Probing RNA structure genome-wide using high throughput sequencing

Michael Kertesz (kertesz@weizmann.ac.il)
Dept. of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel

Yue Wan
Howard Hughes Medical Institute, Program in Epithelial Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

Elad Mazor
Dept. of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel

John Rinn
The Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA 02142, USA

Robert Nutter
Life Technologies, Foster City, CA 94404, USA

Howard Chang
Howard Hughes Medical Institute, Program in Epithelial Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

Eran Segal
Dept. of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel

Method Article

Keywords: RNA, Structure, Secondary Structure, High-throughput sequencing

Posted Date: September 7th, 2010

DOI: https://doi.org/10.1038/nprot.2010.152

Abstract

The functions of RNA in cellular biology are diverse. While it has traditionally been thought to be a passive carrier of information from DNA to proteins, the role of RNA has expanded greatly to include other functional classes such as ribozymes, riboswitches, long and short non-coding RNAs[1-6]. The regulation of coding and non-coding RNAs occur at transcriptional and posttranscriptional levels. While in some cases this regulation occurs due to sequence specificity of transcripts, in other cases RNA regulation is governed by its structure[7]. RNA structure has been shown to be involved in the catalytic activity, stability, localization and rate of translation of transcripts in the cell[8-11]. Understanding what these structural motifs are, how they are organized, and how they might interact with other molecules in the cells, such as other RNAs, proteins and ligands, is critical to our understanding of cellular biology. Traditionally, the secondary structure of an RNA can be probed in solution, one at a time, using chemicals and enzymes that either cleave or modify single or double stranded regions[12-14]. The double or single stranded regions are identified by running the cleaved fragments onto a sequencing gel or by reverse transcription followed by capillary electrophoresis. The intensity of the bands on the gels indicates the extent of chemical/enzymatic cleavage[15]. This method becomes tedious when the structures of long RNAs are of interest. As only a few hundred bases can be resolved at a time, probing the structures of RNAs that are a few kilobases long require running multiple gels. While it is possible to separately clone and probe a region of interest in a long RNA without probing the structure of the full length transcript, doing so risks detecting structural rearrangements that may occur because of the loss in long range contacts. The tedious nature of probing one RNA at a time and the difficulty in probing the structures of long RNAs require the development of a technique that can probe multiple RNA species at a time, independent of their length. We developed a method, Parallel Analysis of RNA Structure \(PARS), to increase the throughput of obtaining experimental RNA structure data independent of RNA size. PARS couples traditional RNA footprinting with high throughput sequencing to identify double and single stranded regions of many RNAs in solution simultaneously. The method provides genome-wide RNA structural information at single nucleotide resolution, greatly expanding the number of RNA secondary structures probed. Here, we describe the PARS protocol in detail. Briefly: RNAs are cleaved by RNase V1 or S1 nuclease independently and the cleavage sites are captured by adaptor ligation followed by high throughput sequencing. The number of sequencing reads mapped to a base in the appropriate library indicates the amount of RNase V1 or S1 nuclease cuts, providing information on whether the base is double or single stranded \(Fig1). PARS can be readily applied to probe RNA structural changes that occur under different solution conditions, enhancing our understanding of RNA structure.

Reagents

Critical: All reagents need to be RNase free and use the highest quality reagents when possible.
PolyA selected RNA at a concentration of at least 25ng/ul.
RNase V1 \(Ambion cat#AM2275)
S1 nuclease \(Fermentas cat#EN0321)

Superase Inhibitor \(Ambion, cat #AM2696)
RNA structure buffer, 10X \(Ambion cat#AM2275)
Alkaline hydrolysis solution \(Ambion cat#AM2275)
19:1 Acrylamide:bis-acrylamide
Urea
TBE 10X \(Invitrogen, cat#15581-044)
Gel loading dye II \(Ambion, cat#AM8546G)
Gel loading solution \(Ambion, cat#AM8556)
Solid Small RNA Expression Kit \(Ambion)
Formamide \(Ambion, cat#AM9342)
Antarctic phosphatase \(NEB cat#M0289)
Costar® Spin-X® centrifuge tube filters \(Sigma-Aldrich, cat#CLS8162-96EA)
Sterile scalpels
RNA Century Markers \(Ambion, cat#AM7140)
20/100 Ladder \(IDT)
Nuclease free H2O \(Ambion, cat#AM9930)
RNA loading dye\(II) \(Ambion, cat#AM8547)
Ethanol \(Goldshield)
Phenol:Chloroform:Isoamyl Alcohol \(25:24:1,v/v) \(Invitrogen, cat#15593-031)
TOPO TA cloning kit \(Invitrogen, cat#K4500-01SC)
Glycogen \(Ambion, cat#AM9510)
Ethidium Bromide Solution \(10mg/ml) \(Bio-rad, cat#161-0433)
Reagent setup

6% denaturing PAGE gel: Make 500ml of 6% Acrylamide-bis acrylamide, 1X TBE, 7M urea PAGE gel mix and store in a light sensitive bottle at 4C. Polymerize the gel mix just before use by adding 1:100 v/v of 10% APS and 1:1000v/v of Temed.
6% native PAGE gel: Make 500ml of 6% Acrylamide-bis acrylamide, 1X TBE, gel mix and store in a light sensitive bottle at 4C. Polymerize the gel mix just before use by adding 1:100 v/v of 10% APS and 1:1000v/v of Temed.
Gel elution buffer: Add 5ml of TE buffer, pH 8, \(10mM Tris-HCl, pH8, 1mM EDTA) to 5ml of 5M ammonium acetate.

