Cold Spring Harbor Protocols

High-Throughput Illumina Strand-Specific RNA Sequencing Library Preparation

Silin Zhong, Je-Gun Joung, Yi Zheng, Yun-ru Chen, Bao Liu, Ying Shao, Jenny Z. Xiang, Zhangjun Fei and James J. Giovannoni

Email Alerting Service	Receive free email alerts when new articles cite this article - **click here.**
Subject Categories	Browse articles on similar topics from *Cold Spring Harbor Protocols*. Analysis of Gene Expression (142 articles) Analysis of Gene Expression, general (99 articles) High-Throughput Analysis, general (97 articles) Libraries (97 articles) Libraries, general (97 articles)

Protocol

High-Throughput Illumina Strand-Specific RNA Sequencing Library Preparation

Silin Zhong,[1,2,5] Je-Gun Joung,[1] Yi Zheng,[1] Yun-ru Chen,[1] Bao Liu,[2] Ying Shao,[3] Jenny Z. Xiang,[3] Zhangjun Fei,[1,4,5] and James J. Giovannoni[1,4,5]

[1]Boyce Thompson Institute for Plant Research, Cornell University, Ithaca, NY 14853, USA
[2]Institute of Genetics and Cytology, Northeast Normal University, Changchun, 130024, China
[3]Weill Medical College, Cornell University, New York, NY 10021, USA
[4]U.S. Department of Agriculture/Agriculture Research Service, Plant, Soil, and Nutrition Laboratory, Ithaca, NY 14853, USA

INTRODUCTION

Conventional Illumina RNA-Seq does not have the resolution to decode the complex eukaryote transcriptome due to the lack of RNA polarity information. Strand-specific RNA sequencing (ssRNA-Seq) can overcome these limitations and as such is better suited for genome annotation, de novo transcriptome assembly, and accurate digital gene expression analysis. This protocol describes a simple and robust method to generate ssRNA-Seq libraries for the Illumina sequencing platform. It has significantly increased the throughput to 96 libraries in a two-day preparation while simultaneously lowering the reagent costs to below ten dollars per library. It is compatible with both single-read and paired-end multiplex sequencing and, most importantly, its data can also be used with existing conventional RNA-Seq data. This is a significant advantage, because it enables researchers to switch to ssRNA-Seq even if a large amount of data has already been generated by the nonstrand specific methods.

RELATED INFORMATION

In the conventional Illumina RNA-Seq library preparation, the RNA orientation information is lost at the second-strand cDNA synthesis step, as the sequencing reads derived from the first- and second-strand cDNA are indistinguishable. The Illumina ssRNA-Seq kit uses a strategy similar to that of its small RNA sequencing kit, in which two adapters are sequentially ligated to the 3′- and 5′-end of the mRNA (Vivancos et al. 2010). Since the sequencing primer binding site is located in the 5′-end adapter, all reads would start from the 5′-end of the first-strand cDNA. However, this method is labor intensive and RNA adapters and preadenylated DNA oligos are difficult and costly to synthesize. In addition, because a barcode sequence could not be incorporated into the 5′ adapter RNA oligo easily without generating bias for the RNA ligation, this method is not ideal for multiplexing (sequencing multiple RNA samples in one lane, also known as "indexed" or "barcoded" sequencing), which is crucial for large-scale sequencing projects.

To circumvent this, we have reverse-engineered the conventional RNA-Seq library preparation procedure to make it faster, more economical, and compatible with large scale ssRNA-Seq (Fig. 1). In this method, the rate-limiting mRNA fragmentation and ethanol precipitation steps are replaced by a single-step fragmentation and elution. We also use magnetic particles to purify the RNA/cDNA hybrid and perform second-strand synthesis with dUTP. The strand-specificity is achieved by enzymatically removing the dUTP-containing second-strand cDNA prior to polymerase chain reaction (PCR) enrichment (Parkhomchuk et al. 2009). The time-consuming agarose gel electrophoresis step is replaced by sequential removal of the short (adapter dimer) and long (library with large insertion) DNA fragments with carboxyl group coated magnetic particles, leaving a clean and uniformly sized library for sequencing. Using this approach, we have generated hundreds of ssRNA-Seq libraries for the tomato and rice sequencing projects since 2009, and it has also been successfully used for the maize and sorghum projects.

