Video Article

A Rapid High-throughput Method for Mapping Ribonucleoproteins (RNPs) on Human pre-mRNA

Katherine H. Watkins[1], Allan Stewart[1], William G. Fairbrother[1,2]

[1]Department of Molecular and Cellular Biology, Brown University
[2]Center for Computational Molecular Biology, Brown University

Correspondence to: William G. Fairbrother at William_Fairbrother@brown.edu

Abstract

Sequencing RNAs that co-immunoprecipitate (co-IP) with RNA binding proteins has increased our understanding of splicing by demonstrating that binding location often influences function of a splicing factor. However, as with any sampling strategy the chance of identifying an RNA bound to a splicing factor is proportional to its cellular abundance. We have developed a novel in vitro approach for surveying binding specificity on otherwise transient pre-mRNA. This approach utilizes a specifically designed oligonucleotide pool that tiles across introns, exons, splice junctions, or other pre-mRNA. The pool is subjected to some kind of molecular selection. Here, we demonstrate the method by separating the oligonucleotide into a bound and unbound fraction and utilize a two color array strategy to record the enrichment of each oligonucleotide in the bound fraction. The array data generates high-resolution maps with the ability to identify sequence-specific and structural determinates of ribonucleoprotein (RNP) binding on pre-mRNA. A unique advantage to this method is its ability to avoid the sampling bias towards mRNA associated with current IP and SELEX techniques, as the pool is specifically designed and synthesized from pre-mRNA sequence. The flexibility of the oligonucleotide pool is another advantage since the experimenter chooses which regions to study and tile across, tailoring the pool to their individual needs. Using this technique, one can assay the effects of polymorphisms or mutations on binding on a large scale or clone the library into a functional splicing reporter and identify oligonucleotides that are enriched in the included fraction. This novel in vitro high-resolution mapping scheme provides a unique way to study RNP interactions with transient pre-mRNA species, whose low abundance makes them difficult to study with current in vivo techniques.

Protocol

Pool design and oligo recovery

1. The first step is to design the pre-mRNA pool to be studied. This can be done using the UCSC Genome Browser and downloading particular genes, splice junctions, or other areas of interest.
2. Once the windows of interest have been selected, tile across them computationally using the following conditions: read length should be 30 nucleotides with a 10 nucleotide overlap, therefore each oligo is shifted 20 nucleotides from the prior one. *Oligo overlap should be increased in proximity to splice sites to ensure adequate coverage; increasing the overlap from 10 to 20 nucleotides can accomplish this.
3. Flank each 30-nucleotide oligo with universal primer sequences, which will be used to amplify the pool downstream. Tiled oligo orders should be submitted to Agilent, where the sequences will be printed on a custom oligonucleotide microarray.
4. To recover the DNA oligos from the microarray surface begin by placing the array cover slide in an array hybridization chamber and gently pipetting 500 µL of dH$_2$O onto the slide
5. Sandwich the array on top of the cover slide, taking care to avoid air bubbles that can disrupt the process. Be sure to place the array face down so the side with the oligos is touching the surface of the water. A good rule of thumb is "Agilent touches Agilent" meaning that the Agilent label on the array faces the Agilent label on the cover slide
6. Close the hybridization chamber and rotate overnight at 99°C in a hybridization oven
7. The following day, carefully remove the array from the chamber and draw off the water, which now contains the oligos liberated from the array surface. This is the oligo pool
8. Place the pool in a 1.5 mL Eppendorf tube and sonicate the pool at 50% amplitude for 3-5 second pulses
9. Next, amplify the pool by low cycle PCR, using the universal primers with a T7 tag appended to the end. For the first round, denature at 94°C for 1 minute, anneal at 55°C for 20 seconds, and elongate for 1 minute at 72°C. For subsequent rounds, do 10 seconds, 20 seconds, and 10 seconds at each respective temperature, with a final elongation step of 5 minutes at 72°C. Because the number of cycles necessary to amplify the pool depends on the efficiency of oligo recovery, it may be necessary to try multiple cycle numbers. Be careful not to over-amplify the pool, which is signified by a smeared band on an acrylamide gel. PCR amplified samples can be stored at 4°C until needed.
10. The final step in oligo preparation is to transcribe RNA using the MEGAshortscript™ Kit from Ambion. *The following protocol has been adapted from Ambion
    a. Thaw the T7 10x Reaction Buffer, four ribonucleotide solutions and water at room temperature while keeping the T7 Enzyme Mix on ice
    b. Assemble the reaction mixture in an RNase-free microcentrifuge tube at room temperature. Components of the reaction buffer can cause template DNA to precipitate if the reaction is mixed on ice.
    c. In the following order, pipette 3 µL nuclease-free water, 2 µL of the 10x reaction buffer, 2 µL of each 75 mM nTP solution, 5 µL of template DNA (the PCR amplified oligo pool) and 2 µL of T7 enzyme, for a final volume of 20 µL
    d. Flick the tube gently to mix the reaction then briefly microfuge to collect the mixture at the bottom of the tube.
    e. Incubate the reaction in a thermocycler at 37°C for 2 hours. Incubation time will be template-dependent, therefore it may be necessary to use a time-course experiment to determine the optimal incubation time for maximal yield
    f. After 2 hours, remove the template DNA by adding 1 µL of Turbo DNase to the reaction and incubate at 37°C for an additional 15 minutes
    g. Terminate the reaction and recover the RNA by phenol/chloroform extraction and ethanol precipitation.
        a. For a 20 µL reaction volume, add 115 µL of dH$_2$O and 15 µL of 3 M sodium acetate to the reaction.

