Chapter 23

Next-Generation Sequencing Applied to Flower Development: RNA-Seq

Jun He and Yuling Jiao

Abstract

Genome-wide study of gene expression, or transcriptome profiling, is critical for our understanding of biological functions, including developmental processes. Recent breakthroughs in high-throughput sequencing technologies have revolutionized gene expression profiling to study the transcriptome at the nucleotide level, which is known as RNA-seq. RNA-seq, also called "whole transcriptome shotgun sequencing," uses next-generation sequencing technologies to sequence cDNA in order to infer a sample's RNA content. Here we describe a detailed bench-ready protocol to generate RNA-seq libraries for high-throughput single-end or pair-end sequencing compatible with the Illumina sequencing platform.

Key words Next-generation sequencing, RNA-seq, Transcriptome

1 Introduction

High-throughput gene expression study is central to many biological processes including flower development, which is under the control of complicated transcriptional regulatory networks [1]. Evolved from Southern blotting, DNA microarrays containing tens of thousands of DNA probe spots attached to a solid surface expanded gene expression profiling to the genome scale [2]. However, hybridization-based microarray methods have several limitations, including reliance on genome sequence information, high background levels owing to cross-hybridization, and a limited dynamic range of detection owing to both background and saturation of signals [3]. Moreover, it is hard to compare expression levels among different experiments [3]. Recent advances in next-generation sequencing technologies have made it possible to directly sequence the transcriptome [4–12]. Deep-sequencing based RNA-seq has overcome the limitations encountered by microarray technologies [13]. In addition, RNA-seq has the single-base resolution that is needed to study detailed features of each transcript, such as splicing isoforms, transcription start site(s), and

José Luis Riechmann and Frank Wellmer (eds.), *Flower Development: Methods and Protocols*, Methods in Molecular Biology, vol. 1110, DOI 10.1007/978-1-4614-9408-9_23,

polyadenylation site(s). In fact, RNA-seq has been rapidly adopted for many biological studies, including flower development [14–16]. However, the analysis of the data generated in RNA-Seq experiments can be complex, and the methods for data processing are still evolving. Whereas RNA-Seq data analysis is not described in this chapter, the reader is referred to several recent reviews addressing that topic [17–22].

2 Materials

Prepare all solutions with diethyl pyrocarbonate (DEPC)-treated water to avoid RNA degradation.

2.1 Purification of Total RNA

1. Nuclease-free mortar and pestle.
2. Liquid nitrogen.
3. Total RNA extraction kit or reagent (such as the RNeasy plant mini kit, Qiagen, Hilden, Germany).
4. DEPC-treated water: add 1 mL of DEPC to 1 L of distilled water. Mix thoroughly with magnetic stirrer overnight at room temperature. Autoclave and then cool to room temperature before use.

2.2 mRNA Purification from Total RNA

1. Dynabeads mRNA purification kit (Life Technologies, Carlsbad, CA, USA).
2. Magnetic stand (Promega, Madison, WI, USA).
3. Washing buffer: 10 mM Tris–HCl, 0.15 M LiCl, 1 mM EDTA, pH 7.5 at 25 °C.
4. Beads binding buffer: 20 mM Tris–HCl, 1 M LiCl, 2 mM EDTA, pH 7.5 at 25 °C.
5. 10 mM Tris buffer: 10 mM Tris–HCl, pH 7.5 at 25 °C.

2.3 Fragmentation of mRNA

1. 5× Fragmentation buffer: 200 mM Tris-acetate, 500 mM potassium acetate, 150 mM magnesium acetate.
2. Stop buffer: 200 mM EDTA, pH 8.0 at 25 °C.
3. Glycogen (20 μg/μL, Life Technologies, Carlsbad, CA, USA).
4. 3 M NaOAC, pH 5.2 at 25 °C.
5. 70 % ethanol, dilute with DEPC-treated water.

2.4 cDNA Synthesis

1. Superscript II reverse transcriptase (200 U/μL, Life Technologies, Carlsbad, CA, USA).
2. 5× First-strand buffer (Life Technologies, Carlsbad, CA, USA).
3. 0.1 M dithiothreitol (DTT, Life Technologies, Carlsbad, CA, USA).

4. Random primers (3 μg/μL, Life Technologies, Carlsbad, CA, USA).
5. RNase inhibitor (40 U/μL, such as RiboLock RNase inhibitor, Fermentas, Burlington, ON, Canada).
6. 10× Second-strand buffer: 500 mM Tris–HCl, 50 mM $MgCl_2$, 10 mM DTT, pH 7.8 at 25 °C.
7. 25 mM dNTP mix.
8. RNase H (5 U/μL, Fermentas, Burlington, ON, Canada).
9. DNA polymerase I (10 U/μL, Enzymatics, Beverly, MA, USA).
10. PCR purification kit (such as Zymoclean DNA clean & concentrator-5, Zymo, Orange, CA, USA).