www.cshprotocols.org

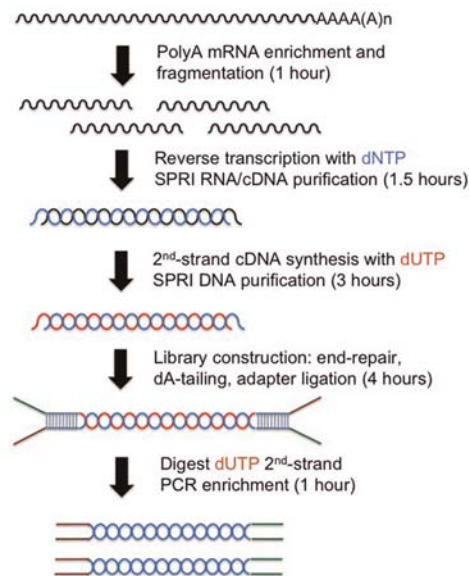Cold Spring Harbor Protocols
www.cshprotocols.org

FIGURE 1. Strand-specific RNA sequencing library preparation. PolyA RNA (shown in black) is captured by oligodT beads and fragmented when eluted from the beads. The first-strand cDNA was synthesized by reverse transcriptase using dNTP (blue), whereas the second-strand cDNA was generated using a dUTP mix (red). After end-repair, dA-tailing and adapter ligation, the dUTP-containing second strand is digested by uracil DNA glycosylase, rendering it unamplifiable in the subsequent PCR step. Hence the resulting library contains only PCR product of the first-strand cDNA.

MATERIALS

RECIPES: Please see the end of this article for recipes indicated by <R>.

It is essential that you consult the appropriate Material Safety Data Sheets and your institution's Environmental Health and Safety Office for proper handling of equipment and hazardous materials used in this protocol.

Reagents

Actinomycin D (Sigma A1410, dissolved in dimethylsulfoxide [DMSO] to 1 µg/µL)
Agarose gel
AMPure XP (Agencourt A63881)
Barcode adapter oligo A (the Ns represent barcode sequence): 5′-ACACTCTTTCCCTACACGACGCT CTTCCGATCTNNNNNT-3′
Barcode adapter oligo B (5′ phosphorylated oligo; the Ns represent barcode sequence): 5′-pNNNNN AGATCGGAAGAGCGGTTCAGCAGGAATGCCGAG-3′
<R>Binding buffer (2×)
Blue Buffer (10×; Enzymatics) or NEBuffer 2 (New England Biolabs)
Dithiothreitol (DTT) (100 mM stock), supplied with SuperScript III RT (Invitrogen)
DNA Polymerase I (Enzymatics P705L)
dNTPs (dATP, dTTP, dGTP, dCTP; dUTP is used in place of dTTP in Step 19)
Dynabeads oligo(dT)25 (Invitrogen 610-05)
End-Repair Buffer (10×; Enzymatics), provided with End-Repair Mix (Low Concentration), or PNK buffer (10×; New England Biolabs)
End-Repair Mix (Low Concentration) (Enzymatics Y914-LC-L)
Ethanol (75%)
Klenow 3′–5′ exo⁻ (Enzymatics P701-LC-L)
PE (paired-end) primer A: 5′-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCT CTTCCGATCT-3′
PE primer B: 5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCC GATCT-3′
Phusion II with HF Buffer (5×) (New England Biolabs)
Random hexamer, synthesized $(N)_6$
Rapid Ligation Buffer (2×; Enzymatics), provided with T4 DNA Ligase HC
RNAClean XP (Agencourt A66514)
RNase H (Enzymatics Y922L)

RNasin Plus RNase Inhibitor (Promega N2615)
SuperScript III Reverse Transcriptase (RT; Invitrogen 18080-085) with 5× first-strand buffer
T4 DNA Ligase HC (Enzymatics L603-HC-L)
TE buffer (10 mM Tris HCl [pH 7.5], 1 mM EDTA)
Uracil DNA Glycosylase (UDG; Enzymatics G501L)
<R>Washing buffer for ssRNA-Seq

