Chapter 10

Small RNA Library Construction from Minute Biological Samples

Jessica A. Matts, Yuliya Sytnikova, Gung-wei Chirn, Gabor L. Igloi, and Nelson C. Lau

Abstract

Increasingly, the discovery and characterization of small regulatory RNAs from a variety of organisms have all required deep-sequencing methodologies. However, the crux to successful deep-sequencing analysis depends upon optimal construction of a cDNA library compatible for the high-throughput sequencing platform. Challenges to small RNA library constructions arise when dealing with minute tissue samples because certain structural RNA fragments can dominate and mask the desired characterization of regulatory small RNAs like microRNAs (miRNAs), endogenous small interfering RNAs (endo-siRNAs), and Piwi-interacting RNAs (piRNAs). Here, we describe methods that improve the chances of constructing a successful library from small RNAs isolated from minute tissues such as enriched follicle cells from the *Drosophila* ovarium. Because the ribosomal RNA (rRNA) fragments are frequently the major contaminants in small RNA preparations from minute amounts of tissue, we demonstrate the utility of antisense oligonucleotide depletion and an acryloylaminophenylboronic acid (APB) polyacrylamide gel system for separating the abundant 2S rRNA in *Drosophila* from endo-siRNAs and piRNAs. Finally, our methodology generates libraries amenable to multiplex sequencing on the Illumina Hi-Seq platform.

Key words Small RNAs, Illumina deep-sequencing, Library construction

1 Introduction

Among the various methods to detect small endogenous RNAs in organism samples, the procedure to convert RNAs into a library of cDNAs followed by sequencing is considered to be the most comprehensive technique because it enables discovery of the molecules' sequence as well as confirmation of its identity. For example, the seminal discoveries that expanded the catalog of microRNAs beyond lin-4 and let-7 resulted from the sequencing of cDNA libraries from small RNAs purified from nematode, fly and human cells [1–3]. With the commercialization and adoption of high-throughput next generation

Jessica A. Matts and Yuliya Sytnikova have equally contributed to this work

Mikiko C. Siomi (ed.), *PIWI-Interacting RNAs: Methods and Protocols*, Methods in Molecular Biology, vol. 1093,
DOI 10.1007/978-1-62703-694-8_10,

sequencing technologies such as 454, Illumina, and ABI SoLID, the detection of molecules based on the number of reads have lent to unparalleled sensitivity and accuracy in small RNA quantitation.

Despite the immense depth of sequencing now achievable by high-throughput systems like the Illumina Hi-Seq 2000 capable of routinely generating >200 million reads per flow cell lane, the success of a small RNA sequencing run depends most on the composition of the library in terms of being free of undesired degradation remnants or naturally abundant short ribosomal RNA fragments. For example, *Drosophila* and other insects generate an abundant 31-nucleotide (nt) 2S rRNA that migrates very closely to the piRNAs that range from 24 to 32 nt long [4]. In addition, *Xenopus* eggs also contain a slew of ribosomal RNA fragments that range from 23 to 35 nt long, which also co-migrate in size with *Xenopus* piRNAs and miRNAs [5]. When there are plenty of tissues and cells from which to generate a cell lysate, this lysate can be subjected to immunoprecipitation (IP) of ARGONAUTE and PIWI proteins, or with cation exchange chromatography [6], both of which can quite effectively deplete undesired contaminating RNAs.

However, when one wishes to profile small RNAs from minute samples, such as a single egg or from a very small population of cells enriched in a particular cell type, the IP and chromatography methods are not practical, and typically just total RNA is isolated. From these total RNA preparations, rRNA fragments persist and will become a major nuisance. For example, when total small RNAs from single *Xenopus* eggs were profiled, the rRNA fragments reduced the representation of other small regulatory RNAs down to below 20 % of the library [5]. This issue can be partially mitigated by sequencing libraries on the Illumina Hi-Seq platform versus the Illumina Genome Analyzer (GA) platform because the ~20-fold increase in depth from the former platform may yield enough desired small regulatory RNA reads despite sacrificing the non-useful rRNA contaminants. However, new considerations in the multiplexing of small RNA libraries must be followed for the Illumina Hi-Seq platform because the lower stringency of base-calling versus the GA platform can also reduce the yield of reads passing quality. In the first set of methods detailed in this chapter, we describe our experiences in isolating a small sample of tissues, an enriched population of follicle cells from the *Drosophila* ovarium, and our considerations in generating small RNA libraries from these minute samples for the Illumina Hi-Seq platform, which includes a new format of linkers that are amenable for multiplex sequencing of small RNA libraries.

In the second set of methods, we describe the adaptation of a boronate affinity gel matrix applied to the resolution of small amounts of RNA from *Drosophila* ovary cells. The boronate gel matrix consists of a denaturing polyacrylamide gel impregnated with acryloylaminophenylboronic acid (APB), and short RNAs (<~80 nt) with unmodified 2′–3′ cis-diols will exhibit a stronger dynamic affin-

ity to the boronate than endo-siRNAs and piRNAs which are naturally methylated at the 3′ end by Hen-1 on the 2′ OH [7, 8]. With standard polyacrylamide electrophoresis, the abundant 2S rRNA and other rRNA fragments can co-migrate or resolve poorly from piRNAs and endo-siRNAs. However, on an APB gel the rRNA fragments are retarded while bonafide piRNAs and endo-siRNAs migrate faster, thus facilitating further the removal of the contaminating RNAs from the regulatory small RNAs.

