
Final step:	Indefinite	4°C (hold).

101 Recombine the four PCR tubes.

102 Precipitate DNA (see steps 61 to 66)

103 Resuspend in 10 µl nuclease-free H₂O.

104 Incubate for 25 min on ice.

Perform size selection no. 4

105 Run the samples on 6% TBE-Urea polyacrylamide gel and excise bands greater than 135 bp as in steps 69 to 79.

106 Resuspend the DNA in 12 µl nuclease-free H₂O

107 Allow the samples to resuspend for 25 min on ice.

You should now have a completed set of PIP-seq libraries for high-throughput sequencing that allow the simultaneous analysis of RNA secondary structure and RNA-protein interactions.

REAGENTS AND SOLUTIONS

Use deionized, distilled water in all recipes and protocol steps.

Cross-linking buffer

810 µl of 37% Formaldehyde (1% final) (Sigma, cat. no. 252549)

6 µl SilWet L-77 (0.02% final) (Lehle Seeds, cat. no. VIS-30)

Distilled, deionized H₂O, up to 30 ml

Prepare fresh immediately before use

DNase solution, 1x

550 µl DEPC-treated H₂O

3.85 ml RDD buffer (Qiagen, cat. no. 79254)

Use solution to resuspend lyophilized RNase free DNase (Qiagen, cat. no. 79254)

Divide into 80-µl aliquots

Store up to 1 year at -20°C

DSN Hybridization buffer

200 μl of 1 M HEPES buffer solution (200 mM final) (Life Technologies, cat. no. 15630)

400 μl of 5 M NaCl (2 M final) (Life Technologies, cat. no. AM9937)

400 μl nuclease-free H_2O

Divide into 500- μl aliquots

Store up to 1 year at -20°C

DSN STOP buffer

10 mM EDTA

RIP buffer

12.5 ml of 1 M Tris-Cl, pH 8.6 (25 mM final)

37.5 ml of 2 M KCl (150 mM final)

5 ml of 0.5 M EDTA, pH 8.0 (5 mM final) (Fisher Scientific, cat. no. BP2483)

2.5 ml Igepal CA-630 (0.5% final) (Sigma, cat. no. I8896)

Distilled, deionized H_2O , up to 500 ml

Store at up to 1 year 4°C

Prepare 10 ml fresh immediately before use by adding:

1 tablet Protease Inhibitor Mini Cocktail (Roche, cat. no. 11836170001)

5 μl RNaseOUT (Life Technologies, cat. no. 10777019)

50 μl 0.1 M DTT (Life Technologies, cat. no. 18064-014)

RNase stop buffer, 16×

5 ml of 20% SDS (10% final) (Denville Sciences, cat. no. CS5585-28)

100 μl of 0.5 M EDTA, pH 8.0 (5 mM final) (Fisher Scientific, cat. no. BP2483)

1 μl of 20 mg/ml Proteinase K (Denville Sciences, cat. no. CB3210-5)

Distilled, deionized H_2O , up to 10 ml

Prepare fresh immediately before use

COMMENTARY

Background Information

The study of post-transcriptional regulation is limited only by the techniques currently available. While the past decade has seen the development of numerous approaches

involving single protein immunoprecipitation followed by sequencing (Ule et al., 2003), more recent protocols have attempted to identify RNA sequences bound globally throughout a sample. In addition to PIP-seq, another approach is global photoactivatable ribonucleoside enhanced cross-linking and immunoprecipitation (gPAR-CLIP). This technique utilizes the ribonucleoside 4-thiouridine, which forms covalent bonds with adjacent proteins when exposed to 365 nm UV light. This covalent bond causes misincorporation of guanine by reverse transcriptase, resulting in uracil-to-cytosine mismatches in the sequencing reads at protein bound sites (Baltz et al., 2012). This technique requires both the uptake and incorporation of 4-thiouridine and UV cross-linking treatment. Therefore, it is primarily applicable in a tissue culture system. Alternatively, PIP-seq utilizes a formaldehyde cross-linking step, allowing it to be used on whole tissues.

Critical Parameters and Troubleshooting

The PIP-seq protocol is dependent on complete digestion by structure-specific RNases. Unlike previous protocols that use low levels of RNase V1 to study single-hit kinetics (Wan et al., 2014), PIP-seq fully digests the double- and single-stranded RNA in each corresponding sample. Therefore, the relatively high concentration of these RNases, and the long incubations are necessary to best gauge secondary structure and protein binding. Additionally, the high levels of SDS in the 16× RNase stop buffer are necessary to fully denature the RBPs in both samples.

After preparing high-throughput sequencing libraries two common concerns are the presence of unwanted RNA populations (rRNA and tRNA), as well as adapter duplexes and other sequencing artifacts. In order to reduce highly abundant RNA classes this protocol utilizes a DSN library treatment step, and recommends a 5 hr 68°C reannealing time, which has been used for plant tissue samples (Gosai et al., 2015). The reannealing time must be determined empirically for different samples, which we have found can range from 4 to 9 hr. Following sequencing library completion, a small aliquot of library can be taken and used for TOPO cloning and Sanger sequencing in order to estimate the percentage of clones containing unwanted RNA populations. Additionally, a higher quantity of low quality sequencing reads can be attributed to adapter duplex artifacts that accumulate during the cloning process. Therefore, it is crucial to remove any unwanted bands during the size selection steps of this protocol (Fig. 27.5.2).

