

Probe-Directed Degradation (PDD) for Flexible Removal of Unwanted cDNA Sequences from RNA-Seq Libraries

Stuart K. Archer,^{1,2,5} Nikolay E. Shirokikh,^{1,3,5} and Thomas Preiss^{1,4}

¹Genome Biology Department, The John Curtin School of Medical Research (JCSMR), The Australian National University, Acton, Canberra, Australian Capital Territory, Australia

²Present address: Monash Bioinformatics Platform, Monash University, Clayton, Victoria, Australia

³Present address: Moscow Regional State Institute of Humanities and Social Studies, Ministry of Education of Moscow Region, Kolomna, Moscow Region, Russia

⁴Victor Chang Cardiac Research Institute, Darlinghurst (Sydney), New South Wales, Australia

⁵These authors contributed equally to this work.

Most applications for RNA-seq require the depletion of abundant transcripts to gain greater coverage of the underlying transcriptome. The sequences to be targeted for depletion depend on application and species and in many cases may not be supported by commercial depletion kits. This unit describes a method for generating RNA-seq libraries that incorporates probe-directed degradation (PDD), which can deplete any unwanted sequence set, with the low-bias split-adaptor method of library generation (although many other library generation methods are in principle compatible). The overall strategy is suitable for applications requiring customized sequence depletion or where faithful representation of fragment ends and lack of sequence bias is paramount. We provide guidelines to rapidly design specific probes against the target sequence, and a detailed protocol for library generation using the split-adaptor method including several strategies for streamlining the technique and reducing adapter dimer content. © 2015 by John Wiley & Sons, Inc.

Keywords: RNA-seq • removal of rRNA • cDNA library • duplex-specific nuclease • PDD

How to cite this article:

Archer, S.K., Shirokikh, N.E., and Preiss, T. 2015. Probe-directed degradation (PDD) for flexible removal of unwanted cDNA sequences from RNA-seq libraries. *Curr. Protoc. Hum. Genet.* 85:11.15.1-11.15.36.
doi: 10.1002/0471142905.hg1115s85

INTRODUCTION

The described method for sequence-specific depletion of RNA-seq libraries relies on complementarity matching of unwanted nucleic acids with specifically designed oligonucleotide probes, and use of Duplex-Specific Nuclease (DSN) for elimination of the perfectly annealed duplexes (Archer et al., 2014). The method includes design of oligonucleotide probes to target degradation of specified sequences, construction of representative split-adaptor single-stranded circular cDNA library, annealing of the circular cDNA with the probes followed by DSN treatment (the PDD itself), amplification and analysis of the efficiency and selectivity of depletion in the resulting rectified cDNA (outlined in Fig. 11.15.1A). Several important solutions are implemented to overcome major intrinsic difficulties and limitations that occur otherwise, such as bioinformatics tools for



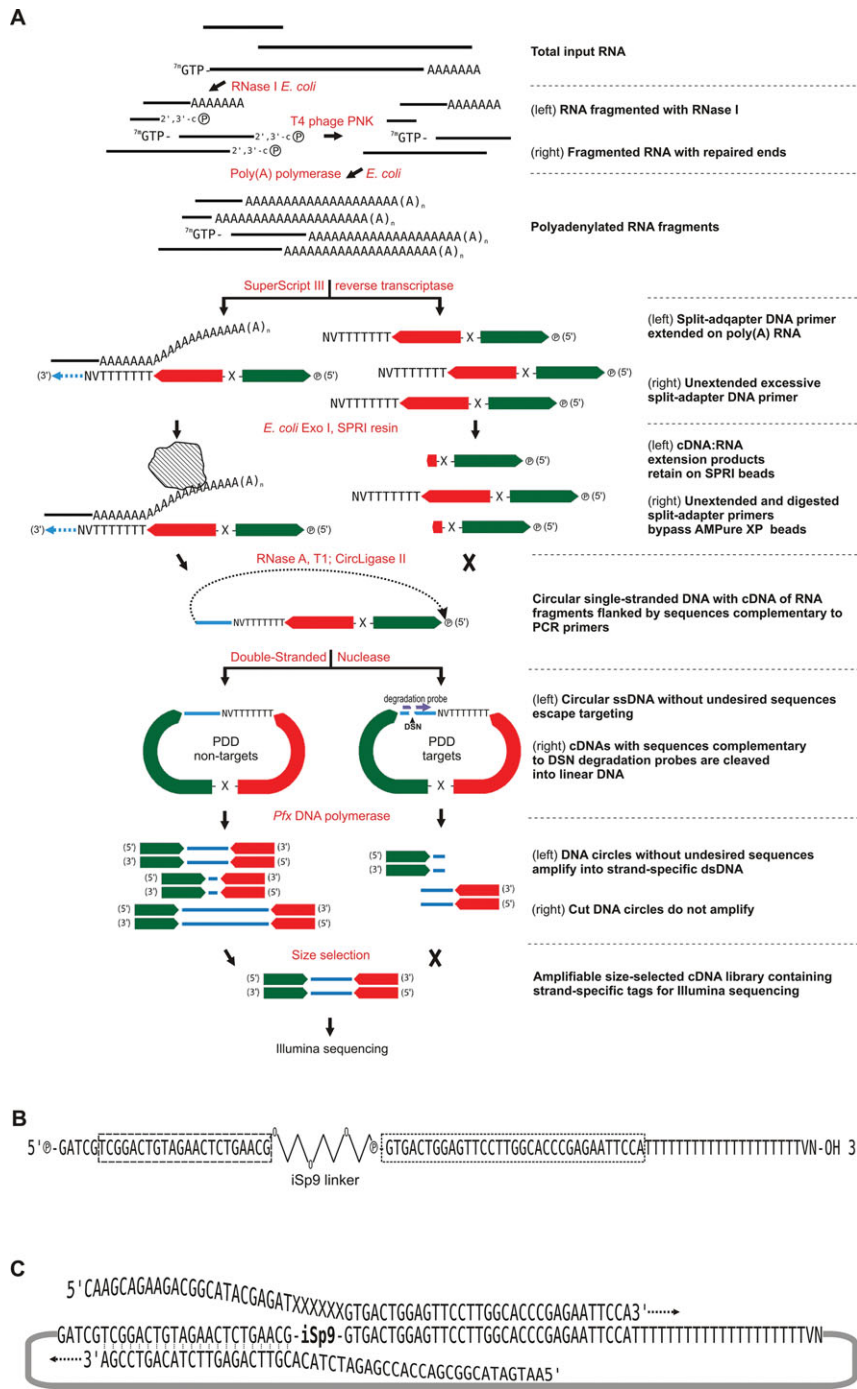


Figure 11.15.1 (A) Schematic illustration of the PDD method starting from total RNA sample and finishing with a depleted library compatible with high-throughput sequencing, as described in the unit. Structures of the nucleic acids involved in each main step of the library construction are outlined on the left together with enzymes and procedures required to move on to the next protocol step (denoted by the arrows). Text on the right describes features of the library molecules important to perform the PDD. PDD variant with the use of *E. coli* RNase I fragmentation of initial RNA is shown. Cleavage site of the Duplex-Specific Nuclease (DSN) is indicated with an arrow on the PDD targets circular DNA molecules with annealed PDD degradation probes. **(B)** Structure of the split-adaptor primer used for first-strand cDNA synthesis and library circularization in the PDD method. 5' phosphate, 3' hydroxyl and triethylene glycol spacer groups are shown. Regions required for annealing of the forward (dotted line) and reverse (dashed line) primers are indicated. **(C)** Schematic representation of synthesis of linear, strand-specific indexed products from a circular template in the course of library amplification by PCR.