Equipment

Vertical gel electrophoresis unit \(Sigma, cat#Z375055)
Vertical gel electrophoresis unit \(R. Shadel, cat# 1211)
High voltage power supply \(Bio-rad, cat #164-5056)
thermal cycler
heat block
-80C freezer

-20C freezer
Rocking Shaker

Procedure

PARS experiments are done using RNase V1 and S1 nuclease. Combing structural information obtained from these 2 enzymes independently creates greater confidence for whether a base is single or double-stranded than used alone. The procedure below describes the processing of one sample that is either cleaved by RNase V1 or S1.
RNA folding and enzymatic cleavage

1. Add 2ug of double polyA selected mRNA in 80ul of nuclease free H2O in a 200ul thin wall PCR tube.
2. Heat the RNA at 90°C for 2min in the thermal cycler with heated lid on; then immediately place the tubes on ice for 2min.
3. Add 10ul of 10X RNA structure buffer \(ice cold); mix by pipetting up and down several times.
4. Transfer the tubes from ice to the thermal cycler; set the thermal cycler such that the temperature slowly increases from 4°C to 23°C over 20min.
<CRITICAL STEP> Bring RNA to 23C slowly allows the RNA to fold properly.
5. Add 10ul of S1 nuclease or 10ul of 1:100 dilution of RNase V1; mix by pipetting up and down. Incubate the samples for 15min.
<CRITICAL STEP> Partial digestion of RNAs by RNase V1 and S1 nuclease allows the enzymes to cleave RNAs at single hit kinetics. This prevents capturing structures that are formed as a result of secondary rearrangements.

Inactivate enzymes and precipitate the cleaved RNA

6. Transfer the reaction to a 1.5ml eppendorf tube containing 100ul of phenol:chloroform:isoamyl alcohol. Vortex the tube vigorously.
7. Spin the tubes in a micro-centrifuge at 4°C, 14000rpm, for 10min.
8. Remove the top aqueous layer carefully and transfer to a new 1.5ml eppendorf tube.
9. Add 10ul of 3M sodium acetate, 1ul of glycogen, and 300ul of 100% cold ethanol.
10. Precipitate the RNA by incubating at -20°C overnight.
11. Spin at 14000rpm, 4°C, for 15min in a centrifuge.
12. Remove the supernatant; wash the RNA pellet in 70% cold ethanol.
13. Spin at 14000rpm, 4°C, for 15min in a centrifuge.
14. Remove the supernatant; then quick spin for a few seconds and remove residual 70% ethanol. Dissolve the RNA pellet in 4ul of nuclease free H2O and transfer to PCR tube. Keep on ice.

Fragmentation of RNA

15. Add 20ul of 1X alkaline hydrolysis buffer to a PCR tube; then heat at 95°C for 1min using a thermal cycler.

16. Transfer 4ul of RNA in PCR tube from ice to a thermal cycler set at 95°C; then heat for 10sec. These pre-heating steps prevent the RNA from being fragmented before the RNA is fully denatured.
17. Add 16ul of the heated 1X alkaline hydrolysis buffer to the heated RNA; mix by pipetting up and down several times. Incubate for 3min at 95°C, then immediately place the RNA on ice.
<CRITICAL STEP> Accurate timing of RNA fragmentation is important to concentrate most of the RNA fragments to around 200bases.
18. Add 2ul of 3M sodium acetate to stop the fragmentation reaction.

Size selection of RNA by gel electrophoresis.

This size selection step removes very short fragments of RNA, generated from the fragmentation step, that could ligate to adaptors.