    b. Mix thoroughly then add 75 µL of acidic phenol and 75 µL of chloroform. Mix the solution by vortexing, then microfuge for one minute at 13,000 rpm at room temperature. Transfer the aqueous layer to a new tube and repeat the extraction

    c. Transfer the aqueous layer to a new tube and add 150 µL of chloroform. Vortex the mixture and centrifuge as before. Transfer the aqueous layer to a new tube.

  h. Recover the RNA by adding two volumes of 100% cold ethanol to the aqueous layer and mixing well. Chill the mixture for a minimum of 15 minutes at -20°C then centrifuge at 4°C for 15 minutes at maximum speed to pellet the RNA. Carefully remove the supernatant and resuspend the pellet in 105 µL of 0.1x Tris-EDTA (TE) Buffer..

11. It is difficult to remove all of the free nucleotides from the T7 Kit reaction, so for the most accurate results, quantify the RNA using RiboGreen.
    a. To begin the quantitation, first dilute enough 1x TE, 2,300 µL + 50µL/sample.
    b. Next, make the RiboGreen. You will need 1000 µL + 50 µL/sample. For high range samples (1 µg/mL to 20 ng/mL) mix the RiboGreen 1:200 in 1x TE; for the low range samples (50 ng/mL to 1 ng/mL) mix the RiboGreen 1:2000 in 1x TE; Pipette 50 µL of RiboGreen solution into each well of an opaque plate, starting with A1 and moving down, not across
    c. Dilute an RNA stock solution to 2 µg/mL for high range samples or 100 ng/µL for low range samples
    d. Using the stock solutions, create a standard curve by six subsequent 1:2 dilutions of the original 2 µg/mL or 100 ng/mL stock; Pipette 240 µL of the stock into a PCR strip tube. In the remaining 7 tubes, pipette 120 µL of 1x TE. Take 120 µL from the RNA stock well, and mix with tube 2. Take 120 µL from tube 2 and mix with tube 3. Repeat until tube 7. Tube 8 will only have TE and will be your blank. Pipette 50 µL of the standard curve solutions into the opaque plate, A1:H2 (run the curve twice)
    e. In a new tube, mix 45 µL of TE with 5 µL of the oligo pool and add all 50 µL to a well that contains 50 µL of RiboGreen.
    f. Using a plate reader set the machine to read fluorescence from the top after a one-minute mix with an excitation wavelength of 485 nm and emission wavelength of 530 nm and read the plate
    g. Use the standard curve generated to quantify the recovered RNA, which should be stored at -20°C.