generation of low-bias sets of degradation probes, prevention of split-adaptor self-ligation, and use of DSN reaction conditions that provide high specificity of degradation.

Compared to conventional methods for ridding of undesired nucleic acids from RNA-seq libraries, PDD offers a cost-effective alternative where the range of degraded sequences is specified by a user-defined and controlled set of non-modified, inexpensive DNA oligonucleotides. The target range can therefore be easily adjusted by including or omitting the corresponding DNA probes in the DSN reaction. In contrast to the approaches utilizing capture of unwanted nucleic acids with oligonucleotide probes modified with affinity tags, PDD is less sensitive to initial degradation of the targeted RNA and provides better selectivity as only perfectly matched strands are eliminated. The longer perfect duplexes required for effective DSN action [about 10 bp, (Anisimova et al., 2008)] over techniques utilizing RNase H [requiring about 6 complementary paired nucleotides, (Wu et al., 1999)], as well as the higher temperatures at which the DSN-catalyzed degradation can take place (Anisimova et al., 2008) make PDD highly target-specific, as off-target degradation is undetectable (Archer et al., 2014). This minimum effective duplex length makes possible construction of targeting probes whose potential to bind and cause degradation of off-target sequences is minimized both at the design stage and by the reaction conditions. Further, the degradation reaction occurs only after adaptor ligation, rendering it incapable of generating artifactual fragment ends. Therefore, PDD is particularly suited for applications where high fidelity and lack of bias are more important than achieving extremely high depletion efficiency. To capitalize on its already low potential to introduce bias, we combined PDD here with the split-adaptor library generation method (Ingolia et al., 2009; Ingolia et al., 2012), which introduces much less sequence bias or 3'-5' bias than other library preparation methods (Levin et al., 2010; Jackson et al., 2014).

As an example, we apply PDD for construction of a representative RNA-seq library from total RNA of *Saccharomyces cerevisiae* cells fragmented by chemical or enzymatic hydrolysis. The modal length of the fragments obtained in our example is in the range of 100 to 200 nt and the minimal starting amount of RNA required for reliable generation of one RNA-seq library is about 100 ng, although we recommend starting with 500 ng. Few parameters need to be adapted additionally if a user would wish to scale up the experiment or fragment RNA to a different extent. The PDD implementation presented here is organized into eight protocols in accordance with the anticipated time course of the experiment. Support Protocol 1 provides a guideline for design of oligonucleotide degradation probes; Support Protocol 2 describes starting material and initial fragmentation of RNA; Support Protocol 3 includes setting up RNA end-repair and polyadenylation reactions; Support Protocol 4 contains procedures for reverse transcription of RNA from split-adaptor primer; Support Protocol 5 is used to remove unextended split-adaptor to prevent its self-ligation; Support Protocol 6 (and its alternative) contain directions for intramolecular circularization of cDNA; Basic Protocol 1 implements PDD with previously designed DNA oligonucleotide probes and circular cDNA; and Basic Protocol 2 describes amplification and size-selection of the rectified cDNA.

DESIGN OF LOW-BIAS DEGRADATION DEOXYRIBONUCLEOTIDE PROBES

PDD requires the preparation of nonmodified DNA probes corresponding to the sense-strand of the RNA to be targeted. The spacing of the probes depends on the minimum planned insert size of the library. The furthest allowable distance apart that will still theoretically promote the degradation of all target molecules is the minimum insert size minus 20 nt; however, somewhat closer inter-probe spacing is preferable to allow some flexibility to alter and refine individual probe sites. Probes should be more than 22

SUPPORT PROTOCOL 1

Transcriptional Profiling

11.15.3

nucleotides long and with a predicted melting temperature $T_m > 48^\circ\text{C}$ (by the data of Allawi and SantaLucia, 1997). We have not detected co-depletion of off-targets even with considerable sequence similarity to the probes; however, we still recommend that probes should lack sequence homology with other transcripts as much as possible. We therefore prescreen the target transcript for short regions of homology with other transcripts, after which probes can be designed outside these regions of cross-homology. If it cannot be avoided to target a region with appreciable homology to another transcript, and the T_m of the match is above the hybridization temperature (typically 48°C), then the predicted off-target transcript segment can be flagged or masked in the downstream bioinformatics analysis.

Users with bioinformatics experience should be able to generate scripts to automate much of this process; however, the following procedure avoids the need for writing scripts or using proprietary software. It uses an alignment tool familiar to most biologists (online BLAST) and can produce robust probe libraries within a reasonable time. It requires some use of the console commands, and working installations of Biopython and ApE software.

Materials

- Computer with internet connection (command line instructions given are for the Bash Unix shell)
- ApE (A Plasmid Editor, we use v2.0.47)
- A working installation of Biopython (www.biopython.org)
- Transcript sequences for your organism (refseq_rna database at NCBI and the target RNA sequence, from an established genome project)

Define number of PDD probes

1. Calculate optimal probe density.

Probes are designed to be spaced fairly regularly along the target RNA in the sense orientation. The maximum inter-probe distance is the minimum library insert size minus 20 nt. The period (distance from the start of one probe to the start of the next) is this distance plus the minimum probe length. Thus, for a library with a 200 nt lower size cut-off, the maximum inter-probe distance is 180 nt, thus the maximum period for probes of at least 25 nt is $180 + 25 = 205$ nt. However, we recommend aiming for a period of ~ 15 nt less ($= \sim 190$ nt in this case) to ensure robustness and allow some flexibility in target sites. This would allow coverage of nearly 10 kb of target sequence using 50 probes (which is the maximum number we have used for PDD).