19. Add 2ul of the RNA century ladder to a 1.5ml eppendorf tube; then add 6ul of gel loading dye \(II). Mix by pipetting up and down. Incubate at 70°C for 10min in a heat block; then immediately place on ice and load into a lane of a 6% polyacrylamide sequencing gel \(19:1 acrylamide:bis-acrylamide, 1X TBE, 7M urea).
20. Add 2ul of the 20/100 ladder to a 1.5ml eppendorf tube; then add 6ul of gel loading dye \(II). Mix by pipetting up and down. Incubate at 70°C for 10min in a heat block; then immediately place on ice and load into a lane of a 6% polyacrylamide sequencing gel.
21. Add 22ul of Gel loading dye\(II) to the reaction and load all of the reaction into a lane of a 6% polyacrylamide sequencing gel.
22. Perform electrophoresis at 45W until the lower blue dye has run about 2/3 the length of the gel.
23. Add 1ml of nuclease free H2O to a 1.5ml eppendorf tube; then add 1ul of ethidium bromide.
24. Peel the gel off the glass plates onto a piece of saran wrap, add 1:1000 diluted ethidium bromide over the gel, and cover the gel with another piece of saran wrap so that the gel is sandwiched between 2 pieces of saran wrap. Smooth out the gel so that the ethidium bromide is spread out evenly across the gel. Stain for 5min.
25. Visualize the gel under long UV; then use a sterile scalpel to cut the region of the gel between 50-200 bases as shown by the molecular weights of fragments in your RNA century and 20/100 ladder lanes.
26. Puncture the bottom of an RNase free 600ul tube 2-3X using a 18.5G needle; then place this tube in a 1.5ml eppendorf tube. Transfer the cut gel slice into the 600ul tube and spin at 14000rpm for 1min. This step shreds the gel slices, allowing maximum surface area for the RNA to diffuse out of the gel.
27. Add 300ul of nuclease free H2O to the shredded gel; incubate at 4°C overnight on a rocking platform.
28. Cut the end of a 1ml pipette tip and use it to transfer the gel slices into a Costar® Spin-X® centrifuge tube filter. Spin at 10000rpm for 2min in a centrifuge.
29. Add 30ul of sodium acetate, 3ul of glycogen and 900ul of cold ethanol to the supernatant and precipitate the RNA at -20°C, overnight.
30. Spin the tubes at 14000rpm for 15min, at 4°C, in a centrifuge.
31. Remove the supernatant and wash with 70% cold ethanol; then spin the tubes at 14000rpm for 15min at 4°C.

32. Remove the supernatant; then quick spin the tubes and use a p10 tip to remove any traces of ethanol.
33. Resuspend the RNA in 3ul of nuclease free water in a 200ul thin wall PCR tube.

Adaptor ligation \(I)

The library preparation steps below are modifications made to the SOLiD™ Small RNA Expression Kit protocol. This kit has been replaced by the SOLiD™ Total RNA Seq Kit \(PN 4445374.), refer to manufacturer's instructions for ligation conditions.

34. Add 2ul of adaptor mix A \(Solid small RNA expression kit) and 3ul of hybridization buffer \(Solid small RNA expression kit) to the RNA; then mix by pipetting up and down.
35. Heat the tube to 65°C, in a thermal cycler, for 10min; then incubate it at 16°C for 5min. Put the sample on ice.
36. Add 10ul of 2X ligation buffer \(SOLiD™ Small RNA Expression Kit) slowly.
37. Add 2ul of the ligation enzyme mix \(SOLiD™ Small RNA Expression Kit). Mix well by flicking the tube such that the sample looks homogenous. Briefly spin the tubes to collect the sample at the bottom of the tubes.
38. Incubate the sample in a thermal cycler at 16°C overnight, with heated lid off.

Remove 3’P to form 3’OH for more efficient 3’end ligation.

39. Add 5ul of 10X Antarctic phosphatase buffer to the sample; mix by pipetting up and down.
40. Add 2.5ul of Superasin RNase Inhibitor; mix by pipetting up and down.
41. Add 2.5ul of Antarctic phosphatase enzyme; mix by pipetting up and down.
42. Add 20ul of nuclease free H2O; mix by pipetting up and down. The total volume of reaction is 50ul.
43. Incubate sample at 37°C for an hour.

Inactivation of the Antarctic phosphatase

44. Transfer the reaction to a 1.5ml eppendorf tube. Add 150ul of nuclease free H2O; then add 100ul of \(100mM Tris pH8); mix by pipetting up and down.
45. Add 300ul of phenol:chloroform:isoamyl alcohol mix. Vortex.
46. Spin at 14000rpm, for 10min, at 4°C.
47. Transfer the supernatant to a new 1.5ml tube and add 30ul of 3M sodium acetate, 3ul of glycogen and 900ul of 100% ethanol.
48. Incubate at -20°C overnight or -80°C for an hour.
49. Spin at 14000rpm, for 15min, at 4°C.
50. Remove supernatant; then add 1ml of 70% ethanol to wash the pellet.
51. Spin at 14000rpm, for 15min, at 4°C.
52. Remove supernatant; spin briefly to collect excess 70% ethanol at the bottom of the tube. Remove the residual ethanol using a p10 tip.
53. Dissolve the RNA pellet in 3ul of H2O.

Adaptor ligation \(II)

54. Add 2ul of adaptor mix A \(SOLiD™ Small RNA Expression Kit) and 3ul of hybridization buffer \(SOLiD™ Small RNA Expression Kit) to the RNA; then mix by pipetting up and down. \(This kit has been replaced by the SOLiD™ Total RNA Seq Kit \(PN 4445374.), refer to manufacturer's instructions for ligation conditions.)
55. Heat the tube to 65°C, in a thermal cycler, for 10min; then incubate it at 16°C for 5min. Put the sample on ice.
56. Add 10ul of 2X ligation buffer \(from SOLiD™ Small RNA Expression Kit) slowly.
57. Add 2ul of the ligation enzyme mix \(SOLiD™ Small RNA Expression Kit). Mix well by flicking the tube such that the sample looks homogenous. Briefly spin the tubes to collect the sample at the bottom of the tubes.
58. Incubate the sample in a thermal cycler at 16°C overnight, with heated lid off.