Equipment

Bioanalyzer, Agilent
Fluorometer, Qubit (Invitrogen Q32866), to use with Qubit dsDNA HS Assay Kit (Invitrogen Q32851)
Magnetic stand
Micropipettor and tips
Microplate reader (BioTek), to use with Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen P7589)
PCR tubes (LPS L232401)
Plates (96-well; Sigma-Corning CLS3360)
Thermocycler

METHOD

PolyA RNA Isolation and Fragmentation

1. Prepare oligo(dT)25 Dynabeads as follows:

   i. Use an appropriate amount of oligo(dT)25 Dynabeads (20 µL per sample).
   ii. Collect the beads on a magnetic stand.
   iii. Wash twice with an equal volume of 1× binding buffer.

2. Resuspend the Dynabeads in an appropriate volume of 2× binding buffer (50 µL per sample). Aliquot the Dynabeads to each well/tube containing total RNA in 50 µL H$_2$O. Mix by pipetting.

3. Close the cap/seal the plate. Heat to 65°C for 2 min on a thermocycler with a heated lid.

4. Incubate at room temperature for 5–10 min with occasional shaking.

5. Collect and wash the beads twice with 150 µL of washing buffer.

6. Elute the mRNA by adding 50 µL of TE buffer and incubating at 80°C for 1 min. Immediately chill the plate on ice for 1 min.

7. Add 50 µL of 2× binding buffer. Repeat the binding and washing (Step 3 to Step 5).

8. Wash with 30 µL of prechilled 1× SuperScript III first-strand buffer to prevent carryover of salt and detergent.

   Lithium ion inhibits reverse transcriptase.

9. Resuspend the Dynabeads in 10 µL of 2× Superscript III buffer supplemented with 10 mM DTT.

10. Incubate at 94°C for exactly 6 min to fragment the mRNA and immediately place on ice.

    Fragmentation time can be modified to generate libraries of different insert size; see Step 34 for details.

11. Collect the Dynabeads on a magnetic stand. Transfer the solution containing fragmented mRNA to a new plate/tubes.

    This is a possible stopping point. Eluted RNA can be stored at −80°C.

First-Strand cDNA Synthesis

12. Assemble the following RT reaction:

    10 µL fragmented mRNA in 2× RT buffer
    0.5 µL random hexamer (1 µg/µL)
    0.5 µL RNasin Plus

13. Heat at 50°C for 1 min. Immediately place on ice. Add 9 μL of the following RT master mix:

    6.88 μL H_2O
    0.12 μL Actinomycin D (1 μg/μL)
    1 μL DTT (100 mM)
    0.5 μL dNTP (25 mM)
    0.5 μL SuperScript III

14. Perform the RT reaction:

    i. 25°C for 10 min
    ii. 50°C for 50 min

15. Immediately add 36 μL of RNAClean XP to each well/tube. Incubate the mixture on ice for 15 min.
    The solution is very viscous; pipette up and down at least 10 times to mix.

16. Collect the solid phase reversible immobilization (SPRI) beads on a magnetic stand.
    When using SPRI beads, always keep the PCR strip/plate on the magnet till the final elution step.

17. Wash twice with 75% ethanol without disturbing the beads.

18. Air-dry the beads for 5 min. Elute the RNA/cDNA hybrid with 10 μL H_2O.

Second-Strand Synthesis with dUTP

19. Prepare the second strand reaction master mix on ice as follows:

    1.5 μL Blue Buffer (10×) (or NEBuffer 2)
    1 μL dNTP mix (10 mM dATP, dCTP, dGTP, and dUTP)
    0.2 μL RNase H (5 U/μL)
    1 μL DNA polymerase I (10 U/μL)
    1.3 μL H_2O

20. Add 5 μL of the master mix to each 10 μL of RNA/cDNA. Incubate at 16°C for 2.5 h.
    The completed second strand reaction can be held in the PCR machine at 4°C overnight.