Identify regions of cross-homology in the target transcript

2. Here we describe the use of BLAST on the NCBI BLAST Web site (www.ncbi.nlm.nih.gov/blast) to find short stretches of high homology to other transcripts in the target RNA. BLAST hits generated using high gap/mismatch penalties are a reasonable and quick measure of potential cross-hybridization, although a few hits containing mismatches may be included (these are split into two or more separate hits in step 3). This procedure assumes that a refseq_rna database exists for the target species.
 - a. In “nucleotide blast” on the blastn tab, enter the target sequence, choose refseq_rna as the “database”, and choose the species name under “organism”.
 - b. For the program settings, choose “Optimize for” → “somewhat similar sequences (blastn)”.
 - c. Under “algorithm parameters”, set “Max target sequences” to 10000; “Expect threshold” to 100000; “word size” to 7; “match/mismatch scores” to 1/−4; and “Gap costs” to “existence: 5 extension: 2”.

- d. Click on the BLAST button, wait for the search to complete, and download the results as an “XML,” which may contain thousands of hits.

Highlight cross-homology in ApE

3. ApE is highly suitable for manual probe design as it outputs the T_m of a sequence selection in real time. By displaying BLAST hits within ApE, probes can more easily be manually designed to target low cross-homology regions in the first iteration of probe selection. Our custom Python script, `blast2ape.py`, can be used to assist in this.
 - a. In ApE, paste the target RNA sequence (sense strand) into an empty ApE nucleotide file and then save it.
 - b. In the same directory, download the BLAST Hit Table text file, and add the `xml2ape.py` script (available at: <https://github.com/stu2/xml2ape>) to the directory.
 - c. Follow the instructions in the `xml2ape` readme file to execute the script on the command line using the other two files as input. It should produce an ApE file.
4. Select probes in ApE as follows:
 - a. Open the file in ApE and set the width of the ApE viewing window to half the maximum probe period calculated above to allow easy visual determination of where probes should fall on every second line.
 - b. Select suitable probe regions using the instant T_m calculation ($T_m > 48^\circ\text{C}$) while avoiding long contiguous overlaps with stretches of cross-homology where possible. However, not exceeding the maximum inter-probe distance is a higher priority than completely avoiding cross homology. Store each probe selection as a new feature, using the default name and feature properties for convenience. Save the ApE file when finished. The probes will be hybridized at the antisense cDNA stage, thus they must correspond to the sense strand of the target.
 - c. All features apart from BLAST hits can be extracted from the ApE file and exported into a table for ordering probes, using the `ape2table.py` script, which accompanies the `xml2ape` script, as outlined in the readme file.

IMPORTANT NOTE: *Order lyophilized probes and ensure that they are resuspended at a high enough individual concentration such that a stock mixture of all probes can be generated containing at least 2 μM of each individual probe.*

STARTING MATERIALS AND CONTROLLED FRAGMENTATION OF TOTAL RNA

This protocol provides directions for fragmentation of total RNA by divalent-cation-catalyzed hydrolysis, as implemented in the NEB Next RNA fragmentation module, which is a low-bias fragmentation method. Alternatively, RNase I fragmentation, which is also capable of cutting in a sequence-independent manner, can be employed (as shown in Figure 11.15.1A; see Alternate Protocol 2) and leaves behind fragment ends with similar chemistry (Meador et al., 1990). Other methods may also be employed provided that the 3' ends produced are either direct substrates for polyadenylation by EPAP (i.e., 3'-OH, in which case end-repair could be skipped) or for 3' end-repair by T4 PNK. This protocol is adapted to fragment 500 ng of total RNA to a modal size of ~ 180 nt in a 5 min hydrolysis reaction (this can be adjusted for different fragment sizes by referring to the NEBNext Magnesium RNA Fragmentation Module protocol). 500 ng of 180 nt fragments is equivalent to approximately 8 pmol of starting RNA fragments. According to our estimates this amount of RNA shall be sufficient for construction of four cDNA libraries and we recommend it as minimal starting point to prevent unnecessary reiterations of experiments in case of any misfortunes with subsequent manipulations. However, if the starting material is limiting, this and subsequent protocols can be scaled

SUPPORT PROTOCOL 2

Transcriptional Profiling

11.15.5

down to the amount needed for just one library (125 ng), without putting aside half the sample as a back-up at the start of Support Protocols 4 and 6. If significantly different fragments lengths are desired then the initial RNA mass must also be adjusted and all subsequent protocols scaled to keep the molarity of the RNA fragments roughly consistent.

Materials

- NEBNext Magnesium RNA Fragmentation Module (NEB, cat. no. E6150)
- 10× fragmentation buffer (NEBNext Magnesium RNA Fragmentation Module; NEB, cat. no. E6150)
- Starting RNA (500 ng of total cellular RNA extracted from cells using standard procedures, such as cell lysis followed by phenol deproteinization, or kits such as TRIzol Plus RNA Purification System by Promega, RNeasy Mini Kit by QIAGEN, etc.)
- Deionized nuclease-free water
- Iced water
- 10× NEBNext RNA Fragmentation stop solution (NEBNext Magnesium RNA Fragmentation Module (NEB, cat. no. E6150)
- 3 M NaOAc, pH 5.0
- Glycogen (molecular biology grade), 10 mg/ml
- 100% (v/v) ethanol
- HE buffer (see recipe)
- 80% ethanol, 20% (v/v) nuclease-free deionized water
- 200- μ l thin-walled PCR tubes
- Thermal cycler with heated lid
- Vortex mixer
- 20°C freezer
- Centrifuge for 200- μ l and 1.5-ml microcentrifuge tubes

1. Calculate the required fragmentation time appropriate for the desired product size range by referring to the NEBNext Magnesium RNA Fragmentation Module protocol (E6150).
2. Make up the hydrolysis reaction (below) in a thin-walled 200- μ l PCR tube:

Reagent	Concentration	Volume	
	stock		
NEBNext Fragmentation buffer	10×	1 × (final)	Add 2 μ l
Starting RNA (~500 ng)	X	X (final)	Add X μ l
Deionized nuclease-free water			to 20 μ l

3. Heat in the thermal cycler with heated lid at 94°C for the fragmentation time calculated above (see step 1). Have stop solution ready.
4. Transfer the tube to ice and add 2 μ l of 10× NEBNext RNA Fragmentation Stop Solution; mix well by briefly vortexing for 5 sec. Collect the resulting solution by centrifuging in a centrifuge briefly at 14,000 × g, 25°C.
5. Precipitate by adding 2.0 μ l of 3 M NaOAc pH 5.0 and 2 μ l glycogen, mixing, and adding 3 volumes of 100% ethanol. Vortex well twice, each time for 10 sec.
6. Transfer to -20°C for at least 2 hr.

It is convenient to store the tubes with the precipitated nucleic acids for long periods at this stage if required at or below -20°C.

7. Centrifuge for 20 min at 14,000 × g, 25°C. Carefully aspirate the supernatant and rinse the pellet with 80% ethanol, 20% (v/v) nuclease-free deionized water.