Reverse transcription and RNase H treatment

The library preparation steps below are modifications made to the SOLiD™ Small RNA Expression Kit protocol. This kit has been replaced by the SOLiD™ Total RNA Seq Kit \(PN 4445374.), refer to manufacturer's instructions for reverse transcription and PCR conditions.

59. Make reverse transcription master mix on ice: Add 13ul of nuclease free H2O, then 4ul of RT buffer, then 2ul of 2.5mM dNTP, then 1ul of arrayscript RT, to a tube.
60. Add 20ul of the RT master mix to the sample. Mix by pipetting up and down several times and flicking the tube. Spin briefly to bring the contents to the bottom of the tube.
61. Place the tubes at 42°C, for 30min, in a thermal cycler. The cDNA can be stored at -20°C.
62. Transfer 10ul of the RT reaction to a new tube; then add 1ul of RNase H. Mix by flicking the tube; then briefly spin it down.
63. Incubate the tubes at 37°C for 30min. The RNase H treated sample can be stored at -20°C.

PCR amplification \(small scale PCR)

64. Prepare the PCR master mix on ice, in a new tube. Add 38.9ul of H2O, then add 5ul of 10X PCR buffer I, then add 1ul of PCR primers \(any set out of the 10 sets is fine), then add 4ul of 2.5mM dNTP, then add 0.6ul of AmpliTaq polymerase.
65. Add 49.5ul of the PCR master mix to a thin wall 200ul PCR tube; then add 0.5ul of the RNase H treated sample. Vortex briefly to mix; then spin the contents down.
66. Place the tubes in a thermal cycler and run following program: 95°C for 5min; then 30cycles of 95°C \(30sec) followed by 62°C \(30sec) followed by 72°C \(30sec); then 72°C for 7min. At 15 cycles, pause the program at 72C, and transfer 10ul of the PCR reaction to a new tube. Place the tube on ice. Un-pause the program and repeat the same for 20, 25 and 30cycles. There should be 4 tubes at the end of the PCR, each with sample that has undergone either 15, 20, 25, 30cycles. This helps to determine the lowest amount of PCR cycles needed to amplify the sample.

67. Load 10ul of 100bp ladder and load 10ul of 1:10 diluted 50bp ladder into 2 separate wells.
68. Add 2ul of the gel loading solution to the PCR product; then load the sample into 1 well in a 6% native polyacylamide gel. Skip at least 1 lane between the sample and the ladders.
69. Run at about 120V until the lower blue dye is near the bottom of the gel.
70. Peel the gel off the glass plates onto a piece of saran wrap, add 1:1000 diluted ethidium bromide over the gel, and cover the gel with another piece of saran wrap so that the gel is sandwiched between 2 pieces of saran wrap. Smooth out the gel so that the ethidium bromide is spread out evenly across the gel. Stain for 5min.
71. Visualize the gel under UV. The amplified samples should have a smear at around 200bases all the way till 300bases. Choose the lowest PCR cycle, which shows such a smear, to amplify the samples for large scale PCR amplification. Over-amplified samples tend to have smears with higher molecular weights.

Large scale PCR amplification

72. Prepare the PCR master mix on ice, in a new tube. For each 100ul reaction: add 77.8ul of H_2O, then add 10ul of 10X PCR buffer I, then add 2ul of PCR primers, then add 8ul of 2.5mM dNTP, then add 1.2ul of AmpliTaq polymerase. Prepare about 8 reactions per sample.
73. Add 99ul of the PCR master mix to a thin wall 200ul PCR tube; then add 1ul of the RNase H treated sample. Vortex briefly to mix; then spin the contents down.
74. Place the tubes in a thermal cycler and run following program: 95°C for 5min; then X cycles of 95°C \(30sec) followed by 62°C \(30sec) followed by 72°C \(30sec); then 72°C for 7min. \(X is determined by the lowest PCR cycles necessary in the small scale PCR amplification) to visualize the amplified production between 200-300 bases. A successful library typically requires PCR cycles of 20 cycles or less. Excessive PCR will result in the distortion of the cleavage signal.

Cleanup of the large scale PCR amplification

75. Pool replicate large-scale PCR reactions from the same sample together.
76. Add 5 volumes of buffer PB to the sample; mix by pipetting up and down several times. \(If the volume is too big to be contained in a 1.5ml tube, split the reaction into several tubes).
77. Add 750ul of the sample to a MinElute column \(Qiagen).
78. Spin at 14000rpm, for 1min. \(Load the sample multiple times through the column until all the samples have passed through the column).
79. Add 750ul of buffer PE \(add ethanol prior to using PE) to the column. Spin at 14000rpm for 1min
80. Spin at 14000rpm for 1 min to dry the column.
81. Elute in 2X 20ul of nuclease free H_2O in a new 1.5ml tube.