Co-immunoprecipitation of the oligo pool with an RNP of interest

1. Prior to beginning the co-immunoprecipitation, prepare the Protein A and Protein G Dynabeads from Invitrogen by mixing them in a 1:1 ratio up to a final volume of 50 µL per reaction. For example, for two reactions you would add 50 µL each of each bead.
2. Use a Magnetic Separation Stand like this one from Novagen to hold the beads in place while you remove the supernatant and wash the beads twice with an equal volume of cold 1x PBS
3. After the second wash, resuspend the beads in an equal volume of cold 1x PBS and add 2 ng of the antibody per reaction. This amount may vary depending on the antibody, so experimentation is necessary to find the optimal conditions.
4. Incubate the antibody/bead mixture overnight at 4°C on a rotating platform
5. In the morning, add 2 µg per reaction of sonicated yeast total RNA to block non-specific binding, and rotate for an additional 30 minutes at 4°C.
6. Using the magnetic rack to hold the beads, remove the supernatant and wash the beads twice with cold 1x PBS, leaving the PBS from the second wash in the tube.
7. Take out new tubes, one for each reaction and aliquot 50 µL of the bead mixture into each tube. Let the beads sit while preparing a master mix
8. For each reaction add 200 ng of the oligo pool, an equimolar ratio of protection primers, 2 µg of sonicated yeast total RNA, 20 µL of HeLa nuclear extract (this is the RNP source), and buffer E to bring the volume up to 120 µL. The protection primers are RNA versions of the universal primers. Order transcription templates of the universal primers with a T7 tag appended to the end, then transcribe RNA from the templates using the MEGAshortscript™ Kit .
9. Remove the supernatant from the bead aliquots and resuspend beads in 120 µL of the master mix. Rock the reactions at 4°C for 1-2 hours.
10. After the incubation, for each reaction, remove a small aliquot (including the beads) for a "total" RNA sample
11. Place the remainder of the reactions on the magnetic separation stand and remove the supernatant. Although it is not necessary to analyze this flow-through sample downstream, it may be useful to save.
12. Wash the beads 1-2 times with 50 µL of cold 1x PBS. The RNA bound to the antibody on the beads is the "IP" sample.
13. Resuspend the "total" and the "IP" beads in 50 µL of 1% SDS in 0.1x TE buffer and heat the samples at 65°C for 15 minutes to elute the RNA.
14. Remove the supernatant, which now contains the RNA that was previously bound to the beads and clean by phenol/chloroform extraction followed by an ethanol precipitation. Dry the pellet and resuspend it in 50 µL 0.1x TE buffer.

Reverse transcription and amplification of co-IP samples in preparation for aminoallyl labeling

1. Next, the RNA samples must be reverse transcribed and PCR amplified. For the reverse transcription reaction, use the reverse universal primers.
2. For each sample, begin by mixing 1 µL of the reverse universal primer, with 1 µL of the template RNA (from the immunoprecipitation reaction) and 10 µL of dH_2O. Heat the reactions at 70°C for 5 minutes then chill on ice or keep at 4°C.
3. While the samples cool, make the following master mix. For each reaction combine 4 µL of 10 mM dNTPs, 2 µL of 10x Arrayscript buffer, 1 µL RNase inhibitor, and 1 µL reverse transcription enzyme. Mix thoroughly and add 8 µL of the master mix to each reverse transcription reaction.
4. Use a PCR machine to incubate the samples at 42°C for 2 hours, followed by 95°C for 5 minutes, then 4°C until the samples are needed for amplification.
5. After reverse transcribing the RNA samples, PCR amplify them using the universal primers, with one of the primers containing a T7 tag. The number of cycles needed to amplify the pool will depend on the efficiency of the co-IP so it may be necessary to try multiple cycles. Each reaction requires 5 µL of 10x Taq reaction buffer, 0.2 µL of each primer (100 mM), 1 µL of the cDNA template (from the reverse transcription reaction), 1 µL of 10 mM dNTPs, 42.2 µL of dH_2O, and 0.4 µL of Taq enzyme.
Begin the PCR by incubating the samples at 94°C for 3 minutes. Next, for each cycle, denature at 94°C for 45 seconds, anneal at 55°C for 30 seconds, and elongate at 72°C for one minute, with a final elongation at 72°C for 10 minutes.
6. Next, transcribe the samples again using the MEGAshortscript™ Kit, however, this time add 2 µL of aminoallyl UTP, instead of the T7 UTP. The aminoallyl UTP will be used to label the RNA with the Cy dyes.
7. Phenol/chloroform and ethanol precipitate the samples. Do NOT use Glycoblue or ammonium acetate, both of which may interfere with hybridization of the samples to the array. Dry the pellet and resuspend in 4.5 µL of 0.1 M sodium carbonate.