CAUTION: *The pellet may detach from the side of the tube.*

8. Centrifuge for 5 min at $14,000 \times g$, 25°C . Completely aspirate the supernatant.

This can be achieved by first taking away the major portion of the supernatant, brief spinning of the tube at $14,000 \times g$, and next carefully removing the residual liquid from the pellet with a fine pipet tip.

9. Dry the pellet for 10 min at room temperature or until no drops of liquid remain.

Dry the pellets with the lid open in a dust- and RNase-free environment, or covered with clean aluminum foil, avoiding intense airflow over the tube opening.

10. Resuspend the pellet in 10 μl HE buffer.

The pellet may require vigorous pipetting to redissolve, with re-centrifugation.

CONTROLLED FRAGMENTATION OF TOTAL RNA WITH *ESCHERICHIA COLI* RNASE I

ALTERNATE PROTOCOL 1

This alternate protocol may be used to fragment initial RNA in case divalent-cation fragmentation is not desired. Frozen, ground $0.5 \times$ total cell lysate is used in this protocol as starting material, making it compatible with approaches such as RNA protection assays and ribosomal profiling. This material is typically generated by weighing a cell pellet, resuspending in an equal amount of HBB buffer (thus $0.5 \times$), followed by freezing and grinding. Other methods resulting in approximately the same concentration of lysate are interchangeable. The protocol is adapted to obtain at least 500 ng of total fragmented RNA with most of the fragments occurring in the 20 to 300 nt range, resulting in at least 8 pmol of starting RNA fragments.

Materials

0.5 \times cell lysate [about 100 mg; prepared, e.g., by grinding frozen suspension of 0.5 \times wet cell pellet in 0.5 \times HBB (w/v) with a cryomill]

Iced water

HEPES breakage buffer (HBB; see recipe)

E. coli RNase I (Thermo Scientific, cat. no. EN0601)

1 M Tris·Cl, pH 2.0

3 M NaOAc, pH 5.0

Stop solution (see recipe)

Deionized nuclease-free water

5:1 Phenol:chloroform mixture, pH 4.5

Absolute ethanol

HE or TE buffer (see recipes)

1.5-ml microcentrifuge tubes

Vortex mixer

Microcentrifuge

NanoDrop ND-1000 spectrophotometer or similar

Fragment the RNA in a cytoplasmic lysate with *E. coli* RNase I

1. Weigh out 100 mg of the 0.5 \times cell lysate into a 1.5-ml polypropylene microcentrifuge tube and put it on ice.

The 0.5 \times cell lysate can be generated by weighing a cell pellet, resuspending in an equal amount of HBB buffer, followed by freezing and grinding.

2. Add 200 μl (2 volumes v/w) of ice-cold HBB to the lysate.

Include RNase and protease inhibitors, DDT or 2-mercaptoethanol in the additional HBB, if desired. Note that most of these additives will interfere with subsequent A₂₆₀ measurements, however.

3. Briefly vortex the tube several times until lysate is thawed. Then vortex again for 5 sec and keep on ice.

Avoid bubbling and generation of foam.

4. Centrifuge the tube with the lysate for 10 min at 12 000 × g, 0° to 4°C.
5. Transfer the supernatant into a new ice-cold 1.5-ml tube.
6. Repeat steps 4 and 5 once more.

Make sure not to contaminate the solution with fats and denatured proteins from the precipitated cell debris. Use a fine pipettor tip and release the plunger slowly.

7. Measure the optical density of the lysate at 260 nm against HBB.

Use measurement against appropriate buffer if RNase/protease inhibitors or DTT/2-mercaptoethanol were added to the HBB as they will affect 260 nm readings.

8. Add 3 U of *E. coli* RNase I per each optical unit of the lysate at 260 nm and mix the resulting solution carefully by pipetting.
9. Heat the tube with the mixture to 25°C and incubate it for 30 min at 25°C.

Extract proteins from the lysate and collect nucleic acids by ethanol precipitation

10. Add 50 μl of 1 M Tris·Cl, pH 2.0, 50 μl of 3 M NaOAc, pH 5.0, 50 μl of stop solution, and deionized nuclease-free water to 500 μl to the mixture in the tube from step 9.
11. Add 500 μl of 5:1 phenol:chloroform mixture, pH 4.5, and vortex intensively at high speed for 2 min.

Ensure the phases mix well forming white milky suspension.

12. Separate the phases by centrifugation for 10 min at 14,000 × g, room temperature.
13. Transfer the upper water phase to a new tube avoiding its contamination with phenol phase or denatured proteins on the interface.

Use a finer pipet tip when the major portion of the upper water phase was transferred. Do not attempt to transfer the entire upper phase as the interface molecules will contaminate it. Do not disturb the tube with the separated phases.

14. Repeat steps 11 to 13 once.
15. Measure the resulting volume of the clarified water phase, split it into two tubes if necessary, add 2.5 volumes of absolute ethanol and carefully mix the solution by turning upside down and vortexing briefly for 5 to 10 sec.
16. Precipitate and recover RNA as in Support Protocol 2, steps 6 to 10.
17. After dissolving the dried pellet in 10 μl of HE or TE buffer, acquire optical density readings against the corresponding buffer at 220, 260, 270, 280, and 320 nm with NanoDrop ND-1000.

Extraction of nucleic acids and deproteinization can be considered successful with 260 nm to 280 nm ratio of more than 2. Nonzero 320 and high 220 nm values may point towards contamination with other impurities. A sharp peak at 270 nm is characteristic of phenol contamination.

END-REPAIR AND POLYADENYLATION OF FRAGMENTED RNA

Use this protocol to remove the 2',3'-cyclic phosphate groups left by divalent cation-mediated hydrolysis (or RNase I treatment) at the 3' ends of RNA fragments, to allow addition of long poly(A) stretches by the template-independent polymerase activity of *E. coli* poly(A) polymerase I. The 2',3'-cyclic phosphate groups are converted into 3' hydroxyls by the 3'-phosphatase activity of bacteriophage T4 polynucleotide kinase (T4 PNK) (Wang et al., 2002). It is therefore important that the T4 PNK reaction is complete to avoid introduction of the bias by preferential polyadenylation of naturally-occurring RNA 3' ends carrying hydroxyl groups. If dealing with very short RNA fragments, attention also should be paid to generate 3'-end poly(A) sequences of at least 100 nt because these single-stranded poly(A) ends are later used to help separate extended split-adaptor reverse transcription primers from unextended primers (described in Support Protocol 5).