As biologists begin to interrogate the small RNA profiles of particular niches of cells, from stem cells to specialized neuronal cell types, the need to improve methodologies to generate libraries from minute samples will become more evident. These procedures we have developed will increase the likelihood that properly diverse cDNA libraries can be constructed, and although our reagents are based on lab-made stocks, the antisense oligo-mediated depletion step and the boronate affinity gel matrix can be applied to steps from commercial small RNA library construction protocols.

2 Materials

2.1 Drosophila Culture and Ovary Dissection

1. Standard fly food in bottles with enough extra yeast added to make a fine layer on the top of the food.
2. Flynap or carbon dioxide venting flypad for anesthetizing flies.
3. Stereo dissection microscope.
4. Paintbrush for fly pushing.
5. Fine pointed Dumont 50 tweezers.
6. Watch glass dish chilled in a small box of ice.
7. Chilled 1× PBS.

2.2 Follicle Cell Enrichment

1. Trypsin (Sigma).
2. 4′,6-diamidino-2-phenylindole (DAPI) stain and Vectashield (Vector labs).
3. 1 mg/ml Concanavalin A solution (Sigma).
4. 1 % Formaldehyde in 1× PBS (for fixing follicle cells).
5. 3.7 % Formaldehyde in 1× PBS (for fixing ovaries).
6. Programmable tube mixer (i.e. Thermomixer, Eppendorf).

2.3 Small RNA Purification and Antisense Oligo Depletion of 2S rRNA

1. TriReagent or TriZol solutions (Molecular Research Center or Invitrogen).
2. 2× RNA Urea-TE loading buffer: 8 M Urea, 100 mM Tris–HCl (pH 8.0), 50 mM EDTA, 0.05 % w/v bromophenol blue, 0.05 % w/v xylene cyanol FF.
3. 10 bp ladder (Invitrogen).
4. Gamma-^{32}P-labeled markers (18mer and 34mer).

5. SyberGreen II RNA stain (Invitrogen).
6. Phosphorimaging cassette and phosphorimager (i.e., Typhoon Scanner, GE Healthcare).
7. Siliconized 1.7-mL microcentrifuge tubes.
8. 0.3 M NaCl.
9. 10 μM biotinylated DNA oligo antisense to 2S rRNA: TACA ACCCTCAACCATATGTAGTCCAAGCATACAACCCTCA ACCATATGTAGTCCAAGCAGTCGA-3′biotin.
10. 0.5 M EDTA (pH8.0).
11. 20× SSC: 3 M NaCl, 0.3 M sodium citrate (pH 7.0).
12. 0.5 M NaCl.
13. Glycogen from blue mussel, 20 mg/mL (Sigma).

2.4 Acryloylamino-phenylboronic Acid Polyacrylamide Gel

1. Acryloylaminophenylboronic acid (APB, tRNA Probes, Inc.) [originally produced by the protocol of Ref. 9].
2. Sequagel Urea gel system (National Diagnostics).
3. 50× Tris-Acetate–EDTA (TAE) buffer: 2 M Tris-Acetate, 50 mM EDTA (pH 8.0).
4. Vertical electrophoresis system (i.e., Model V16 from Whatman, Inc.).
5. 2× Urea-TE loading buffer: 8 M Urea, 10 mM Tris–HCl (pH 8.0), 50 mM EDTA, 0.05 % w/v bromophenol blue, 0.05 % w/v xylene cyanol FF.
6. Synthetic RNA Marker oligonucleotides of any sequence at 18, 22, and 32 nt long.
7. Synthetic piRNA with standard terminal 2′-OH: AGGAA AGUUGUGCACACUUGUAAUCCGAAA.
8. Synthetic piRNA with terminal 2′-*O*-Me: UGGGAUUACA AGUGUGCACAACUUUCCUGmC.

2.5 Linker-Ligation Small RNA Library Construction from Minute Samples

1. Chimeric marker oligonucleotides where uppercase bold represents an RNA base, lowercase italics is a standard DNA base, and "dU" is deoxy-uridine, which can be degraded by uracil-deoxyglycosylase (UDG):

   34 nt chimeric marker: **CAGUAC** *ggatcca* dUdUdUdU *tat-gctc* **AGCGUACGAA**

   18 nt chimeric marker: **C** agtac dUdUdU gctag **CUAA.**
2. 3′ Linker adaptor: p-*cgtcgtatgccgtcttctgcttgt*-/3AmMO/, where /3AmMO/ is a 3′ amino modifier and is chemical phosphorylated on the 5′ end.
3. Original 5′ Adaptor: **GUUCAGAGUUCUACAGUCCG-ACGAUC**.