Size selection of PCR products between 150-250bases.

82. Add 8ul of gel loading dye to 40ul of elute PCR product.
83. Load 10ul of 100bp ladder and load 10ul of 1:10 diluted 50bp ladder into 2 separate wells in a 6%

native PAGE gel.
84. Skip at least 1 lane between the sample and the ladders; then add 12ul of PCR product into 1 well in a 6% native PAGE gel. \(Load the equivalent of 200ul of large scale PCR product into 1 well to prevent overloading of the lane). Continue loading the wells until all the sample is loaded.
85. Run at about 120V until the lower blue dye is near the bottom of the gel.
86. Peel the gel off the glass plates onto a piece of saran wrap, add 1:1000 diluted ethidium bromide over the gel, and cover the gel with another piece of saran wrap so that the gel is sandwiched between 2 pieces of saran wrap. Smooth out the gel so that the ethidium bromide is spread out evenly across the gel. Stain for 5min.
87. Visualize the gel under long UV. Use a sterile scalpel to cut out a gel slice between 150bases to 250 bases for all the samples. Change scalpels between different samples to avoid sample contamination. <CRITICAL STEP> Cutting a band above 150bases prevents contamination of adaptor dimers.
88. Puncture the bottom of an RNase free 600ul tube 2-3X using a 18.5G needle; then place this tube in a 1.5ml eppendorf tube. Transfer the cut gel slice into the 600ul tube and spin at 14000rpm for 1min. This step shreds the gel slices, allowing maximum surface area for the RNA to diffuse out of the gel.
89. Add 300ul of gel elution buffer to the shredded gel; incubate at room temperature overnight.
90. Transfer 300ul of the gel elution buffer into a new tube on ice. Add another 300ul of gel elution buffer to the shredded gel; incubate at room temperature for another 2 hours.
91. Cut the end of a 1ml pipette tip and use it to transfer the elution buffer from the first overnight incubation into a Costar® Spin-X® centrifuge tube filter. Spin at 10000rpm for 2min in a centrifuge. Repeat this step with the elution buffer from the second incubation.
92. Add 6ul of glycogen, then add 450ul of isopropanol to the elution buffer from the Spin-X® centrifuge filter. Incubate at room temperature for 10min.
93. Spin at 14000rpm for 20min.
94. Remove the supernatant; then add 1ml of 70% ethanol. Gently pipette the solution up and down without disturbing the pellet. Remove the 70% ethanol and repeat this three times.
95. Dissolve the DNA pellet in 5 of water; then pool DNA from the same sample together into 1 tube.
96. Quantitate the amount of DNA present by nanodrop.
97. Check the size of the library using Agilent Bioanalyzer.

TA cloning

This step clones the products present in the library which can then be sequenced by traditional capillary electrophoresis \(CE) sequencing to confirm that the products are mostly different mRNAs in the cell.
98. Transfer 2ul of the cDNA library to a new tube; then add 2ul of H2O, then add 1ul of salt solution \(TOPO TA cloning kit, Invitrogen); then add 1ul of vector \(TOPO TA cloning kit, Invitrogen). Mix by flicking the tube; then briefly spin the tube.
99. Incubate on benchtop for an hour; then transform 2ul of the reaction into Top10 competent cells \(Invitrogen) by following the manufacturer's instructions.
100. Pick 20 colonies to send for CE sequencing.

Sequencing of the libraries

101. The libraries \(from step 96) are prepared for SOLiD™ ABI cluster generation and sequencing following the manufacture's standard protocols. The sequenced reads are then mapped to the yeast genome and transcriptome

Timing

Day 1: Steps 1-27
Day 2: Steps 28-38
Day 3: Steps 39-58
Day 4: Steps 59-89
Day 5: Steps: 90-99

Troubleshooting

1. No amplified products between 200-300 bases. The most common problem behind no amplified products is nuclease contamination. As this protocol manipulates the RNA across a few days- it is critical to keep all reagents RNase free. Wipe down all pipettes with RNaseOUT \(Ambion) before starting. Too little or excessive fragmentation could also result in a failed experiment. After fragmentation of RNA, some of the fragmentation products can be run on the Agilent Bioanalyzer to determine the final fragmentation size. The fragmentation size should peak at around 200bases. Reduce the amount of fragmentation time from 3min to 2min if there is too much fragmentation, alternatively, increase the fragmentation time to 4min if there is too little fragmentation.
2. Most of the TA cloning products are adaptor linkers. The gel slice is cut too low in molecular weight. Re-run the 6% native PAGE and cut slightly higher- above 150 to 250bases for good separation from the adaptor dimers. Check that the gel does not heat up during running. Heating up may denature the double stranded cDNA causing changes in migration.

3. Most of the TA cloning products are ribosomal RNA. The polyA selection steps to deplete ribosomal RNA are not efficient, polyA selection need to be repeated with new total RNA. Before starting the protocol- run total RNA, one time polyA selected RNA and double selected RNA on the Agilent Bioanalyzer. Double selected polyA mRNA should be a smear and the ribosomal RNA bands should be very faint.