Materials

Solution containing 500 ng of nucleic acids in 10 μ l (see Support Protocol 2)
Deionized nuclease-free water
10 \times T4 PNK Buffer (New England Biolabs, cat. no. B0201S) **WITHOUT** ATP
T4 Polynucleotide Kinase (T4 PNK; New England Biolabs, cat. no. M0201), **NOT**
the T4 PNK 3' phosphatase negative mutant
RNaseOUT RNase inhibitor (Life Technologies)
10 mM EDTA
5 \times E-PAP Buffer (Ambion/Life Technologies, cat. no. AM1350)
25 mM MnCl₂ (Ambion/Life Technologies, cat. no. AM1350)
50 mM rATP with neutral pH
50 mM DTT (freshly made from 1 M stock)
2 U/ μ l E-PAP (Ambion/Life Technologies, cat. no. AM1350)
3 M NaOAc, pH 5.0
Absolute ethanol
HE buffer (see recipe)

1.5-ml microcentrifuge tubes
Thermostat for 1.5-ml microcentrifuge tubes
Vortex mixer

Repair the RNA 3' ends with T4 PNK and heat inactivate the reaction

1. Obtain solution containing 500 ng of nucleic acids in a 1.5-ml microcentrifuge tube.

From the following steps and on until step 1 of Support Protocol 6 it is recommended to adhere to the regular precautions of RNA handling to avoid sample degradation. The samples should also be treated as "pre-PCR" and not prepared in an area where post-amplification libraries are handled, until step 8 of Basic Protocol 2.

2. Use the following template to calculate the amount of water required for the reaction (total volume 15.25 μ l).

10 μ l of 50 ng/ μ l fragmented RNA (32 ng/ μ l final)
5.25 μ l deionized nuclease-free water

3. Add 5.25 μ l of deionized nuclease-free water to the RNA solution and mix.
4. Heat the tube in the thermostat for 2 min at 70°C and immediately transfer it into ice.

This step is used to heat denature RNA to make 3' ends as accessible as possible.

5. Add 10× T4 PNK buffer, RNaseOUT and T4 PNK in the amounts indicated below and mix it with the solution (total volume: 20 μl).

15.25 μl denatured RNA (from step 4)
2 μl of 10× T4 PNK buffer (NEB, without ATP; 1× final)
0.75 μl of 40 U/μl RNaseOUT (1.5 U/μl final)
2 μl of 10 U/μl T4 PNK (wild-type; 1 U/μl final)

Use the wild-type, 3' phosphatase-positive version of PNK. Many suppliers market 3' phosphatase-negative PNK that will not repair 3' ends. ATP should be absent from the reaction mixture to prevent phosphorylation of RNA 5' ends that will otherwise interfere with the subsequent ligation reaction. It is recommended to assay the 3' phosphatase activity of the T4 PNK using test DNA or RNA molecules in advance of working with the actual libraries.

6. Bring the contents of the tube to the bottom by brief pulsing at 14,000 × g, 25°C.

Avoid bubbling and generation of foam.

7. Incubate for 2 hr at 37°C.

8. Add 2 μl of 10 mM EDTA.

9. Heat the tube with reaction mixture in the thermostat for 10 min at 65°C.

Steps 8 and 9 are required to heat-inactivate T4 PNK. It will otherwise 5' phosphorylate the RNA fragments in the subsequent polyadenylation reaction.

Polyadenylate 3' ends of the RNA fragments and clean up the resulting RNA

10. Use the following template to calculate amounts of the reagents required for the polyadenylation reaction (total volume 50 μl):

22 μl 3'-end repaired RNA
9 μl deionized nuclease-free water
10 μl of 5× E-PAP buffer (1× final)
5 μl of 25 mM MnCl₂ (2.5 mM final)
1 μl of 50 mM rATP (1 mM final)
1 μl of 50 mM DTT (freshly made from 1 M; 1 mM final)
1 μl of 40 U/μl RNaseOUT (0.8 U/μl final)
1 μl of 2 U/μl E-PAP (0.04 U/μl final)

11. First mix the 3'-end repaired RNA solution with deionized nuclease-free water, E-PAP buffer, manganese chloride, rATP, and dithiothreitol.

The EDTA in the sample may chelate approximately 0.4 mM of the MnCl₂²⁺; this does not adversely affect the reaction.

12. Add RNaseOUT and E-PAP to the tube and thoroughly mix with the reaction solution by vortexing three times, each time for 5 sec. Bring the contents of the tube to the bottom by brief centrifugation for 5 sec at 14,000 × g, 25°C.

Avoid bubbling and generation of foam.

13. Incubate the resultant mixture in the thermostat for 1 hr at 37°C.

14. Centrifuge for 5 min at 14,000 × g, 25°C, to pellet any pyrophosphate precipitate formed. Transfer the supernatant containing the polyadenylated RNA to a fresh tube.

Upon completion of the reaction, a cloudy pyrophosphate precipitate may appear, forming a white pellet upon centrifugation. Absence of the pyrophosphate precipitate may be a

sign of an inactive preparation of the poly(A) polymerase or poorly repaired 3' ends of the RNA fragments. In this case, it is suggested to confirm the activity of these enzymes and carry out polyadenylation of a test RNA (e.g., yeast tRNA) before continuing with the library preparation.

15. Add 5 μ l of 3 M NaOAc, pH 5.0.
16. Add 150 μ l of absolute ethanol to each tube and mix thoroughly by vortexing twice, each time for 10 sec.
17. Precipitate the RNA following steps 6 to 9 of Support Protocol 2.
18. Resuspend the dried pellet in 20 μ l of HE buffer, pipetting vigorously to dissolve if necessary. Store long-term at -20°C or below if necessary.

Upon polyadenylation, optical density readings will no longer correspond to the original mass of the RNA in the sample.

SYNTHESIS OF COMPLEMENTARY DNA USING SPLIT-ADAPTER PRIMER

SUPPORT PROTOCOL 4

This protocol provides instructions for generation of directional first-strand cDNA from the polyadenylated RNA fragments using DNA oligonucleotide primer split into 5' and 3' parts by an artificial linker that blocks traversal of polymerases (commonly known as split-adapter primer) (Ingolia et al., 2009). The split-adapter primer contains sequences required for subsequent PCR amplification, including those compatible with Illumina primers and sequence tags unique for each library that are used for cluster formation and detection during amplification. It also contains a region complementary to poly(A) and the last two non-adenosine nucleotides of RNA fragments for priming of the cDNA first strand (Fig. 11.15.1B, discussed in Commentary). Although special measures described in the subsequent protocols and commentary are intended for elimination of split-adapter self-ligation products from the libraries, occurrence of amplifiable split-adapter primer dimers can still pose a major consideration while performing the method. It is therefore important to use minimal molar excess of the split-adapter primer providing efficient cDNA synthesis. Three- to five-fold molar excess of the split-adapter primer over the fragmented RNA can be sufficient for quantitative synthesis of cDNA according to our experience. Because of the pipetting time involved in certain steps, we do not recommend processing more than four samples in parallel.