4. 5′ Hi-Seq adaptors with barcodes underlined, RNA bases in uppercase bold, DNA bases in lowercase italics, and "N" represents random incorporation of all four DNA bases:
   Barcode CAA: gttcagagttctacagtccgacgatc *NNN* CAA**AA**
   Barcode ACC: gttcagagttctacagtccgacgatc *NNN* ACC**AA**
   Barcode GUU: gttcagagttctacagtccgacgatc *NNN* **GUUAA**
   Barcode UGG: gttcagagttctacagtccgacgatc *NNN* **UGGAA**.
5. Reverse Transcription primer & 5′ PCR primer: caagcagaagacggcata.
6. 3′ PCR primer: aatgatacggcgaccaccgacaggttcagagttctacagtccga.
7. Mth RNA ligase (New England Biolabs).
8. Polynucleotide kinase (PNK) (New England Biolabs).
9. Uracil-DNA glycosylase (New England Biolabs).
10. Phusion Polymerase (New England Biolabs).
11. T4 RNA ligase I enzymes (New England Biolabs).
12. RiboLock RNase inhibitor, 40 U/μl (Thermo Scientific).
13. 5× RNA ligase buffer: 250 mM HEPES (pH 8.3), 50 mM $MgCl_2$, 16.5 mM DTT, 50 μg/mL BSA, 41.5 % glycerol.
14. 10 mM ATP.
15. Low melting temperature agarose (i.e., Agarose II, Amresco).
16. Gel extraction kit (Qiagen).
17. RNA Clean & Concentrator kit (Zymo Research).
18. SuperScript III Reverse Transcriptase (Invitrogen).
19. Quant-iT Pico-green dsDNA assay kit (Invitrogen).
20. Zero-Blunt TOPO PCR cloning kit for sequencing (Invitrogen).

3 Methods

3.1 Harvesting Ovaries and Enriching for Follicle Cells

About 3 days prior to dissections, flies should be given extra yeast in the food in order to fatten up the ovaries so that they are easily distinguishable and removable from the rest of the abdomen (*see* **Note 1**). Typically, ~400 female flies are dissected in one sitting to obtain ~100 μl of tissue. Follicle cell enriching must be performed immediately after one dissection sitting, and practice is required to enable obtaining tissue efficiently. The amount of trypsin to add varies depending on amount of tissue (usually 7.5 μl of a 10× trypsin stock from Sigma per 100 μl of tissue yields a good enriched follicle cell sample).

1. Anesthetize fattened flies with either flynap or carbon dioxide, and sort for females (larger yet lighter abdomen).
2. Transfer females to a watch glass containing cold 1× PBS. Use fine tweezers to tease out the ovaries by pulling away from the abdomen and place ovaries in a microcentrifuge tube containing ice-cold 1× PBS.

3. Wash ovaries once in 1 mL cold 1× PBS (save a 20 µl drop for DAPI staining).
4. Add 700 µl of cold 1× PBS to the ovaries, then add 77 µl of a 10× trypsin stock. Vortex briefly to mix.
5. Place tube in a thermoshaker to shake at 1,400 rpm (or the maximum rpm) for 20 min at 30 °C. Shaking must be vigorous (>1,200 rpm) to ensure that the majority of the nurse cells and oocytes are properly lysed during enzymatic digestion.
6. Rest the tube on ice to allow undigested tissue to settle and remove the supernatant of liberated follicle cells to a new tube which is centrifuged at 4 °C for 8 min at 2,000 × *g* to pellet the follicle cells.
7. Repeat **steps 4–6** at least two more times on remaining undigested tissues to collect additional enriched follicle cell samples. The final remaining tissue mostly consists of remaining mostly consists of fertilized eggs (up to 70 %) whose egg shells are resistant to trypsin digestion and can be discarded.
8. Wash and centrifuge the follicle cell pellet once in ice-cold 1× PBS (2 min, 1,600 × *g*), and resuspend pellet in 300 µl of cold 1× PBS. Remove 5 µl for DAPI staining verification. Spin down cells, remove PBS, and store the cell pellet at −80 °C until ready to start the small RNA library construction.
9. To verify efficacy of follicle cell enrichment, drops of dissected ovaries and follicle cells are placed on a microscope glass slide previously coated with Concanavalin A, which promotes cell adhesion to glass. After 5 min at room temperature, cells are fixed with formaldehyde in PBS for 5 min, then dipped in a 50 mL solution of 10 µg/ml DAPI in 1× PBS in a coplin jar for 5 min. Cells and ovaries are then mounted in Vectashield with a cover slip and observed under a fluorescent microscope with a DAPI filter to visualize the clear difference in size and morphology between intact ovaries and follicle cells (Fig. 1c).

3.2 Small RNA Purification and Antisense Oligo-Mediated 2S rRNA Depletion

Typically, two rounds of dissections and follicle cell enrichment would yield 10 µg of total RNA. However, we have been able to perform library construction from as little as 1–3 µg of total RNA [5]. Total RNA is extracted using TriReagent (also known as Trizol), and the very small pellet is resuspended in 20 µl of water.