21. Purify double-stranded DNA (dsDNA) using 1.8 volumes of AMPure XP beads. Elute with 10 μL of H_2O.
    This is a possible stopping point. Eluted dsDNA can be stored at −20°C.
    See Steps 48–57 for the general procedure for using AMPure XP beads.

End-Repair

22. Prepare an appropriate amount of the end-repair master mix on ice as follows:

    1.5 μL End-Repair Buffer (10×) (or NEB PNK buffer)
    3 μL dNTP mix (1 mM)
    0.5 μL End-Repair Mix (Low Concentration)

23. Add 5 μL of the master mix to 10 μL of the dsDNA. Incubate at 20°C for 30 min.

24. Purify using 1.8 volumes of AMPure XP beads. Elute with 10 μL of H_2O.

dA-Tailing

25. Prepare an appropriate amount of the master mix on ice as follows:

    1.5 μL Blue Buffer (10×) (or NEBuffer 2)
    0.5 μL dATP (10 mM)
    2.5 μL H_2O
    0.5 μL Klenow 3′–5′ exo

Cold Spring Harbor Protocols

www.cshprotocols.org

26. Add 5 µL of the master mix to each sample and incubate at 37°C for 30 min.

27. Purify dsDNA using 1.8 volumes of AMPure XP beads. Elute with 8 µL of H_2O.

Y-Shape Adapter Ligation

28. Add 1 µL of the desired barcode adapter (5 µM) to each sample.

29. Prepare the master mix on ice as follows:

    8.5 µL Rapid Ligation Buffer (2×)
    0.5 µL T4 DNA Ligase HC (600 U/µL)

30. Add 9 µL of the master mix to each well. Mix by pipetting up and down. Incubate at room temperature for 15 min.

    This is a possible stopping point. The ligated library can be stored at −20°C.

Triple-SPRI Purification and Size Selection

31. Purify DNA using 1 volume (18 µL) of AMPure XP beads.

    Polyethylene glycol (PEG) and NaCl facilitate DNA binding to the SPRI beads. Because the ligation buffer already contains 7.5% PEG, less AMPure XP beads are used to prevent small adapter dimer binding.

32. Elute with 10 µL of TE buffer.

33. Purify the library again using exactly 1.4 volumes of AMPure XP beads (14 µL) and elute with 10 µL of TE buffer.

34. To perform size selection, add exactly 1 volume of AMPure XP beads (10 µL) to the eluted DNA and incubate at room temperature for 5 min.

    This is a critical step; an equal volume of beads will bind the large DNA fragments. A library without proper size fractionation will show a broader peak with a "tail" on the Bioanalyzer (Fig. 2A). If the fragmentation time is decreased in Step 10 to generate larger libraries, the amount of beads used in this step should be reduced accordingly. For example, larger insert libraries for PE sequencing could be generated by using 5 min fragmentation time (Step 10) and 0.75 volumes of SPRI beads for subtraction (Step 34).

35. Pull the beads to the side of the plate/tube on a magnetic stand. Carefully pipette the supernatant into a tube with 10 µL of AMPure XP beads.

36. Mix by pipetting. Incubate for 15 min at room temperature.

    This will bring up the NaCl and PEG concentration equivalent to two volumes of AMPure XP beads to facilitate DNA binding.

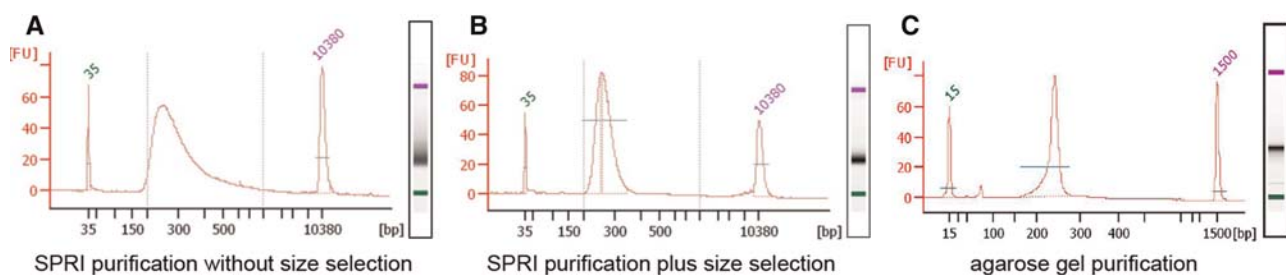37. Collect the beads on a magnetic stand. Wash twice with 75% ethanol without disturbing the beads.