NOTE: It is preferable to proceed with Support Protocol 5 directly after completing Support Protocol 4. Therefore, check the required materials for that section also before proceeding.

Materials

- Half (10 μ l) of the 20 μ l Polyadenylated RNA fragments solution (from step 18 of Support Protocol 3)
- 10 pmol/ μ l split-adapter primer 5' /5Phos/ GA TCG TCG GAC TGT AGA ACT CTG AAC G /iSp9/ G TGA CTG GAG TTC CTT GGC ACC CGA GAA TTC CAT TTT TTT TTT TTT TTT TTT TVN 3' in nuclease-free deionized water (iSp9: 9-carbon spacer; 5Phos: 5 phosphorylation);
- 10 mM dNTPs (Ambion/Life Technologies, cat. no. AM8200)
- Iced water
- 100 mM DTT (freshly made from 1 M stock)
- 5 \times First-Strand Buffer (Life Technologies, cat. no. 18080-044)
- 40 U/ μ l RNaseOUT RNase inhibitor (Life Technologies)

Transcriptional Profiling

11.15.11

200 U/μl SuperScript III reverse transcriptase (Life Technologies, cat. no. 18080-044)

Deionized water

200-μl thin-wall low-binding nuclease-free PCR microcentrifuge tubes

Sterile nuclease-free 1.5-ml microcentrifuge tubes

Thermal cycler with heated lid

NOTE: Primer sequence carries Illumina adapter sequences. Oligonucleotide sequences © 2007-2013 Illumina, Inc. All rights reserved. Derivative works created by Illumina customers are authorized for use with Illumina instruments and products only. All other uses are strictly prohibited.

Anneal the polyadenylated RNA with the split-adapter primer and synthesize first strand cDNA using SuperScript III reverse transcriptase

1. Put half (10 μl) of the polyadenylated RNA fragments solution from Support Protocol 3 into 200-μl thin-wall low-binding nuclease-free PCR tube. Retain the rest as a back-up.
2. Use the following template to mix the reagents required for the priming reaction (total volume 13 μl):

10 μl of 400 fmol/μl Poly(A) tailed RNA

2 μl of 10 pmol/μl split-adapter primer

1 μl of 10 mM each dNTPs

We have found that 1 U/μl Avian Myeloblastosis Virus (AMV) reverse transcriptase can also be used to specifically extend the split-adapter primer under similar conditions in a reaction mixture additionally containing 5 mM MgCl₂ with the use of incubation at temperature decreased to 45°C.

3. In a fresh 1.5-ml tube on ice, make up an appropriate amount of RT Master Mix II using the following template, mixing well by pipetting or vortexing for 5 sec:

Reagent	Concentration of stock	Volume to add (<i>n</i> = number of samples)
DTT	100 mM	<i>n</i> × 1.1 μl
First Strand Buffer	5×	<i>n</i> × 4.4 μl
RNase OUT	40 U/μl	<i>n</i> × 1.1 μl

The final volume of Master Mix II (6.6 μl per reaction) is 10% higher than nominally required per reaction to account for pipetting error.

4. Heat the first 200-μl tube (containing the 13 μl RNA/primer/dNTP mix) in a thermal cycler with heated lid for 3 min at 75°C using a heated lid option set to 105°C.

At this and subsequent steps until circular cDNA is generated (Support Protocol 6) avoid excessive heating and prolonged incubation of sample nucleic acids under harsh conditions, as well as their contact with enzymes possessing 5'-phosphatase activities, to preserve 5'-end phosphate groups of the split-adapter primer.

5. Cool the tube with reaction mixture to 65°C and then slow-ramp (1° to 4°C per minute) to 55°C.

6. Pause the thermal cycler program and add 6 μl of RT Master Mix II. Briefly mix the solution by pipetting.

7. *Do not leave the tube with open lid longer than is required to pipet.* Add 1 μl of 200 U/μl SuperScript III reverse transcriptase to the solution and briefly mix it by quick pipetting.

8. Close the lid tightly and continue the slow-ramp down to 50°C in the thermal cycler.
9. Incubate the tube with resultant mixture in the thermal cycler for 30 min at 50°C using heated lid.

Reagents for Support Protocol 5 can be prepared during the incubation.

PURIFICATION OF cDNA AWAY FROM EXCESS SPLIT-ADAPTER PRIMER

This protocol describes steps necessary for removal of the excess or unextended split-adapter primer from cDNA starting immediately after completion of the first strand cDNA synthesis (see Support Protocol 4). The removal of split-adapter primer is achieved by degradation of the single-stranded, unextended split-adapter molecules with *E. coli* Exonuclease I followed by isolation of the molecules possessing long polynucleotide chains (corresponding to the polyadenylated RNA fragments) on AMPure XP beads (illustrated in Fig. 11.15.1A; see also Archer et al., 2014). Exo I cleaves single-stranded DNA in a 3'→5' direction leaving intact dinucleotides while reaching 5' ends (Lehman and Nussbaum, 1964). Therefore, application of Exo I alone may lead to accumulation of the molecules that potentially can interfere with ligation and can generate invalid ligation products. Detachment of the unextended, but uncleaved, split-adapter primers from the RNA poly(A) stretches under ligation conditions can also replenish the pool of split-adapter self-ligation substrates. We found that substantial suppression of split-adapter dimer amplification in the subsequent protocols required a combination of Exo I treatment with purification on AMPure XP beads. This results in levels of adapter dimer amplification that can be effectively expunged during the library size-selection stage if mobility of the target fragments sufficiently differs from mobility of the adapter dimers (see Basic Protocol 2).

NOTE: Steps 6 to 11 (washing and dividing the AMPure XP beads) can be performed up to several hours in advance to simplify the later execution of sample processing.

Materials

First-strand cDNA solution (see Support Protocol 4, step 9)
20 U/μl *E. coli* Exo I (NEB, cat. no. M0293)
500 mM EDTA, pH 8.0
Deionized nuclease-free water
Agencourt AMPure XP purification system (Beckman Coulter, A63880)
PN binding buffer (see recipe)
PN washing buffer (see recipe)
70% ethanol, 30% (v/v) deionized nuclease-free water solution
HE buffer (see recipe)
RNase A/T1 Mix (Thermo Scientific, cat. no. EN0551)

Thermal cycler with heated lid
Vortex mixer
Low-binding nuclease-free 1.5-ml microcentrifuge tubes (e.g., Eppendorf DNA LoBind tubes)
Two magnetic racks for 1.5-ml microcentrifuge tubes
Heated oven to contain a magnetic rack set to 60°C
Heated block for 1.5-ml microcentrifuge tubes set to 60°C

NOTE: If dealing with many samples, a 37°C heat block for 200-μl tubes, or a second thermal cycler at 37°C, may be required at step 2 to facilitate batch operations.