Amino-allyl labeling of RNA with Cy Dyes

1. To begin the labeling protocol dissolve each of the fluorescent dyes, cy3 and cy5, in 45 µL of DMSO. To avoid light, wrap the tubes in aluminum foil and store at -20°C in the dark
2. Add 2.5 µL of nuclease-free water to RNA samples and mix thoroughly
3. Next, add 3 µL of Cy3 to the "total" sample, and 3 µL of Cy5 to the "IP" sample. Incubate the samples in the dark for 1 hour at room temperature; The Cy-3 sample will turn red or pink while the Cy-5 will turn blue.
4. To terminate the reaction, add 6 uL of 4 M hydroxylamine-HCl to each sample, mix thoroughly and spin briefly. Incubate for 15 minutes at room temperature.
5. Purify the labeled samples by phenol/chloroform extraction followed by ethanol precipitation and resuspend pellet in 50 µL of 0.1x TE

Hybridize samples to the microarray using the Agilent Array Hybridization Kit

1. Begin the hybridization by gently mixing 50 µL of 10x blocking buffer, 140 µL of nuclease-free $H2_O$, 25 µL of the Cy3-labeled "total" pool, 25 µL of the Cy5-labeled "IP" pool, 10 µL of the 25x fragmentation buffer and 250 µL of the 2x hybridization buffer, taking care to avoid introducing bubbles. Microfuge the sample briefly to collect the solution in the bottom of the tube and also to reduce bubbles.
2. Place the o-ring cover slide in the hybridization chamber and carefully pipette the solution onto the slide.
3. Place the detection array on top of the cover slide to "sandwich" the solution. Remember to orient the array slide so that the Agilent label on the array faces the Agilent label on the cover slide.
4. Clamp the slide/array sandwich in the hybridization chamber and rotate at 50°C for 3 hours.
5. While the array is rotating, make the following 50 mL solutions in 50-mL conical tubes: one conical filled with 50 mL of dH_2O, two 1x SCC buffer solutions, each in a separate tube, and one 2x SCC buffer.
6. After the three-hour incubation, carefully unclamp the array/slide sandwich and place it in the 50 mL conical containing the 2x SCC buffer. Using a pair of tweezers, gently separate the array from the cover slide and pull out the cover slide. Be careful not to touch the array as the slide is pulled out. Close the conical and gently invert for 60 seconds.
7. Using the tweezers, transfer the array to the 1x SCC solution, cap the tube, and gently invert for 1 minute then transfer the array to the second 1x SCC solution and again gently invert for another minute.
8. Lastly, transfer the array to the tube of dH_2O and gently invert for 30 seconds.
9. Dry the array by placing the edges of the slide on a Kimwipe. Continue to rotate the array on the wipe making sure to not disrupt the side containing the oligos.
10. Place the dry array in an empty 50-mL conical tube and scan with a Microarray scanner

Interpreting the Microarray

1. The microarray scanner will create a TIF image of the array. Run this image through the Agilent Feature Extraction software, using the grid file (supplied by Agilent) which is unique to your specific pool.
2. The Agilent software will interpret the signal at each feature to create a file, which is parsed with a Perl script. The signal at each feature is normalize to the total intensity at that channel. For example at a particular feature in the green channel, the intensity of the green channel for the individual feature is divided by the total intensity of the green channel over all features. The final output is the log ratio of the red channel over the green channel at each feature on the array.
3. The oligo pool data from the microarray can be re-mapped back to the genome to get a picture of where, and how strongly, the RNP of interest bound.

Figure 1. Experimental schematic
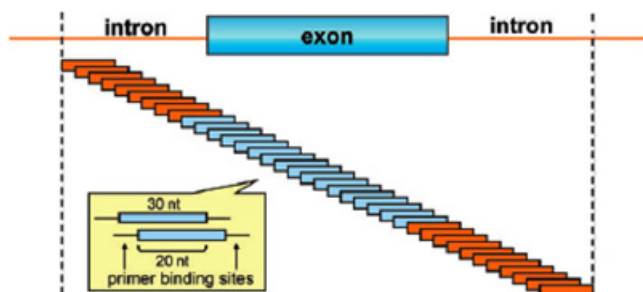
A. Tiling scheme. After selecting and downloading pre-mRNA sequences of interest, the selecting area is tiled through in 30 nucleotide windows that shift in 10 nucleotide increments. Universal primer binding sites flank each side of the window.
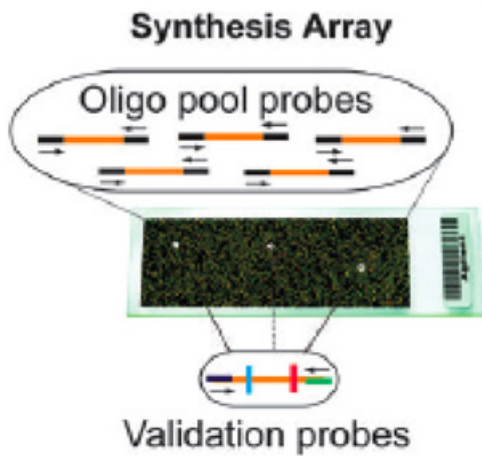
Synthesis Array