FIGURE 2. The effect of SPRI size selection. (*A*) Library purified twice with SPRI-beads without size selection shows a broad peak on the Agilent Bioanalyzer. (*B*) The large insert libraries tend to cluster poorly on the flowcell and cause potential sequencing bias, whereas a sharp peak is observed from the library prepared by SPRI size fractionation as described in this protocol. (*C*) A Bioanalyzer image of a library prepared by agarose gel electrophoresis.

38. Elute the DNA with 10 µL water.

This is a possible stopping point. The size-fractionated DNA can be stored at −20°C.

PCR Enrichment

39. Digest the second strand DNA with 0.5 µL of UDG at 37°C for 15 min.

40. Prepare PCR reaction as follows:

5 µL UDG-digested DNA
0.2 µL primer A (10 µM)
0.2 µL primer B (10 µM)
4 µL Phusion HF Buffer (5×)
0.2 µL dNTP (10 mM)
9.9 µL H_2O
0.5 µL Phusion II

41. Perform initial denaturation at 94°C for 2 min, followed by 10–12 cycles of amplification (98°C for 10 sec, 65°C for 30 sec, 72°C for 30 sec).

This is a critical step. To avoid over-amplification, the PCR cycle should be optimized for each cell/tissue type. For many plant RNA samples: use 14–15 cycles for 0.5–1 µg total RNA, 10–12 cycles for 1–5 µg RNA, 8–10 cycles if >5 µg of RNA has been used.

42. Randomly select a few libraries and run 2 µL of their prepurified PCR product on an agarose gel.

The adapter dimer will be ∼100 bp (Fig. 3). After the second SPRI clean up, there should not be a strong adapter dimer band. The use of phosphorothioate base at the 3′ end of the PCR oligo will further suppress PCR amplification of the adapter dimers.

43. Purify using 1.4 volumes of AMPure XP beads. Elute with 20 µL of TE buffer.

A successful library preparation should yield >200 ng of amplified library.

Mix Barcoded Libraries for Multiplexed Sequencing

44. Measure the DNA concentration of each library using a PicoGreen dsDNA Assay Kit (96-well plate) or Qubit dsDNA HS Assay Kit (single-tube).

45. Combine equal amounts (e.g., 20 ng) of each barcode library.

46. Concentrate the library using 1.4 volume of AMPure XP and elute with 10 µL TE buffer.

Concentration >2 ng/uL is often sufficient.

47. Check the quality of the library on the Agilent Bioanalyzer. A sharp peak ∼200 bp is expected (Fig. 2B).

Over-amplification will result in double- and triple-sized peaks at 400 bp and 600 bp.

General Procedure for Using AMPure Beads

48. Add an appropriate amount of beads to the sample as indicated in the protocol.

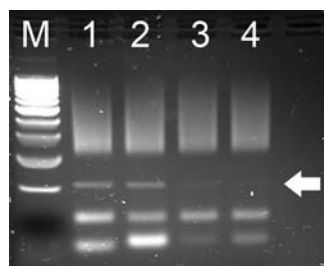**49.** Mix well by pipetting up and down 10 times.

FIGURE 3. Phosphorothioate-modified oligos suppress amplification of the adapter dimer. Adapter dimer containing crude adapter ligation product (with only one round of SPRI purification) is amplified by using SR PCR primers without (*lanes 1* and *2*), or with one (*lane 3*) or two (*lane 4*) phosphorothioate-modified base(s). The position of the adapter dimer is indicated by the arrow. *Lane M*: 100 bp DNA marker.

50. Incubate the plate/tube at room temperature for 10–15 min.

    Unless specified in the protocol, prolonged or low temperature incubation will result in binding of small DNA fragments such as adapter dimers.