1. Small RNAs are purified by gel purification from a 15 % urea denaturing gel by electrophoresis at 30 W for 75 min or until the darker Bromphenol blue dye is near the bottom of the gel. Total RNA is first denatured in 2× Urea Loading buffer with heating at 95° for 5 min, and samples are run with a 10 bp ladder and/or radioactive RNA markers.
2. If using radioactive RNA markers, expose gel covered in plastic wrap against a phosphor plate for 15 min. If only a 10 bp DNA

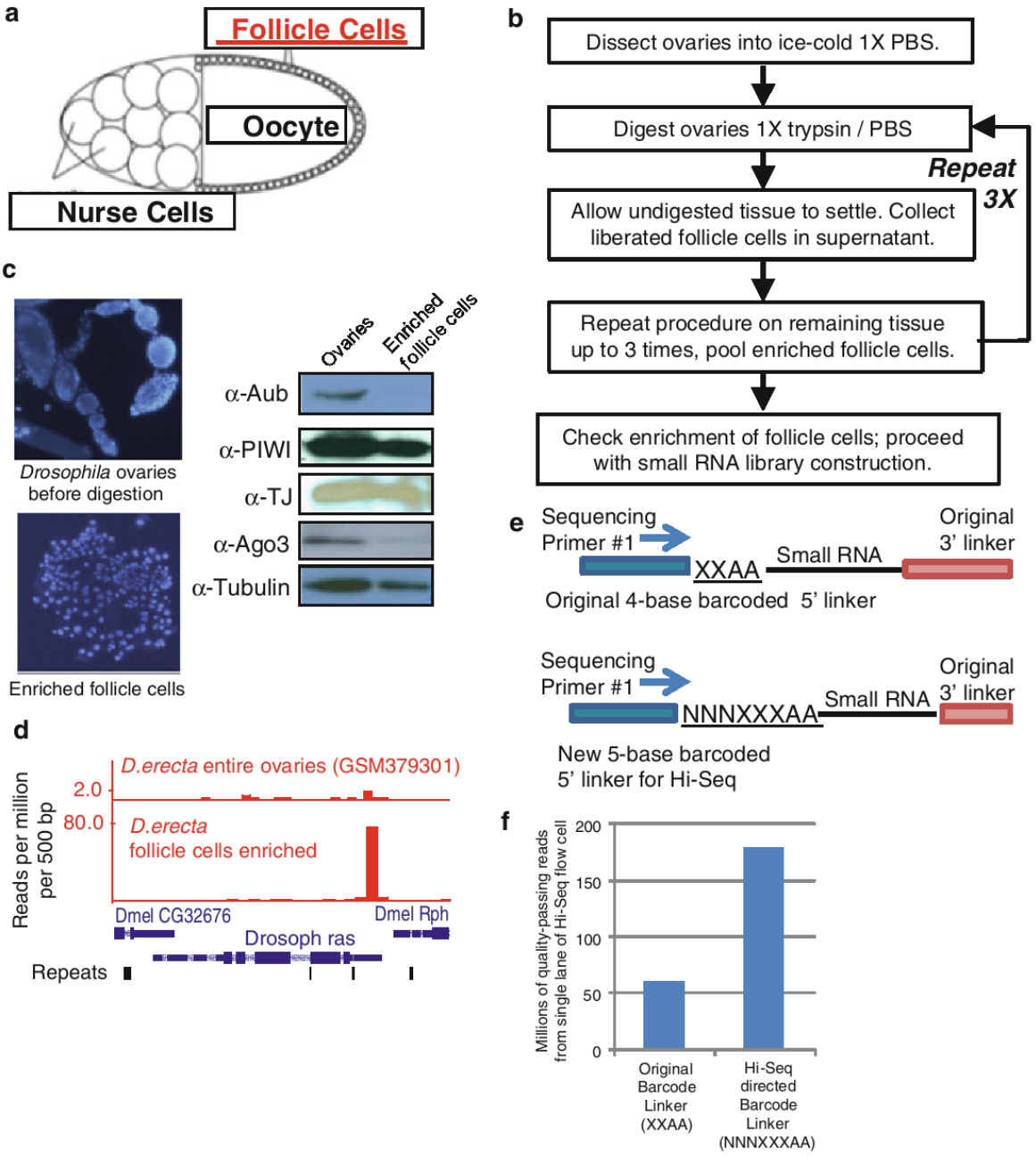

Fig. 1 *Drosophila* ovary follicle cell enrichment and small RNA library construction from this minute tissue sample. (**a**) A schematic of a *Drosophila* egg chamber highlighting the relative small size of the follicle cells. (**b**) Flowchart of the procedure to enrich follicle cells from *Drosophila* ovaries. (**c**) Validation of follicle cell enrichment procedure with 4′,6-diamidino-2-phenylindole (DAPI) staining in left panels and western blots that show retention of Piwi and Traffic Jam (TJ) proteins concomitant with depletion of Aubergine (Aub) and Argonaute-3 (Ago3) proteins. (**d**) Counts of genic 3′UTR-directed piRNAs from the *Raspberry* (*ras*) gene are enriched from the follicle cells compared to sequencing from total ovary RNA. (**e**) Schematic comparing the original 5′ linker barcode format that was suitable for sequencing on the Illumina Genome Analyzer and the new barcode format we developed for the Illumina Hi-Seq platform. (**f**) A small RNA library with the original 4-base barcode suffers from poor read quality discrimination, while the simple addition of a 3-base random element in the new 5-base barcode greatly increases the number of quality-passing reads