2.5 End Repair

1. 10× End repair buffer: 50 mM Tris–HCl, 10 mM $MgCl_2$, 1 mM ATP, 10 mM DTT, pH 7.5 at 25 °C.
2. T4 DNA polymerase (3 U/μL, New England, Biolabs, Ipswich, MA, USA).
3. Klenow DNA polymerase (5 U/μL, New England, Biolabs, Ipswich, MA, USA).
4. T4 polynucleotide kinase (10 U/μL, New England, Biolabs, Ipswich, MA, USA).

2.6 dA Addition

1. 10× A-Tailing buffer (i.e., NEB Buffer 2, New England, Biolabs, Ipswich, MA, USA).
2. 1 mM dATP.
3. Klenow (3′–5′exo-) (5 U/μL, Enzymatics, Beverly, MA, USA).

2.7 Adaptor Ligation

1. 2× Rapid T4 DNA ligase buffer (Enzymatics, Beverly, MA, USA).
2. T4 DNA ligase (600 U/μL, Enzymatics, Beverly, MA, USA).
3. NEBNext adapter for Illumina (15 μM, New England, Biolabs, Ipswich, MA, USA).
4. USER enzyme (1 U/μL, New England, Biolabs, Ipswich, MA, USA).

2.8 Gel Purification and Size Selection

1. Low-Range ultra agarose.
2. TAE buffer.
3. 6× DNA Orange loading dye: 10 mM Tris–HCl, 0.15 % Orange G, 0.03 % xylene cyanol FF, 60 % glycerol, 60 mM EDTA, pH 7.6 at 25 °C.
4. 100 bp DNA ladder.

5. Ethidium bromide solution, 2 μg/mL.
6. Gel purification kit (such as the QIAquick gel extraction kit, Qiagen, Hilden, Germany).

2.9 PCR Amplification

1. NEBNext universal PCR primer (25 μM) (New England, Biolabs, Ipswich, MA, USA).
2. NEBNext index primers (25 μM) (New England, Biolabs, Ipswich, MA, USA).
3. 5× Phusion high fidelity PCR buffer (New England, Biolabs, Ipswich, MA, USA).
4. Phusion high fidelity DNA polymerase (2 U/μL, New England, Biolabs, Ipswich, MA, USA).
5. 25 mM dNTPs.

3 Methods

Keep samples on ice to minimize RNA degradation unless otherwise specified. Perform all centrifugation steps at 4 °C, unless indicated otherwise. Wear gloves and use sterile techniques when working with RNA. All glassware and plasticware should be RNase-free.

3.1 Purification of Total RNA

Total RNA can be purified from floral tissue by standard methods (*see* **Note 1**). The starting total RNA should be of high quality as determined by gel electrophoresis or by using a Bioanalyzer, and quantified by using a UV spectrometer such as NanoDrop 2000c (*see* **Note 2**).

3.2 mRNA Purification from Total RNA

1. Dilute a minimum of 5 ng total RNA with RNase-free water to 50 μL in a non-sticky tube, heat the sample at 65 °C for 5 min to disrupt RNA secondary structure and then place the sample on ice.
2. Mix well the Dynabeads oligo$(dT)_{25}$ suspension with a vortex at 4 °C, aliquot 15 μL of Dynabeads oligo$(dT)_{25}$ into a 1.5 mL non-sticky tube.
3. Wash the beads twice with 100 μL beads binding buffer each time and remove the supernatant by using a magnetic stand (*see* **Note 3**).
4. Resuspend the beads in 50 μL beads binding buffer, add the 50 μL total RNA sample from **step 1**, and rotate the tube at room temperature for 5 min. Remove the supernatant by using a magnetic stand.
5. Wash the beads twice with 200 μL washing buffer each time and remove the supernatant by using a magnetic stand.