51. Place on the magnetic stand for 2–5 min to collect the beads.

    Viscous solutions will require longer collection times.

52. Gently remove the solution without disturbing the bead pellets.

53. Add 200 μL of 75% ethanol.

    Keep the plate on the magnet.

54. Wait for 30 sec and remove the ethanol without disturbing the beads.

55. Repeat the 75% ethanol wash once more.

    Inspect the plate carefully and remove any remaining ethanol droplets.

56. Air-dry the beads. Add an appropriate amount of water (or TE buffer) to elute the DNA.

    Add water when small cracks start to appear on the bead pellet/ring. It is also important to prevent over-drying, because the beads will aggregate and reduce the yield.

57. Place on the magnet. Transfer the eluent to a new plate/tube.

DISCUSSION

Strand-Specific RNA-Seq vs. Conventional RNA-Seq

For comparison, we used the conventional Illumina RNA-Seq method to sequence a tomato RNA sample, which had also been sequenced using the ssRNA-Seq approach described here. We aligned reads from both the strand-specific and conventional RNA-Seq to the published full-length tomato cDNA sequences and derived digital expression for each gene. The effectiveness of ssRNA-Seq becomes obvious when the gene expression level is calculated only using reads aligned to the "correct" strand. The conventional RNA-Seq library overestimated 445 (3.55%) out of 12,539 tomato genes that we analyzed by at least 10% and 204 (1.61%) genes by 10-fold or more (Fig. 4A). When the ssRNA-Seq reads aligned to both strands of the annotated genes were counted, the Pearson correlation of digital gene expression between the two data sets is very high (r = 0.9935, Fig. 4B). Both sequencing methods generate a similar bell curve-like read distribution pattern (Fig. 4C), where RNA-Seq reads mapped across the transcript body and were depleted for both 5′ and 3′ ends as previously reported (Wang et al. 2009). Most importantly, it also shows that ssRNA-Seq data can be used and compared with conventional RNA-Seq data if the reads are allowed to align without considering the strand information. This is a significant advantage over other ssRNA-Seq methods, because it allows researchers to switch to more accurate strand-specific sequencing in ongoing projects, even if large amounts of data have already been generated by traditional RNA-Seq approaches.

False Antisense Read Derived from Uracil-Minus RNA

One potential drawback of any dUTP-based ssRNA-Seq methods is that an RNA molecule containing no uracil will inevitably escape the UDG digestion. We scanned full-length cDNA databases of *Arabidopsis* (20,683), rice (28,469), maize (27,455), tomato (13,227), and poplar (4664), for thymine-minus regions in a 200-nucleotide (nt) window, which is the average size of RNA-Seq libraries, and only two cDNAs from rice were found to contain such regions. We further scanned predicted transcripts in two well-annotated plant genomes, *Arabidopsis* (http://www.arabidopsis.org/) and rice (http://rice.plantbiology.msu.edu/), and only one and four transcripts respectively were found to contain such T-minus regions. While few in number, such rare transcripts could be flagged as a caution in subsequent data analyses.