ladder is used, stain the gel with SyberGreen II for 10 min. Phosphor plate or gel is directly scanned on the image scanner (phosphorimaging or fluorescent setting). Print a full-scale picture of the gel, and placing this behind the gel, cut out the region of small RNAs from ~20 to 30 nt (the 2S rRNA migrates at 31 nt).

3. Elute the gel samples overnight at 4 degrees C in 500 µl of 0.3 M NaCl.
4. Precipitate the RNA in 2 volumes of absolute ethanol. 1 µg of mussel glycogen greatly enhances precipitation of small RNAs, while using siliconized tubes reduces nonspecific adherence of RNA to plastic.
5. Resuspend the RNA in 25 µl of dH_2O.
6. Antisense oligo-mediated depletion of 2S rRNA begins by prebinding oligo to the 25 µl RNA with the following: 10 µl of anti-2S rRNA biotin oligo (10 µM), 2 µl of 0.5 M EDTA (pH 8.0), 10 µl of 20× SSC, and 68 µl of dH_2O. Heat the sample to 70 °C for 5 min and slowly cool to 37 °C.
7. Meanwhile, prepare 600 µl per sample of MagneSphere streptavidin beads by washing according to manufacturer's instruction with 0.5× SSC, and add the 100 µl beads suspension to the RNA/oligo mix.
8. Incubate at room temperature for 20 min with mild agitation (can be placed on a rotisserie rocker to rotate).
9. Bind beads to a magnetic stand for 3 min, remove the supernatant to a new tube, add 200 µl of 0.5 M NaCl, 1 µl of glycogen, and 1 ml of absolute ethanol to supernatant and precipitate overnight at –20° C.
10. Next day pellet RNA and resuspend in 10 µl of dH_2O for library construction.

3.3 Boronate Affinity Gel Electrophoresis to Resolve piRNAs from Other Small RNAs

Another way to purify piRNAs from 2S rRNA and other abundant rRNA fragments is to exploit the differential affinity of modified versus unmodified terminal 3′-ends of nucleic acids to special gel matrixes during electrophoresis. For example, an acrylamide gel matrix impregnated with a boronate group can react in a dynamic equilibrium with the 3′-terminal 2′ and 3′ hydroxyl groups of unmodified short RNAs (Fig. 2a), but this reaction does not occur when there is a 2′-*O*-methyl group. Boronate groups are incorporated by adding APB to a standard denature urea polyacrylamide gel [9]. Short RNAs of similar lengths but with 3′-terminal modifications will not sufficiently resolve simply by size on a standard polyacrylamide gel (Fig. 2b, c), however when APB is added to 10 % of the acrylamide concentration (e.g., for a 20 % polyacrylamide gel we add APB to 2 %), the piRNAs with a 3′-terminal 2′-*O*-methyl group now migrate much faster in the gel and separate very effectively from unmodified rRNAs. Electrophoresis with