B. Synthesis Array. Tiled oligo orders should be submitted to Agilent, where the sequences will be printed on a custom oligonucleotide microarray.
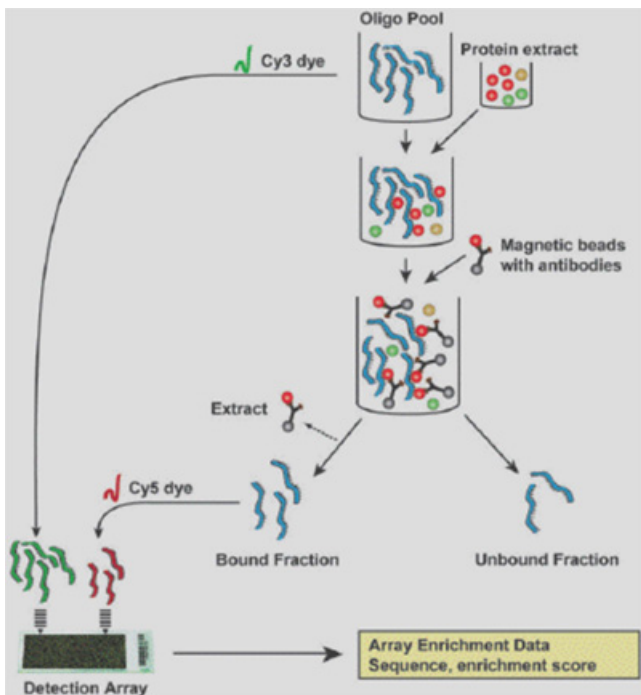
C. Co-immunoprecipitation. After boiling the oligos off the array, incubate in HeLa nuclear extract, then add the extract to magnetic beads that are prepared with an antibody against the RNP of interest. Label the starting oligo pool with Cy3 and the bound IP pool with Cy5 and apply to the detection array and scan.

Figure 2. Annotating the genome with and analysis of enrichment data.
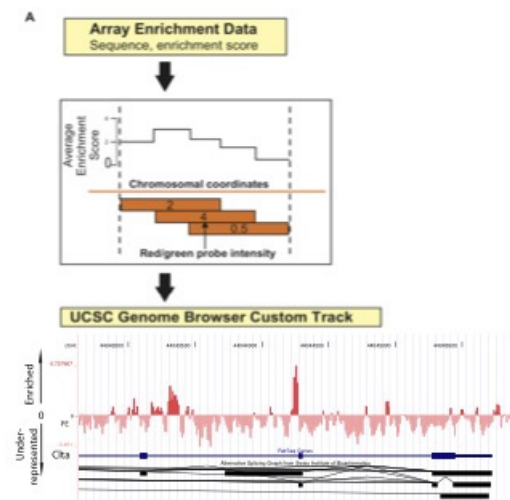
A. Convert the average score at each genomic coordinate from the array data to base-ten log and map that score to the given coordinates. An illustration of this averaging step is given for 3 overlapping 30-nt oligos with scores of 2, 4, and 0.5, where the average enrichment score for each 10-nt window is graphed above. The selected genomic regions can be visualized using a Custom Track in the USCS Genome Browser. The

example browser window given is in the clta gene, where gene features (exons/introns/alternative splicing etc.) are given along the bottom and log average enrichment scores from the PTB-bound oligos are represented by dark red vertical bars. PTB strongly binds just upstream of exon 2 in the polypyrimidine tract.

Discussion

When doing this procedure it's important to remember that creation of the pool is flexible and open to modification at either the RNA or protein level. At the RNA level orthologous regions, disease mutations, polymorphisms, or random mutations can be introduced into the oligonucleotide sequence. At the protein level the RNA binding factor can be modified by phosphorylation or other post translational modifications. Additionally, the binding environment can be manipulated to either increase or decrease the levels of additional factors that interact or compete with the RNA binding protein of interest.

References