Purification and Size Selection on Magnetic Beads

It is of critical importance to obtain a clean library of uniform insert size for high-density cluster generation. Large insert libraries and libraries of a broad distribution of insert size tend to yield many fewer clusters. In order to avoid the gel electrophoresis step in the original RNA-Seq method and still obtain a
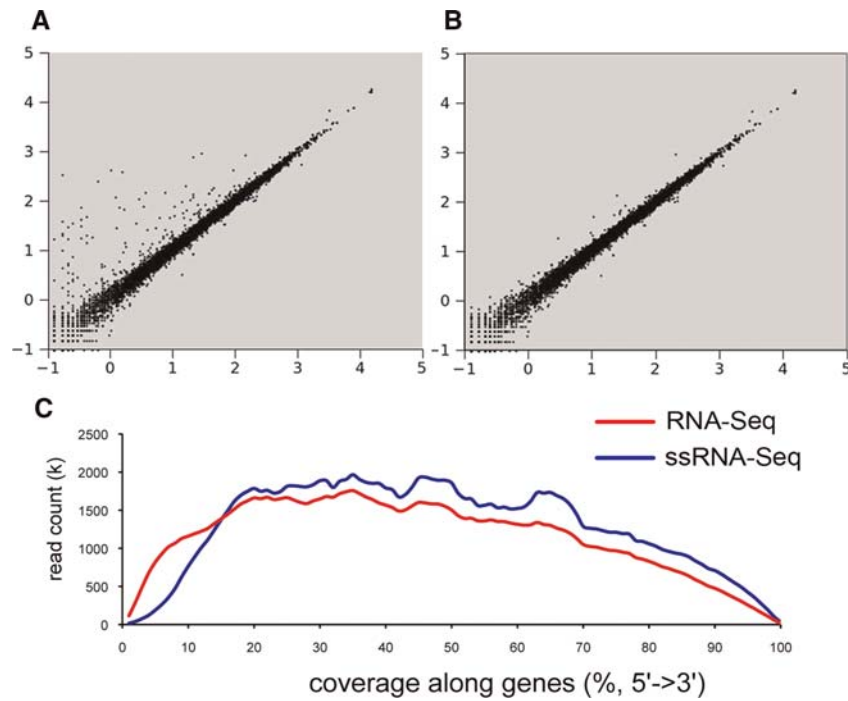
FIGURE 4. Comparison of strand-specific RNA-Seq and conventional RNA-Seq. (*A*) Scatter plot of digital gene expression values (log$_{10}$ transformed) between ssRNA-Seq (*x*-axis) and conventional RNA-Seq (*y*-axis) data. (*B*) The conventional RNA-Seq method overestimated the expression of genes associated with anti-sense RNA. If the alignment was performed without considering RNA strand information, the two libraries showed a high correlation. (*C*) Coverage plot of RNA-Seq reads at each percentile of a gene's length shows the bell-curve distribution pattern. The strand-specificity is obvious from the ssRNA-Seq read coverage that shows a shift toward the 3′ because the strand-specific sequencing starts at the 5′ end of the first-strand cDNA (3′ end of the RNA fragment after reverse-complementation), whereas the conventional library sequencing could randomly start from both ends.

uniform library, we subtract the undesired large DNA fragments from the library by using a modified SPRI-bead protocol, in which the PEG and NaCl concentration are optimized to remove large DNA fragments, while leaving the small libraries in the solution. In our sequencing facility, both SPRI and agarose gel size-selected libraries with a sharp peak at 200 bp could deliver 100–120 million pass-filter clusters on HiSeq2000, while larger or nonsize fractioned libraries normally yield ~80 million pass-filter clusters per HiSeq2000 lane.

We have replaced all column purification steps in the original Illumina method with SPRI beads, which enables preparation of up to 96 libraries with a plate magnetic collector. It could also be easily scaled down to prepare libraries on 8- or 12-well PCR strips. Due to the higher recovery rate of the SPRI beads than silica spin columns, this ssRNA-Seq protocol can construct libraries with less RNA materials. We routinely use 2–5 µg of tomato fruit RNA to construct ssRNA-Seq libraries, and we have tested as little as 500 ng total RNA as input with success. For these low input libraries, the PCR cycles should be proportionally increased as indicated in Step 41.

Different Multiplex Sequencing Strategies for GAII and HiSeq2000 Instruments

A key feature of the Illumina sequencing-by-synthesis technology is that DNA fragments are ligated with the Y-shaped adapter followed by asymmetric PCR to add unique tags to their 5′ and 3′ ends. One way to perform multiplex sequencing is to add extra nucleotides to the end of the Y-shaped adapter (Craig et al. 2008). All sequencing reads will start with the added nucleotides, which serve as the barcode, and are immediately followed by a "T" from the T/A overhang. On the previous GAIIx platform, libraries made with any Craig method adapters could be mixed and sequenced. However, on the HiSeq2000 instrument, such a dominant "T" signal from the Craig adapter T/A overhang will interfere with its base-calling and focusing system if it appears in the first five cycles. Hence, we have modified the barcode length of

the Craig method adapters to prevent having only "T" in one sequencing cycle. Libraries prepared with such 4- and 5-nt barcode adapters can then be mixed and sequenced on the HiSeq2000. It is also possible to use 6-nt, or longer, Craig barcode adapters for multiplexing, because HiSeq2000 uses the first five cycles for focusing. To (re)sequence existing Craig-type multiplex libraries incompatible with HiSeq2000, one could spike in the Illumina supplied PhiX control library or a new library prepared with a shorter barcode adapter to balance the base distribution in the first five cycles.