4. The same fragment comes up multiple times in one TA cloning reaction. The library may be of low complexity and hence the same product is captured repeatedly. Repeat the protocol with twice as much starting material.

Anticipated Results

A PARS experiment consists of at least 2 biological replicates of libraries generated from RNase V1 and S1 nuclease cleavage. The amount of sequencing reads that is mapped to a base in the transcript from

either RNase V1 or S1 nuclease indicates whether the base is double or single stranded. In this experiment, the P9-9.2 domain of the Tetrahymena ribozyme was doped into the pool of RNA and cleaved by RNase V1 and S1 nuclease independently 17\(Fig2). We independently radioactively labeled the 5' end of P9-9.2 and performed traditional footprinting on this fragment. The intensities of the bands on the sequencing gel were quantitated by SAFA15. The positions of the cleavages detected by high throughout sequencing corresponded to the positions of cleavages by traditional footprinting, with the S1 nuclease cleaving in single stranded regions and RNase V1 cleaving in double stranded regions14. Because of the size selection step \(cutting the gel slice between 150-250 bases) that we introduced to prevent contamination of adaptor dimers, the last 30-40bases of the RNA could not be probed.

The choice of RNase V1 and RNase S1 as cleavage enzymes was deliberate because these enzymes cleave leaving a 5'P that allow us to capture the exact cleavage site. While other structural probes can also be used, this protocol would need to be modified. Many of the chemicals/enzymes identify single stranded bases by cleaving the RNA and leaving behind 5'OH13. However, these 5'OH cannot be captured directly by adaptor ligation unless they are converted to 5'P. This phosphate addition step has the potential to introduce noise as degradation products that were originally present in the cell are captured and sequenced. One potential way to overcome this problem could be to kinase and fragment the RNAs without incubating the RNAs with structural probes. This could potentially serve as a background lane for subtraction of noisy signals from the structurally probed RNAs.

On top of being able to probe many RNAs simultaneously, PARS also has the advantage to probe RNA structures independent of RNA length. This allows RNAs to be probed in their full-length context and maintain any long-range interactions that were originally present in the molecule. PARS is readily applicable to probe structural changes under diverse solution conditions to provide an understanding of how structural changes can occur in the presence of ligands, proteins and different temperatures3, 16. The explosion of information on RNA structures has the potential to change our understanding of how RNA structure govern RNA function.

References

1. Regulski, E. E. & Breaker, R. R. In-line probing analysis of riboswitches. Methods Mol. Biol. 419, 53-67 \(2008).
2. Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science 289, 905-920 \(2000).
3. Chowdhury, S., Maris, C., Allain, F. H. & Narberhaus, F. Molecular basis for temperature sensing by an RNA thermometer. EMBO J. 25, 2487-2497 \(2006).
4. Rinn, J. L. et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129, 1311-1323 \(2007).
5. Bartel, D. P. & Chen, C. Z. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nat. Rev. Genet. 5, 396-400 \(2004).

6. Kruger, K. et al. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell 31, 147-157 \(1982).
7. Hogan, D. J., Riordan, D. P., Gerber, A. P., Herschlag, D. & Brown, P. O. Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. PLoS Biol. 6, e255 \(2008).
8. Guo, F., Gooding, A. R. & Cech, T. R. Structure of the Tetrahymena ribozyme: base triple sandwich and metal ion at the active site. Mol. Cell 16, 351-362 \(2004).
9. Wang, Y. H., Sczekan, S. R. & Theil, E. C. Structure of the 5' untranslated regulatory region of ferritin mRNA studied in solution. Nucleic Acids Res. 18, 4463-4468 \(1990).
10. Kudla, G., Murray, A. W., Tollervey, D. & Plotkin, J. B. Coding-sequence determinants of gene expression in Escherichia coli. Science 324, 255-258 \(2009).
11. Jambhekar, A. et al. Unbiased selection of localization elements reveals cis-acting determinants of mRNA bud localization in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. 102, 18005-18010 \(2005).
12. Deigan, K. E., Li, T. W., Mathews, D. H. & Weeks, K. M. Accurate SHAPE-directed RNA structure determination. Proc. Natl. Acad. Sci. U. S. A. 106, 97-102 \(2009).
13. Gornicki, P. et al. Use of lead\(II) to probe the structure of large RNA's. Conformation of the 3' terminal domain of E. coli 16S rRNA and its involvement in building the tRNA binding sites. J. Biomol. Struct. Dyn. 6, 971-984 \(1989).
14. Auron, P. E., Weber, L. D. & Rich, A. Comparison of transfer ribonucleic acid structures using cobra venom and S1 endonucleases. Biochemistry 21, 4700-4706 \(1982).
15. Das, R., Laederach, A., Pearlman, S. M., Herschlag, D. & Altman, R. B. SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. RNA 11, 344-354 \(2005).
16. Winkler, W., Nahvi, A. & Breaker, R. R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature 419, 952-956 \(2002).
17. Kertesz, M., Wan, Y., Mazor, E., Rinn,J.L., Nutter, R.C., Chang, H.Y. & Segal, E. Genome-wide measurement of RNA secondary structure in Yeast. Nature \(2010) in press.