6. Add 50 μL of 10 mM Tris–HCl to the beads and mix well gently by finger tapping.
7. Heat the sample at 80 °C for 2 min, then immediately put the tube on a magnetic stand. Transfer the supernatant containing mRNA to a new RNase-free non-sticky Eppendorf tube and add 50 μL beads binding buffer.
8. Heat the sample at 65 °C for 5 min to disrupt RNA secondary structure then place the sample on ice.
9. Wash the beads from **step 7** twice with 200 μL washing buffer each and remove the supernatant.
10. Add 100 μL sample from **step 8** to the beads from **step 9**, mix well by finger tapping, and then rotate for 5 min at room temperature. Remove the supernatant by using a magnetic stand.
11. Wash the beads twice with 200 μL washing buffer each and remove the supernatant by using a magnetic stand.
12. Add 17 μL of 10 mM Tris–HCl to the beads, mix well gently by finger tapping, heat the sample at 80 °C for 2 min to elute the mRNA from the beads, and then immediately put the tube on the magnetic stand. Transfer the supernatant containing mRNA to an RNase-free 200 μL thin-wall PCR tube. The resulting volume of mRNA should be approximately 16 μL.

3.3 Fragmentation of mRNA

1. Add 5× fragmentation buffer (4 μL) to the mRNA sample (16 μL).
2. Incubate the tube in a PCR thermal cycler at 94 °C for 3–4 min (*see* **Note 4**).
3. Add 2 μL stop buffer and place the tube on ice.
4. Transfer the solution to a 1.5 mL RNase-free non-sticky Eppendorf tube and add 2 μL NaOAc (3 M, pH 5.2), 2 μL glycogen, and 60 μL ethanol (100 %) to the tube. Incubate at −80 °C for 30 min.
5. Centrifuge the tube at 17,000 × *g* for 25 min at 4 °C in a microcentrifuge.
6. Pipette out the supernatant carefully without dislodging the RNA pellet, then wash the pellet with 300 μL of 70 % ethanol.
7. Centrifuge the tube at 17,000 × *g* for 20 min at 4 °C and carefully pipette out the supernatant.
8. Dry the pellet for 5 min in a SpeedVac or 10 min in a laboratory hood.
9. Resuspend the mRNA pellet in 11.1 μL of RNase-free water.

3.4 cDNA Synthesis

1. Add 1 μL random primers (3 μg/μL) to the mRNA (11.1 μL) and incubate the sample in a PCR thermal cycler at 65 °C for 5 min, then place the tube on ice.

2. Assemble the first strand cDNA synthesis mix in a PCR tube at RT:

mRNA w/random primers	12.1 μL
5× First-strand buffer	4 μL
0.1 M DTT	2 μL
25 mM dNTP mix	0. 4 μL
RNase inhibitor (40 U/μL)	0.5 μL

The final volume should be 19 μL.

3. Incubate the sample at 25 °C in a thermal cycler for 2 min.
4. Add 1 μL Superscript II reverse transcriptase (200 U/μL) to the sample and run the following PCR program:
   (a) 25 °C for 10 min.
   (b) 42 °C for 50 min.
   (c) 70 °C for 15 min.
   (d) Hold at 4 °C.
5. Place the tube on ice, and add 62.8 μL of water to the first strand cDNA synthesis mix.
6. Add 10 μL of 10× second-strand buffer and 1.2 μL of dNTP mix (25 mM), mix well and incubate on ice for 5 min.
7. Add 1 μL of RNase H (2 U/μL) and 5 μL of DNA polymerase I (10 U/μL), mix well and incubate at 16 °C in a thermal cycler for 2.5 h.
8. Purify the sample using a DNA clean and concentrator kit and elute in 50 μL of water.

3.5 End Repair

1. Assemble the end repair mix in a 1.5 mL reaction tube on ice:

Eluted DNA	50 μL
10× End repair buffer	10 μL
25 mM dNTP mix	1.6 μL
T4 DNA polymerase (3 U/μL)	5 μL
Klenow DNA polymerase (5 U/μL)	1 μL
T4 polynucleotide kinase (10 U/μL)	5 μL
Nuclease-free water	27.4 μL

The final volume should be 100 μL.

2. Incubate the sample in a thermal cycler at 20 °C for 30 min.
3. Purify the sample using a DNA clean and concentrator kit and elute in 32 μL water.

3.6 dA Addition

1. Assemble the dA adding mix in a 1.5 mL reaction tube on ice:

Eluted DNA	32 μL
A-tailing buffer	5 μL
1 mM dATP	10 μL
Klenow (3′–5′exo-) (5 U/μL)	3 μL

   The final volume should be 50 μL (*see* **Note 5**).
2. Incubate the sample in a thermal cycler at 37 °C for 30 min.
3. Purify the sample using a DNA clean and concentrator kit and elute in 23 μL water.