Illumina has a different multiplexing strategy, in which the barcode is "read" separately after the first sequencing read. It has the advantage that all libraries share one adapter because the barcode is added during PCR and they can be sequenced on HiSeq2000 without further optimization. (Note that the new Illumina TruSeq multiplex kit has barcode sequence incorporated into the adapter.) However, it cannot be run side-by-side with non-multiplexing libraries because it requires additional sequencing cycles to "read" the barcode.

Using Phosphorothioate Oligos to Prevent PCR Amplification of Adapter Dimers

Phosphorothioate base has been widely used for generating the T-overhang in the Y-shaped adapter, but its application in library amplification has not yet been fully exploited. We show that phosphorothioate nucleotides at the 3′ end of the PCR oligo could significantly reduce adapter dimer contamination by repressing their amplification (Fig. 3). Adapter contamination is often derived from self-ligated adapters without the T-overhang and PCR primers are able to anneal to them despite the fact that the last "T" on the primer cannot base-pair with the self-ligated adapter. However, proof-reading DNA polymerase with 3′ to 5′ exonuclease activity will consider the last "T" as a misincorporated base and cleave it to reinitiate DNA synthesis, leading to amplification of the adapter dimers. Such undesired amplification will be suppressed by using phosphorothioate-modified nucleotides at the 3′ end of the PCR oligo, which prevent the proof-reading DNA polymerase from removing the mismatched "T."

Concluding Remarks

Deep sequencing is transforming biology as a whole by empowering individual investigators with the sequencing capacity equivalent to what heretofore had been limited to genome sequencing centers. With dramatically increasing sequencing capacity and cost reductions, it is now feasible to use global RNA sequencing to generate a snapshot of the dynamic eukaryotic transcriptome at single nucleotide resolution. It has been demonstrated in many cases that strand-specific RNA-Seq can generate a more accurate gene expression profile than its predecessor and it is thus gaining popularity (Wang et al. 2009). However, its application is often limited by its laborious library preparation procedures and high construction/sequencing cost. Here we have presented a high-throughput (96 libraries per 2 d) and cost-effective (less than $10 per library) ssRNA-Seq method, which makes high volume strand-specific RNA sequencing feasible. Such improvements in ssRNA-Seq will help facilitate the implementation of this highly informative approach for gene expression studies in the near future while simultaneously allowing the incorporation and comparison of data derived from earlier methods.

ACKNOWLEDGMENTS

REFERENCES

Craig DW, Pearson JV, Szelinger S, Sekar A, Redman M, Corneveaux JJ, Pawlowski TL, Laub T, Nunn G, Stephan DA, et al. 2008. Identification of genetic variants using bar-coded multiplexed sequencing. *Nat Methods* **5:** 887–893.

Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A. 2009. Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res* **37:** e123. doi: 10.1093/nar/gkp596.

Vivancos AP, Güell M, Dohm JC, Serrano L, Himmelbauer H. 2010. Strand-specific deep sequencing of the transcriptome. *Genome Res* **20:** 989–999.

Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: A revolutionary tool for transcriptomics. *Nat Rev Genet* **10:** 57–63.

RECIPES

Recipes for items marked with <R> are provided here. Additional recipes can be found online at http://www.cshprotocols.org/recipes.

Binding buffer (2×)

1 M LiCl
20 mM Tris-HCl (pH 7.5)
2 mM EDTA
1% lithium dodecyl sulfate

Washing buffer for ssRNA-Seq

0.15 M NaCl
10 mM Tris-HCl (pH 7.5)
1 mM EDTA

Cold Spring Harbor Protocols

www.cshprotocols.org