Degrade the free unextended split-adapter primer with Exo I

1. Heat the first-strand cDNA solution in the thermal cycler with heated lid for 5 min at 60°C to separate unextended primer from poly(A) RNA.
2. Cool to 37°C, and within 1 min, add 1 µl of 20 U/µl Exo I (NEB) and mix it by brief, gentle vortexing four times, each time for 3 sec. Briefly spin the sample down if necessary.

If dealing with many samples, use a second heat block/thermal cycler at 37°C and transfer samples one-by-one. The last ones to be processed can be left up to 10 min at 60°C before transferring to the 37°C block and adding Exo I. Do not keep the tube at room temperature unnecessarily long at this and subsequent steps.

3. Incubate the reaction tube in the thermal cycler for 20 min at 37°C with heated lid.
4. Stop the reaction by mixing the contents of the tube with 1 µl of 500 mM EDTA, pH 8.0.
5. The final volume should be 22 µl. If it is less, make up to 22 µl with deionized nuclease-free water.

Purify the DNA molecules complementary bound to the polyadenylated RNA fragments on AMPure XP beads

6. Thoroughly resuspend AMPure XP beads and transfer 20 µl of the suspension per library sample into a new 1.5-ml microcentrifuge tube.

Use intensive pipetting to resuspend the beads prior to pipetting in this and subsequent steps.

7. Put the tube with the beads into magnetic rack and allow the beads to collect on the tube wall, and then completely remove the supernatant.

It normally takes at least 1 min for the beads to collect on the tube wall in water-like solutions.

8. Remove the tube from the magnetic rack and resuspend the beads in 44 µl of PN binding buffer per library sample by pipetting.
9. Repeat steps 7 and 8 once again.

It may take more than 2 min for the beads to completely collect on the tube wall because of the high viscosity of the PN binding buffer. It is important to fully adhere to the temperatures, volumes and recipes of the buffers provided in this protocol as binding efficiency of AMPure XP beads for differently sized nucleic acids strongly depends on these parameters.

10. Pipet 44-µl aliquots of the bead suspension (1 per library) into low-binding nuclease-free 1.5-ml tubes and cap tightly. Heat in an oven at 60°C for at least 5 min prior to use.
11. Put one of the magnetic racks into the oven heated to 60°C.
12. Transfer the tube from the thermal cycler containing the library (step 5) to the 60°C oven.
13. While keeping everything at 60°C, add the library sample (22 µl) to the pre-heated AMPure XP beads suspension (44 µl) from step 10 and mix well.

The ratio of PN binding buffer to sample is critical for the size-selectivity of the beads.

14. Incubate the resultant mixture for 5 min at 60°C.

15. Place the tube into the preheated magnetic rack (from step 11) and keep it in the oven for 2 to 5 min at 60°C allowing the beads to collect on the tube wall.
16. Transfer the tube from the magnetic rack in the oven into the rack at room temperature and immediately remove the supernatant.

Do not allow the tube to cool down much before the supernatant is removed, and proceed immediately to step 17. Perform steps 15 and 16 serially if processing multiple tubes.

17. Remove the tube from the rack. Add 30 μ l of PN washing buffer.

This and subsequent steps are performed at room temperature.

18. Resuspend the beads by pipetting.
19. Put the tube with the beads into the room-temperature magnetic rack and allow the beads to collect on the tube wall, and then remove the supernatant completely.
20. Repeat steps 17 to 19 twice more, using 200 μ l of 70% ethanol, 30% (v/v) deionized nuclease-free water solution instead of 30 μ l PN washing buffer.
21. Dry the beads pellet at room temperature by opening the tube lid for 4 min.

Do not dry for too long. The pellet should still appear glossy and moist.

22. Resuspend the bead pellet in 20 μ l of HE buffer by pipetting and wait at least 30 sec.
23. Allow the beads to collect on the tube wall in a magnetic rack at room temperature.
24. Transfer the supernatant with eluted nucleic acids into a fresh nuclease-free low-binding 1.5-ml microcentrifuge tube.

Remove RNA from the cDNA by degrading it with RNase A

25. Heat the tube with purified nucleic acids from step 24 for 10 min at 80°C and snap-cool on ice.

This is required to heat denature nucleic acids and render RNA single-stranded to be accessible for the RNases.

26. Add 1 μ l of RNase A/T1 Mix to the solution and mix it by gentle vortexing three times, each time for 3 sec.

NOTE: RNase A/T1 can be substituted by 1 mg/ml RNase A.

27. Incubate the resultant mixture for 20 min at 37°C.
28. Proceed immediately to Support Protocol 6 or freeze the sample for up to 1 year at or below -20°C .

CIRCULARIZATION OF cDNA USING INTRAMOLECULAR LIGATION

These instructions describe steps required to circularize the first-strand cDNA into amplifiable ssDNA circles with CircLigase II (marketed by Epicentre/Illumina) and purify the circularized ssDNA. CircLigase and CircLigase II (a pre-adenylated form of the ligase) are related to RNA ligase 1 from *Thermus scotoductus* bacteriophage, and are also homologous to classical bacteriophage T4 RNA ligase 1 enzyme (Blondal et al., 2005). Preadenylated CircLigase II is consumed during ligation, thus it must be present in at least equimolar concentration to its substrate. Publically available information about terminal nucleotide preferences of CircLigase is scarce (Tate et al., 2012). It follows from Epicentre descriptions that there is dependency of ligation yield on nucleotide nature for

SUPPORT PROTOCOL 6

Transcriptional Profiling

11.15.15

both 5' and 3' nucleotide residues albeit somewhat different from the traditional T4 RNA ligase 1 (Zhenodarova et al., 1989) (discussed in Commentary). Nonetheless, generation of a directed cDNA library using CircLigase circularization is less bias-prone compared to other methods (Jackson et al., 2014). CircLigase requires 5' phosphate groups on the donor and 3' hydroxyls on the acceptor strands and therefore nucleotides and oligonucleotides carried over from the RNA degradation steps of Support Protocol 5 will not affect the circularization reaction. However, these oligonucleotides and manganese ions used for circularization may interfere with subsequent DSN treatment (Shagin et al., 2002) (see Basic Protocol 1) and should be removed from the reaction products by purification on Ampure XP beads or similar methods.