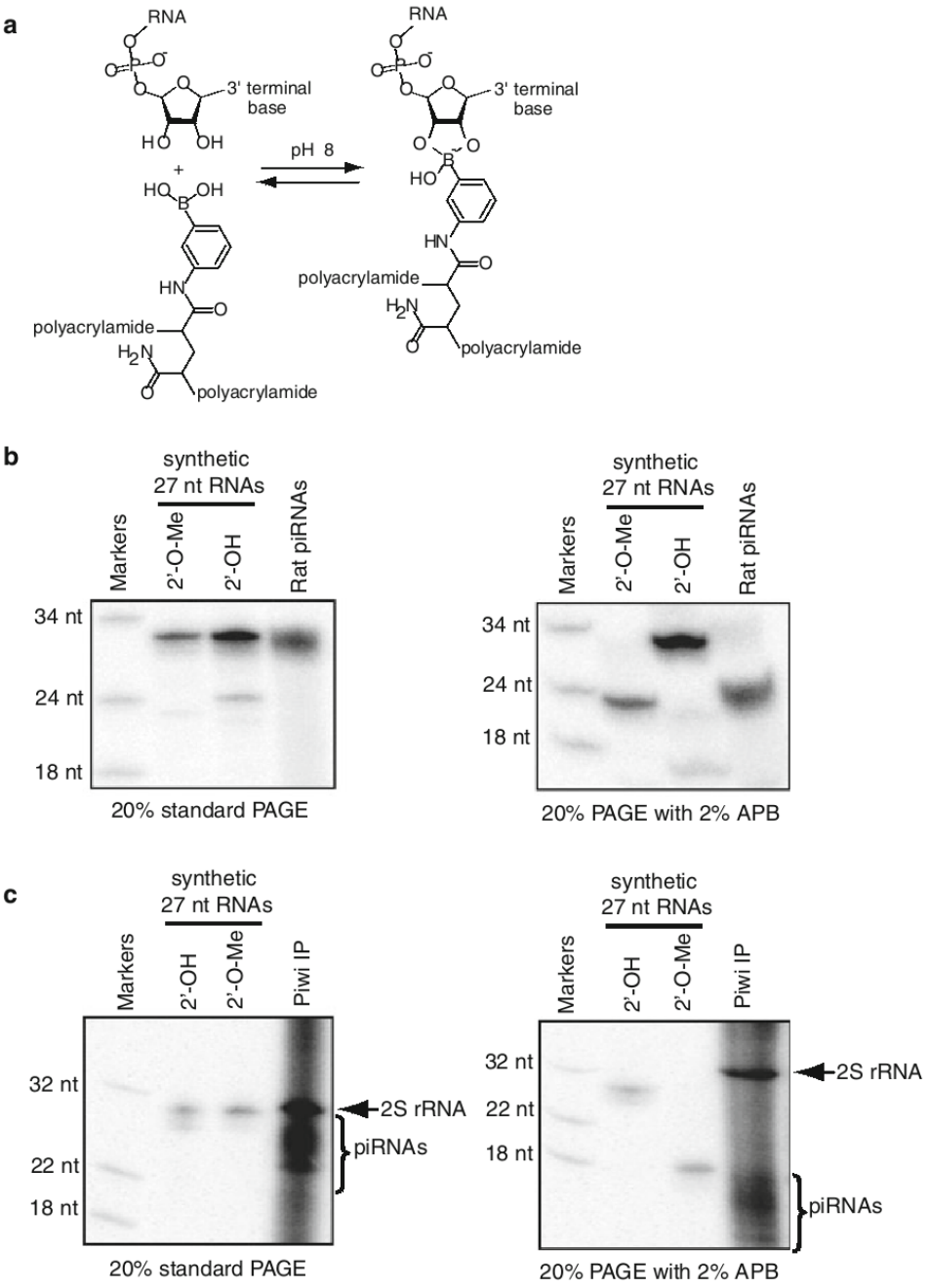

Fig. 2 Boronate affinity gel electrophoresis of small regulatory RNAs. (**a**) Scheme describing the structure of acryloylaminophenylboronic acid (APB) linked to polyacrylamide, and the reversible, dynamic reaction between boronate and the 3′ end of unmodified RNAs such as rRNA fragments. Small regulatory RNAs such as endogenous siRNAs and Piwi-interacting RNAs are methylated on the 3′-terminal 2′-OH, which prevents them from reacting with boronate. (**b**) 5′ radiolabeled synthetic RNAs with a 3′-terminal 2′-OH or 2′-*O*-Me electrophoresed along with radiolabeled rat piRNAs on a standard and APB polyacrylamide gel. The methylated synthetic RNA and rat piRNAs migrate much faster than their typical size when compared to the unmodified RNAs. (**c**) The mobility of *Drosophila* piRNAs are increased in the APB gel, allowing for better resolution of the piRNAs from the contaminating and highly abundant 2S rRNA that co-precipitate nonspecifically in a Piwi immunoprecipitation (IP). The *Drosophila* piRNAs from OSS cells were derived from a Piwi antibody IP performed with protein A/G magnetic beads followed by RNA extraction with TRI Reagent RT

APB gels must utilize Tris-Acetate (TAE) buffer instead of the typical Tris-Borate (TBE) buffer used for polyacrylamide gels because excess borate ions perturb complex formation between the boronate and the RNA.

1. Prepare a 20 % APB-polyacrylamide mix for a 30 ml, 0.8 mm thick gel with 24 ml of Sequagel Concentrate, 3 ml of Sequagel Diluent, 1.5 ml of a 50× TAE stock, 600 mg of APB powder, and 1 ml of dH_2O.
2. Warm the mix to 60 °C till APB powder dissolves completely and then cool back to room temperature.
3. Add 120 μl of 10 % APS and 12 μl of TEMED, mix and pour into the glass plates and spacers, and let the gel polymerize for at least 1 h at room temperature.
4. In our control experiments which are optional for actual library construction, we electrophoresed the same RNA samples on a standard Urea-TAE gel versus an APB gel to demonstrate the different mobility behaviors of piRNAs relative to synthetic RNA markers.
5. Natural RNAs were dephosphorylated with phosphatase, extracted with phenol and chloroform, and then ethanol precipitated.
6. Natural and synthetic RNAs are then labeled with ^{32}P-γ-ATP and PNK and denatured in Urea Loading Dye for 5 min at 95 °C prior to gel loading. Note in our demonstration that on a standard polyacrylamide gel, residual 2S rRNAs persisting in a Piwi immunoprecipitation do not resolve well from piRNAs, however on an APB gel, piRNAs can now be resolved several millimeters away from 2S rRNA (Fig. 2c). *See* **Note 2** regarding being careful not to overload the maximum binding capacity of the APB gel.