Figures

Figure 1

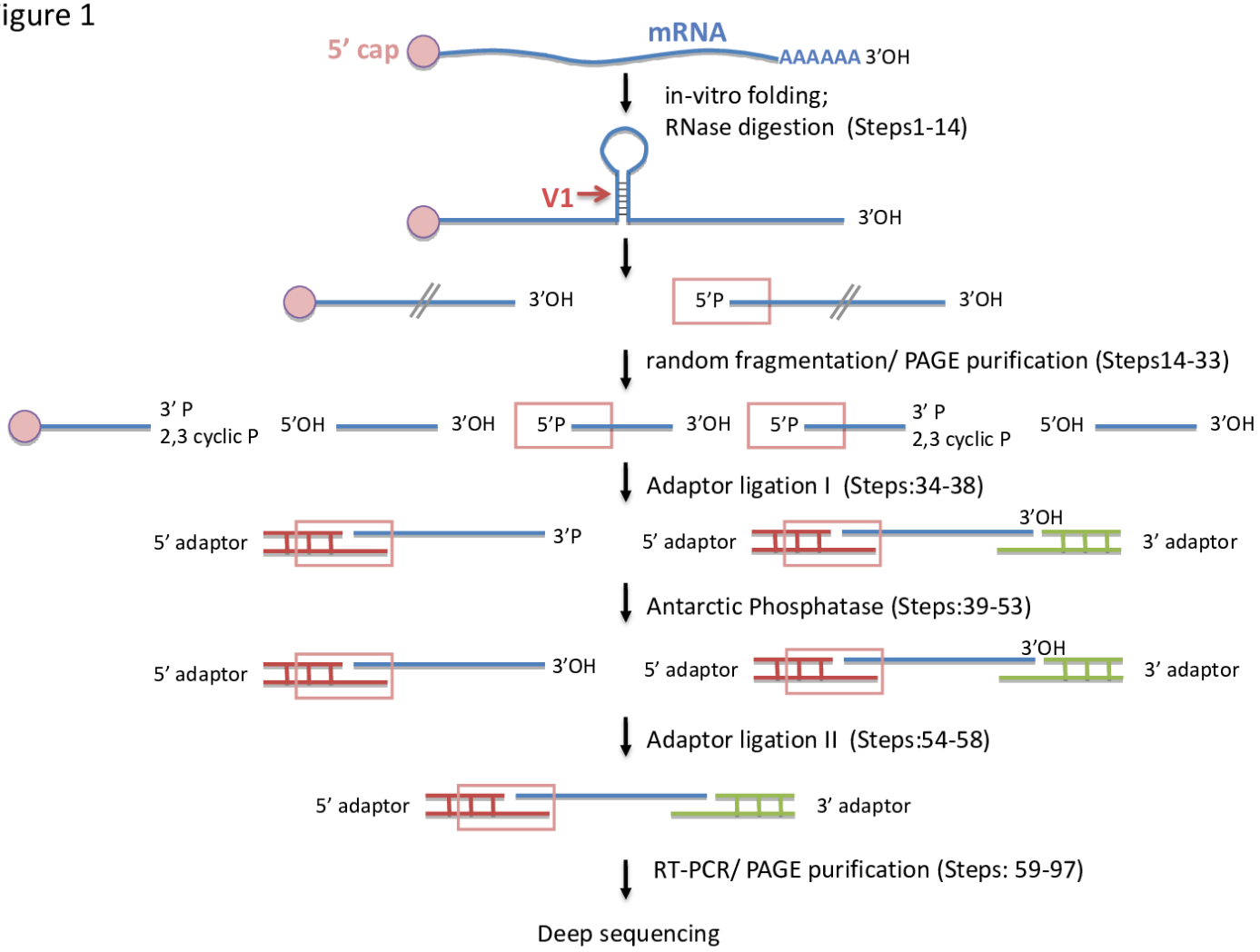

Figure 1

<h3> Detailed schematic of RNA structure probing by high throughput sequencing. </h3> Total RNA is isolated from yeast, selected for poly(A)+ transcripts and renatured in vitro. The folded RNA is then cut by RNase V1 (shown), or RNase S1, resulting in 5'P. The RNA undergoes fragmentation, however only the RNA that has been cut by RNase V1 or S1 contains 5'P that is ligation competent. The RNA then undergoes size selection, followed by 5'adaptor ligation. Fragmentation products with 3'P are converted to 3'OH by Antarctic phosphatase making them able to ligate to 3' adaptors. This is followed by reverse transcription, PCR, and size selection between 150-250 bases to make a library that is suitable for high throughput sequencing.