Anticipated Results

Agilent Bioanalyzer 2100 analysis of RNA from protein interaction profiling should reveal RNA fragments from ~15 nt to ~1000 nt. The footprinting samples will contain larger fragments than the structure only samples (Fig. 27.5.3). This is due to proteins protecting longer regions of RNA from RNase digestion, and indicates high quality samples ready for library preparation.

Time Considerations

After harvesting tissue the formaldehyde cross-linking, protein interaction profiling, and RNA extraction will take 1 to 2 days. The formaldehyde cross-linking step takes only ~45 min, while protein interaction profiling protocol will last ~11 hr. Although both of these

protocols can be completed in one day, the samples can be stored at the end of each protocol. The strand-specific library preparation can takes 4 to 8 days. The samples can be stored in EtOH at -80°C during each precipitation step. Therefore, every precipitation is a stopping point.

Acknowledgments

The authors thank members of the Gregory laboratory for helpful discussions. This work was supported by an NSF Career Award MCB-1053846 and NSF Grant MCB-1243947 to BDG, and NIGMS T32GM008216-29 to SWF.

Literature Cited

- Baltz AG, Munschauer M, Schwanhauser B, Vasile A, Murakawa Y, Schueler M, Youngs N, Penfold-Brown D, Drew K, Milek M, Wyler E, Bonneau R, Selbach M, Dieterich C, Landthaler M. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell*. 2012; 46:674–690. DOI: 10.1016/j.molcel.2012.05.021 [PubMed: 22681889]
- Braunschweig U, Gueroussov S, Plocik AM, Graveley BR, Blencowe BJ. Dynamic integration of splicing within gene regulatory pathways. *Cell*. 2013; 152:1252–1269. DOI: 10.1016/j.cell.2013.02.034 [PubMed: 23498935]
- Ebert MS, Sharp PA. Roles for MicroRNAs in conferring robustness to biological processes. *Cell*. 2012; 149:515–524. DOI: 10.1016/j.cell.2012.04.005 [PubMed: 22541426]
- Gosai SJ, Foley SW, Wang D, Silverman IM, Selamoglu N, Nelson ADL, Beilstein MA, Daldal F, Deal RB, Gregory BD. Global analysis of the RNA-protein interaction and RNA secondary structure landscapes of the arabidopsis nucleus. *Mol Cell*. 2015; 57:376–388. DOI: 10.1016/j.molcel.2014.12.004 [PubMed: 25557549]
- Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*. 2010; 141:129–141. DOI: 10.1016/j.cell.2010.03.009 [PubMed: 20371350]
- Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, Darnell JC, Darnell RB. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*. 2008; 456:464–469. DOI: 10.1038/nature07488 [PubMed: 18978773]
- Silverman IM, Gregory BD. Transcriptome-wide ribonuclease-mediated protein footprinting to identify RNA-protein interaction sites. *Methods*. 2015; 72:76–85. DOI: 10.1016/j.ymeth.2014.10.021 [PubMed: 25448484]
- Silverman IM, Li F, Alexander A, Goff L, Trapnell C, Rinn JL, Gregory BD. RNase-mediated protein footprint sequencing reveals protein-binding sites throughout the human transcriptome. *Genome Biol*. 2014; 15:R3.doi: 10.1186/gb-2014-15-1-r3 [PubMed: 24393486]
- Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB. CLIP identifies nova-regulated RNA networks in the brain. *Science*. 2003; 302:1212–1215. DOI: 10.1126/science.1090095 [PubMed: 14615540]
- Wan Y, Kun Qu K, Zhang QC, Flynn RA, Manor O, Ouyang Z, Zhang J, Spitale RC, Snyder MP, Segal E, Chang HY. Landscape and variation of RNA secondary structure across the human transcriptome. *Nature*. 2014; 505:706–709. DOI: 10.1038/nature12946 [PubMed: 24476892]

Figure 27.5.1.

Overview of PIP-seq protocol. This diagram illustrates all three basic protocols, including formaldehyde cross-linking (Basic Protocol 1), protein interaction profiling (Basic Protocol 2), and strand-specific RNA-seq library preparation (Basic Protocol 3). Although seedlings are illustrated as an example, PIP-seq can be used with a variety of tissue and cell types.

Figure 27.5.2.

Size selection gel from strand-specific library preparation. A gel image taken during Size Selection #4. The DSN-treated library post-amplification is indicated by the asterisk (*) and the unincorporated primers and adapter duplexes are indicated by the arrows.

Figure 27.5.3.

Agilent Bioanalyzer 2100 analysis of RNA from protein interaction profiling. The bioanalyzer traces from a structure only (top) and a footprinting (bottom) sample. The arrow indicates the longer RNA fragments in the footprinting sample.