3.7 Adaptor Ligation

1. Assemble the adaptor ligation mix in a PCR reaction tube on ice:

Eluted DNA	23 μL
2× Rapid T4 DNA ligase buffer	25 μL
Adapter (15 μM)	1 μL
T4 DNA ligase (600 U/μL)	1 μL

   The final volume should be 50 μL.
2. Incubate the sample in a thermal cycler at 20 °C for 15 min.
3. Add 3 μL USER enzyme (1 U/μL), mix well by pipetting, and then incubate at 37 °C for 15 min.
4. Purify the sample using a DNA clean and concentrator kit and elute in 10 μL water.

3.8 Gel Purification and Size Selection

1. Prepare a 2 % low-range agarose gel with 1× TAE.
2. Load the samples as follows (*see* **Note 6**):
   (a) 1 μL DNA ladder in the first well.
   (b) 10 μL DNA elute from 3.7 mixed with 2 μL of 6× DNA loading dye (with Orange G) in the second well.
   (c) 1 μL DNA ladder in the third well.
3. Run the gel at 100 V for 35–40 min or until Orange G reaches the bottom of the gel.
4. Post-stain in ethidium bromide solution (2 μg/mL) for 20 min, then destain in distilled water for 20 min.
5. Place the gel on top of a plastic wrap on a UV box. Be sure to minimize exposure time, and cut the gel at 200 bp (±25 bp) (*see* **Note 7**).
6. Purify the sample using a QIAquick gel extraction kit and elute in 37 μL of elution buffer.

3.9 PCR Amplification

1. Assemble the PCR reaction mix in a 200 μL thin wall PCR tube:

Eluted DNA	37 μL
Universal PCR primer (25 μM)	1 μL
Index primer (X) (25 μM) (*see* **Note 8**)	1 μL
5× Phusion high fidelity PCR buffer	10 μL
25 mM dNTP mix	0.5 μL
Phusion high fidelity DNA polymerase (2 U/μL)	0.5 μL

   The final volume should be 50 μL.

2. Run the PCR with the following program (*see* **Note 9**):
   (a) 98 °C for 10 s.
   (b) 98 °C for 10 s.
   65 °C for 30 s.
   72 °C for 30 s.
   Repeat (b) for 15 cycles.
   (c) 72 °C for 5 min.
3. Purify the PCR product using a DNA Clean Kit and elute in 10 μL elution buffer. This product is the sample library, ready for sequencing (*see* **Note 10**).

4 Notes

1. We recommend the use of RNeasy plant mini kit, TRIzol reagent, or similar RNA extraction kits, and to follow manufacturer's protocols for plant materials. Ensure that the working areas, reaction tubes, and tips are free of RNase contaminants. For floral buds, we routinely start from a 0.1 to 0.5 mL volume of packed tissues, and obtain 50–500 μg total RNA.
2. If a Bioanalyzer is used to check RNA integrity, it is recommended that only samples with an RNA Integrity Number (RIN) value greater than 8 are used for library construction. For a typical RNA-Seq library construction experiment, the use of 5 ng to 10 μg total RNA is recommended. Within that range, using a larger amount of starting total RNA is recommended to reduce the number of PCR amplification cycles during library preparation, which will result in more even distribution of mapped reads within target genes. If the starting amount of total RNA is less than 5 ng, exponential or linear RNA amplification can be used [15, 23].

3. Mix well the beads with buffer by gentle shaking using finger tapping. When removing the supernatant, put the tube on a magnetic stand for about 2 min and remove the buffer without taking the tube out from the magnetic stand.
4. Fragment length distribution depends on the incubation time. A 3-min incubation is recommended if a 400 bp band will be excised in Subheading 3.7, and a 4 min incubation for a 200 bp band.
5. Please note that dATP is very sensitive to freeze-thaw cycles. To avoid freeze-thawing of dATP stocks, store in small aliquots.
6. Using ladders on both sides of the sample lanes helps to locate the gel area to be excised as the DNA sample may not be visible. When running two samples, leave one empty lane between samples and ladders to prevent cross-contamination. Do not run more than two samples on the same gel to avoid cross contamination.
7. The target DNA size may range from 200 to 600 bp depending on your experimental need. Longer target DNA size (such as 400 bp) is recommended for paired-end sequencing.
8. Indexed PCR primers are needed if multiple samples will be combined in the same sequencing lane. The NEBNext multiplex oligos for Illumina contains 12 index primers, each with a different index. For each reaction, only one of the 12 primer indices is used for one library. Make sure to combine samples amplified using different indexed primers into the same sequencing lane.
9. PCR cycle number can be adjusted from 4 to 25 depending on the amount of starting material. For example, use 4 cycles of amplification if starting with 5 μg or more total RNA, 18 cycles for 100 ng, and 25 cycles for 5 ng of starting total RNA.
10. A sample library can be quality controlled by using an Agilent 2100 Bioanalyzer with the Agilent DNA 1000 chip kit to check the size, purity, and concentration of the sample. The concentration of a sample library can be quantified using a Life Technologies Qubit fluorometer. The identity of a sample library can be confirmed by cloning 1 μL of the library into a blunt-end cloning vector, and sequence using Sanger sequencing. For *Arabidopsis*, at least 10 M mapped 50 bp single-end reads are needed to obtain a reasonable coverage of the transcriptome for gene expression quantification. Longer reads, such as 100 bp, can increase mapping accuracy by reducing multi-mapping reads. To study alternative splicing, paired-end sequencing with deeper coverage is recommended [24, 25].