Figure 2

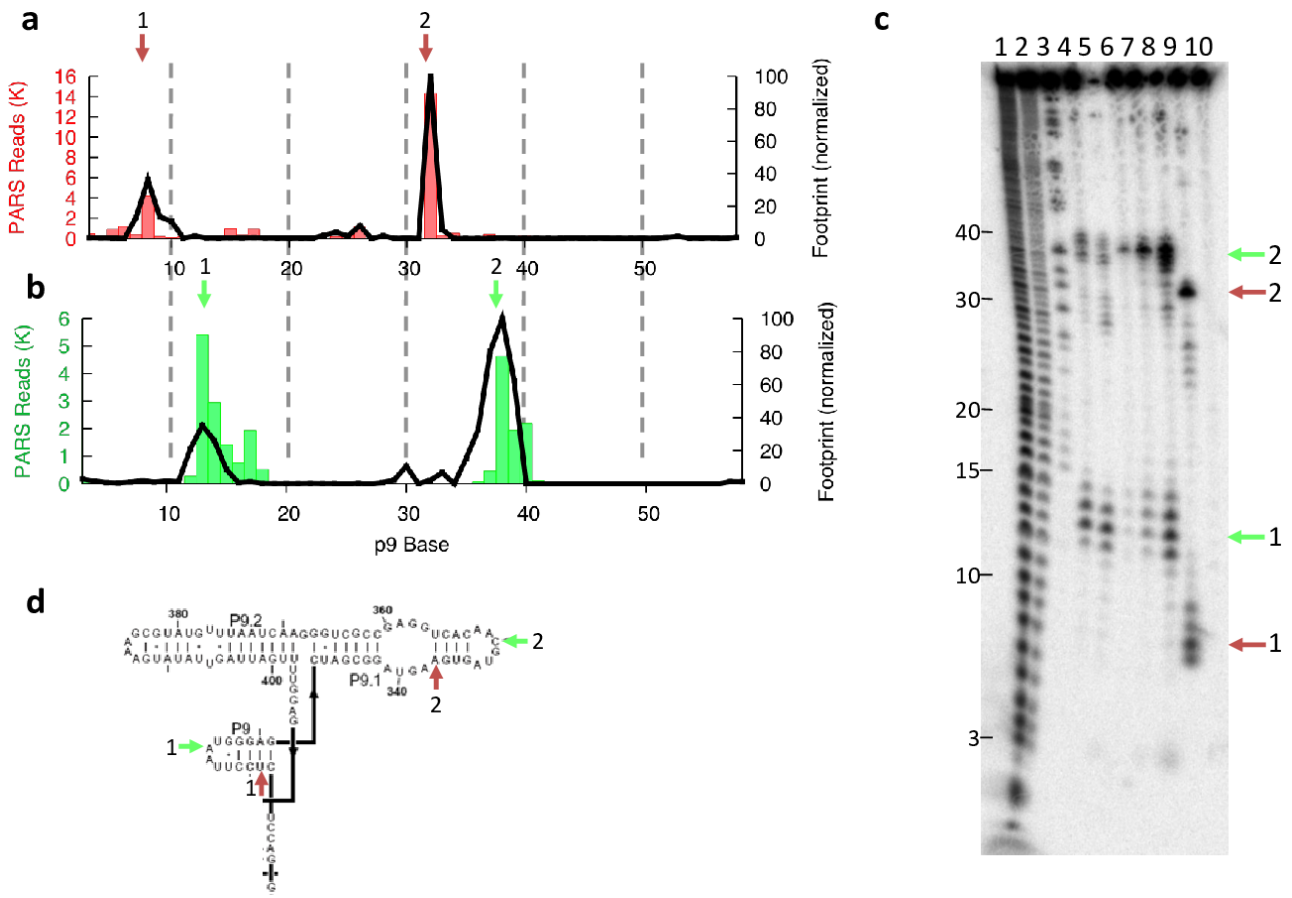

Figure 2

<h3>PARS correctly recapitulates results of RNA footprinting for the p9-9.2 domain of the Tetrahymena ribozyme.</h3> (a) RNase V1 cleaves the folded p9-9.2 domain of the Tetrahymena ribozyme at two distinct sites, which are accurately captured by PARS. The double-stranded signal of PARS obtained using the double-stranded cutter RNase V1 (red bars) is shown as the number of sequence reads mapped along each nucleotide of the P9-9.2 domain. Also shown is the signal obtained on the P9-9.2 domain using traditional footprinting (black line) and semi-automated quantification of the RNase V1 lane shown in (c). Red arrows indicate cleavages that are seen in gel (c). (b) Single-stranded signal of PARS obtained using the single-stranded cutter RNase S1 (green bars), compared to the signal obtained using traditional footprinting (black line). Green arrows indicate cleavages that are seen in gel (c). (c) The gel resulting from RNase V1 (Lane 9) and RNase S1 (Lanes 6,7,8 at pH7 and Lanes 4,5 at pH4.5). Alkaline hydrolysis (Lanes 1,2), RNase T1 ladder (Lane 3) and no RNase treatment (Lane 10) are also shown. (d)Known secondary structure of the p9-9.2 domain. Arrows mark nucleotides that were identified by both PARS and enzymatic probing as double-stranded (red arrows) or single-stranded (green arrows).