3.4 Generating Small RNA Libraries from Minute RNA Quantities for Multiplexing on the Illumina Hi-Seq Platform

Several modifications of our original small RNA library construction protocol [1] have been described to generate libraries for deep-sequencing on the Illumina Genome Analyzer [4, 5, 10–18]. We further extend the list of modifications with a protocol geared towards preparations for the Illumina Hi-Seq sequencer and starting with a small quantity of RNA. A different format for barcoding and multiplexing in the 5′ linker adapter is required because the Hi-Seq machine's colony-calling algorithm now demands significant base diversity in the first 4 bases sequenced, as evidenced by our previous 4-base barcode immediately downstream sequencing primer #1 yield 1/5th of the theoretical maximum number of reads (~60 million versus ~300 million, Fig. 1e). The introduction of three random nucleotides before a 5 bp barcode sequence creates sufficient color diversity for the Hi-Seq colony-calling algorithm to approach sufficient number of quality-passing reads

(Fig. 1e, f). *See* **Note 3** on additional considerations that motivated the modification described here. The steps below highlight additional specific modifications to include in other previously described small RNA library construction procedures.

1. Previously we relied upon chemical adenylation as the most cost-effective method to generate pre-adenylated 3′ Linker adaptors, but the organic chemistry steps are not routine procedures for molecular biologists. We now employ enzymatic adenylation of the 3′ Linker adaptor with the Mth RNA ligase in the 5′-Adenylation Kit sold by New England Biolabs, which nearly quantitatively adds an adenylated moiety to the 5′ end of a monophosphorylated oligonucleotide.
2. Although gel-purified RNAs from passive elution in NaCl is typically concentrated by ethanol precipitation with glycogen carrier in siliconized tubes, for very small samples we have observed decreased loss of RNA by using the RNA Clean & Concentrator kit from Zymo Research, following the manufacturer's protocol for total RNA (>17 nt). We also recommend increasing the volume of ethanol 3× for the mix in the first two steps of the standard manufacturer's protocol. Elutions are performed with 20 μl of dH_2O.
3. The 5′ Linker adaptor ligation reaction components and conditions remain the same as previous protocols except the use of the new barcoded 5′ Linker adaptors for the Hi-Seq sequencer. When one has included the new chimeric RNA/DNA markers of 18 and 34 nt length with internal deoxy-uracils as spiked-in radiolabeled markers to facilitate visualizing the shift of ligated products via gel-electrophorese and phosphorimaging, there is the optional post-reaction step of adding 5 units of uracil DNA glycosylase to the 5′ Linker-ligation reaction and incubating for 20 min at 37 °C.
4. Reverse transcription (RT) with Superscript III reverse transcriptase and PCR with Phusion polymerase can be directly performed from ¼ or ½ of the 5′ Linker-ligation reaction without the need of purifying the ligated molecules from unligated molecules, following the same protocol as previous methods [4, 5, 15–18]. However, a common problem in amplifying a library that began from minute starting RNA is that linker-linker dimers can dominate in the final PCR, regardless of whether one performs or skips the gel purification of the ligated products after the 5′ Linker adapter ligation reaction. By reducing the number of PCR cycles (no more than 15 cycles) in the initial amplification rounds from the RT reaction, this PCR sample can be gel purified on a 4 % low melting temperature agarose gel for the desired 95–115 nt long amplicons from the linker-linker dimer that is ~80 nt long.

5. From this purification, an additional 15 cycles of PCR can be applied to generate enough material for small scale cloning with a standard PCR cloning vector like in the Zero-Blunt TOPO PCR cloning kit followed by Sanger sequencing verification. Purified PCR-amplified DNA can be suitably quantitated for concentration with the Quant-iT Pico-green dsDNA assay kit. This library sample is thus ready for deep sequencing onto the Hi-Seq sequencer.

4 Notes

1. Female flies that have been fattened up with yeast prior to ovary dissection yield larger egg chambers and facilitate the manual dissection of ovaries, which in turn enhances the harvest of follicle cells. Our enrichment procedure significantly depletes the oocytes and nurse cells in the *Drosophila* egg chamber, presumably from loss of structural integrity after trypsinization, however all cell-cycle stages of follicle cells are isolated in our procedure. For more extensive purifications of specific staged follicle cells, *see* refs. 19, 20.
2. The boronate affinity gel was first described as a means to resolve and detect modifications on transfer RNAs [9]. These authors noted the limitation of the boronate gel in terms of overloading the capacity of the gel and the maximum resolving abilities for shorter RNAs. In our experiments, we found the maximum capacity of a 1 mm thick APB gel with a 1 cm wide well was 2 μg of total RNA after which the retention of unmodified RNAs was skewed. When dealing with minute tissue samples, however, total RNA is typically limiting, such as where a single follicle cell enriched sample might typically yield 5 μg of total RNA. When electrophoresing an RNA sample on the APB gel, we recommend resolving no more than 2 μg of RNA for a 1 cm wide, 0.8 mm thick gel to avoid overloading the capacity of the APB. Since modified RNAs run very differently compared to unmodified RNAs of the same size, we recommend that a standard synthetic RNA and a 2′-*O*-methylated RNA oligo of similar size to piRNAs be radiolabeled and run alongside your organismal RNA sample to guide where to cut out the gel region of interest from which to elute the RNA in 0.5 M NaCl for downstream small RNA library construction.
3. Although sequencing chemistry for the Illumina platform continues to improve in providing longer read lengths to over 100 nt, the fidelity and quality of reading the first couple of bases extending from a sequencing or indexing primer are still considerably higher than readings towards the end of the molecule. Thus, we have chosen to place our linker barcode sequence at the first set of bases to be read by the small RNA