Acknowledgements

We thank B.A. Williams and B.J. Wold for sharing protocols. This work was supported by the Ministry of Agriculture of China (Grant 2011ZX08010-002), by the Knowledge Innovation Program of CAS Grant KYQY-150, and by the Hundred Talents Program of CAS.

References

1. Krizek BA, Fletcher JC (2005) Molecular mechanisms of flower development: an armchair guide. Nat Rev Genet 6:688–698
2. Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467–470
3. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63
4. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 18:1509–1517
5. Morin R, Bainbridge M, Fejes A, Hirst M, Krzywinski M, Pugh T, McDonald H, Varhol R, Jones S, Marra M (2008) Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. Biotechniques 45:81–94
6. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5:621–628
7. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. Science 320:1344–1349
8. Cloonan N, Forrest AR, Kolle G, Gardiner BB, Faulkner GJ, Brown MK, Taylor DF, Steptoe AL, Wani S, Bethel G, Robertson AJ, Perkins AC, Bruce SJ, Lee CC, Ranade SS, Peckham HE, Manning JM, McKernan KJ, Grimmond SM (2008) Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat Methods 5:613–619
9. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell 133:523–536
10. Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, Penkett CJ, Rogers J, Bahler J (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature 453:1239–1243
11. Ozsolak F, Milos PM (2011) RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 12:87–98
12. Martin JA, Wang Z (2011) Next-generation transcriptome assembly. Nat Rev Genet 12: 671–682
13. Malone JH, Oliver B (2011) Microarrays, deep sequencing and the true measure of the transcriptome. BMC Biol 9:34
14. Jiao Y, Meyerowitz EM (2010) Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. Mol Syst Biol 6:419
15. Schmid MW, Schmidt A, Klostermeier UC, Barann M, Rosenstiel P, Grossniklaus U (2012) A powerful method for transcriptional profiling of specific cell types in eukaryotes: laser-assisted microdissection and RNA sequencing. PLoS One 7:e29685
16. Yang H, Lu P, Wang Y, Ma H (2011) The transcriptome landscape of Arabidopsis male meiocytes from high-throughput sequencing: the complexity and evolution of the meiotic process. Plant J 65:503–516
17. Oshlack A, Robinson MD, Young MD (2010) From RNA-seq reads to differential expression results. Genome Biol 11:220
18. Kvam VM, Liu P, Si Y (2012) A comparison of statistical methods for detecting differentially expressed genes from RNA-seq data. Am J Bot 99:248–256
19. Fang Z, Martin JA, Wang Z (2012) Statistical methods for identifying differentially expressed genes in RNA-Seq experiments. Cell Biosci 2:26
20. Garber M, Grabherr MG, Guttman M, Trapnell C (2011) Computational methods for transcriptome annotation and quantification using RNA-seq. Nat Methods 8:469–477
21. Gao D, Kim J, Kim H, Phang TL, Selby H, Tan AC, Tong T (2010) A survey of statistical software for analysing RNA-seq data. Hum Genomics 5:56–60

22. Cumbie JS, Kimbrel JA, Di Y, Schafer DW, Wilhelm LJ, Fox SE, Sullivan CM, Curzon AD, Carrington JC, Mockler TC, Chang JH (2011) GENE-counter: a computational pipeline for the analysis of RNA-Seq data for gene expression differences. PLoS One 6:e25279
23. Tang F, Barbacioru C, Nordman E, Li B, Xu N, Bashkirov VI, Lao K, Surani MA (2010) RNA-Seq analysis to capture the transcriptome landscape of a single cell. Nat Protoc 5:516–535
24. Marquez Y, Brown JW, Simpson C, Barta A, Kalyna M (2012) Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. Genome Res 22:1184–1195
25. Filichkin SA, Priest HD, Givan SA, Shen R, Bryant DW, Fox SE, Wong WK, Mockler TC (2010) Genome-wide mapping of alternative splicing in Arabidopsis thaliana. Genome Res 20:45–58