sequencing primer #1, in contrast to other methodologies that place the barcode sequence in the 3′ Linker [12, 13]. On the Illumina Genome Analyzer, the platform uses approximately 8 bases of read information to fine tune the assign and distinguish the "molecule colonies" on the flow cell lane, and our original 4 base barcode was suitable for multiplexing. However, the Hi-Seq platform only utilizes the initial 4–6 bases, and thus finds the original 4 base barcode too low in complexity to distinguish colonies, thus drastically reducing the theoretical numbers of quality-passing reads by threefold (Fig. 1f).

Acknowledgments

J.A.M and Y.S. contributed equally to this work. We are grateful to Kuniaki Saito and Haruhiko and Mikiko Siomi for antibody reagents to *Drosophila* Piwi proteins and to Dorothea Godt for the antibody to TJ. We thank Michael Rosbash and his laboratory members for access to the Illumina sequencer. This work was supported by the National Institutes of Health (Core Facilities Grant P30 NS045713 to the Brandeis Biology Department and R00HD057298 to N.C.L.). N.C.L. is a Searle Scholar.

References

1. Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Science 294:858–862
2. Lee RC, Ambros V (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. Science 294:862–864
3. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. Science 294:853–858
4. Lau NC, Robine N, Martin R, Chung WJ, Niki Y, Berezikov E, Lai EC (2009) Abundant primary piRNAs, endo-siRNAs, and microRNAs in a Drosophila ovary cell line. Genome Res 19:1776–1785
5. Lau NC, Ohsumi T, Borowsky M, Kingston RE, Blower MD (2009) Systematic and single cell analysis of Xenopus Piwi-interacting RNAs and Xiwi. EMBO J 28:2945–2958
6. Lau NC (2008) Analysis of small endogenous RNAs. Curr Protoc Mol Biol. Chapter 26, Unit26 7
7. Saito K, Sakaguchi Y, Suzuki T, Siomi H, Siomi MC (2007) Pimet, the Drosophila homolog of HEN1, mediates 2′-O-methylation of Piwi-interacting RNAs at their 3′ ends. Genes Dev 21:1603–1608
8. Horwich MD, Li C, Matranga C, Vagin V, Farley G, Wang P, Zamore PD (2007) The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr Biol 17:1265–1272
9. Igloi GL, Kossel H (1987) Use of boronate-containing gels for electrophoretic analysis of both ends of RNA molecules. Methods Enzymol 155:433–448
10. Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R (2008) Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. Genes Dev 22:2773–2785
11. Chiang HR, Schoenfeld LW, Ruby JG, Auyeung VC, Spies N, Baek D, Johnston WK, Russ C, Luo S, Babiarz JE, Blelloch R, Schroth GP, Nusbaum C, Bartel DP (2010) Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev 24:992–1009
12. Hafner M, Landgraf P, Ludwig J, Rice A, Ojo T, Lin C, Holoch D, Lim C, Tuschl T (2008) Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. Methods 44:3–12
13. Hafner M, Renwick N, Farazi TA, Mihailovic A, Pena JT, Tuschl T (2012) Barcoded cDNA library

preparation for small RNA profiling by next-generation sequencing. Methods 58:164–170
14. Ruby JG, Stark A, Johnston WK, Kellis M, Bartel DP, Lai EC (2007) Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. Genome Res 17:1850–1864
15. Ro S, Yan W (2010) Small RNA cloning. Methods Mol Biol 629:273–285
16. Thomas MF, Ansel KM (2010) Construction of small RNA cDNA libraries for deep sequencing. Methods Mol Biol 667:93–111
17. Havecker ER (2011) Detection of small RNAs and microRNAs using deep sequencing technology. Methods Mol Biol 732:55–68
18. Donovan WP, Zhang Y, Howell MD (2011) Large-scale sequencing of plant small RNAs. Methods Mol Biol 744:159–173
19. Claycomb JM, Benasutti M, Bosco G, Fenger DD, Orr-Weaver TL (2004) Gene amplification as a developmental strategy: isolation of two developmental amplicons in Drosophila. Dev Cell 6:145–155
20. Bryant Z, Subrahmanyan L, Tworoger M, LaTray L, Liu CR, Li MJ, van den Engh G, Ruohola-Baker H (1999) Characterization of differentially expressed genes in purified Drosophila follicle cells: toward a general strategy for cell type-specific developmental analysis. Proc Natl Acad Sci U S A 96:5559–5564