Materials

RNA-free cDNA (~250 ng/4 pmol) (see Support Protocol 5, step 28)
Deionized nuclease-free water
10× CircLigase II Reaction Buffer (Epicentre/Illumina, cat. no. CL9021K)
50 mM MnCl₂
5 M betaine (Epicentre/Illumina, cat. no. CL9021K)
100 U/μl CircLigase II (Epicentre/Illumina, cat. no. CL9021K)
500 mM EDTA, pH 8.0
Agencourt AMPure XP purification system (Beckman Coulter, cat. no. A63880)
PN binding buffer (see recipe)
70% ethanol, 30% (v/v) deionized nuclease-free water solution
HE buffer (see recipe)

Low-binding nuclease-free thin-walled 200-μl PCR microcentrifuge tubes
Vortex mixer
Thermal cycler with heated lid
Low-binding nuclease-free 1.5-ml microcentrifuge tubes
Magnetic rack for 1.5-ml microcentrifuge tubes

Circularize single-stranded cDNA library molecules with CircLigase II

1. Take 10 μl of the 21 μl sample from step 28 of Support Protocol 5 (~2 pmol of cDNA) and freeze the rest as a back-up.
2. Transfer the sample into a low-binding thin-walled 200-μl PCR microcentrifuge tube.
3. Use the following template to calculate the required amounts of the reaction components (total volume 20 μl):

10 μl of 200 fmol/μl ss cDNA (100 fmol/μl final)
4.5 μl of deionized nuclease-free water
2 μl of 10× CircLigase II Buffer (1× final)
1 μl of 50 mM MnCl₂ (2.5 mM final)
2 μl of 5 M Betaine (0.5 M final)
0.5 μl of 100 U/μl CircLigase II (Epicentre/Illumina; 2.5 U/μl final).

4. First add deionized water, CircLigase II Buffer, manganese chloride and betaine to the tube with ss cDNA and carefully mix the resulting solution by pipetting or vortexing three times, each time for 3 sec.
5. Supplement the reaction with CircLigase II and mix the solution by gentle vortexing three times, for 3 sec each time.

6. Incubate the tube with circularization reaction in the thermal cycler for 5 hr at 60°C using lid heated to 105°C.

Most of the molecules are ligated in less than 1 hr. Longer times may be helpful for ligating molecules with less accessible ends, to reduce bias; however, the incubation time can be shortened to 1 to 2 hr without loss of gross ligation product yield.

7. Stop the reaction by mixing the solution with 1 μ l of 500 mM EDTA, pH 8.0, using pipetting or vortexing three times, each time for 3 sec.
8. Inactivate CircLigase by heating the reaction solution to 80°C for 10 min in the thermal cycler with heated lid.

Clean-up the circularized DNA using Ampure XP beads

9. Resuspend Ampure XP beads well and transfer a $n \times 20$ - μ l aliquot (where n is the number of samples) into a low-binding nuclease-free 1.5-ml tube.
10. Place beads on a magnetic rack and wait at least 1 min for them to collect on the tube wall.
11. Aspirate beads and resuspend in $n \times 42$ μ l of PN binding buffer (where n is the number of samples).
12. Repeat steps 10 and 11.
13. Resuspend the beads well and pipet 42- μ l aliquots (1 per sample) into low-binding nuclease-free 1.5-ml tubes.

Preparation of beads in steps 9 to 13 can be performed several hours ahead of time.

14. Add the 21 μ l circularized cDNA sample (from step 8) to the 42 μ l bead aliquot and mix by pipetting and vortexing four times, each time for 3 sec.
15. Place the tube in a magnetic rack and allow the beads to collect on the tube wall for 2 min.
16. Remove the supernatant and resuspend the beads in 200 μ l of 70% ethanol, 30% (v/v) deionized nuclease-free water solution.
17. Repeat steps 15 and 16.
18. Place the tube in a magnetic rack, allow the beads to collect on the tube wall for at least 2 min, and aspirate the supernatant.
19. Dry the bead pellet at room temperature by opening the tube lid for 4 min.

Do not dry for too long. The pellet should still appear glossy and moist.

20. Resuspend the bead pellet in 10 μ l HE buffer by pipetting and wait 30 sec.
21. Allow the beads to collect on the tube wall for 2 min in a magnetic rack at room temperature.
22. Transfer the supernatant with eluted nucleic acids into a fresh low-binding nuclease-free 1.5-ml microcentrifuge tube. Proceed to Basic Protocol 1 or store the obtained solution for 1 year at -20°C or below.

At this stage, an amplifiable circular ssDNA library is formed. It is recommended to perform a test PCR amplification of this library in advance of PDD treatment described in Basic Protocol 1. A total of 1 μ l of the solution from step 22 should be sufficient to generate PCR products after 10 to 20 cycles in amounts allowing analysis in nondenaturing gels stained with common reagents for double-stranded nucleic acids detection. The DNA generated during such PCR amplification can also be analyzed in Agilent 2100

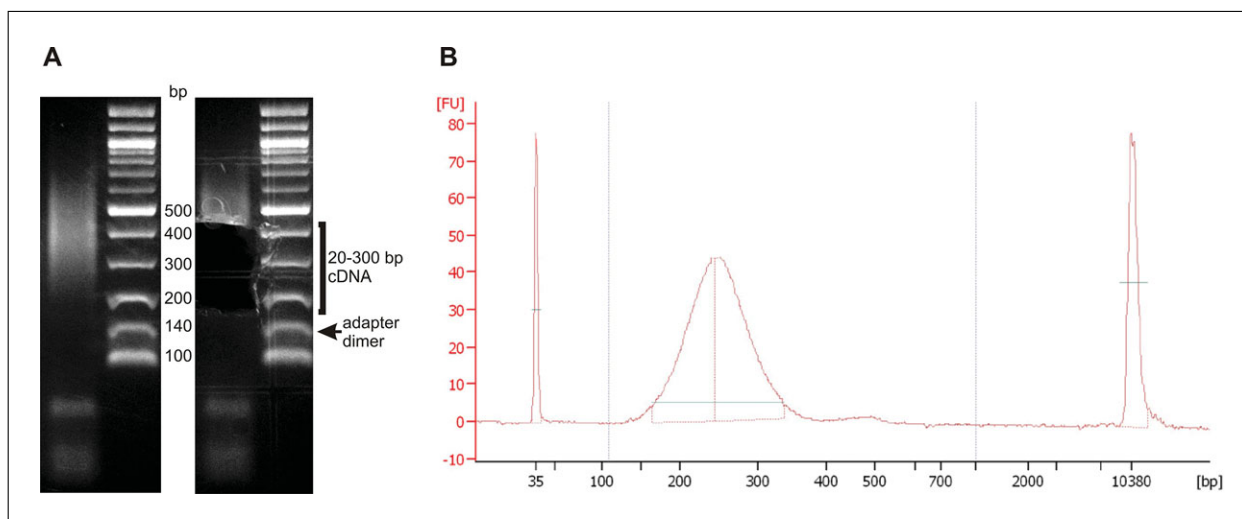


Figure 11.15.2 Size-selection of amplified libraries. **(A)** Agarose gel electrophoresis of an amplified PDD treated library. A non-denaturing ~4% agarose gel was prepared as described in Basic Protocol 2. (left) A PDD-treated library prepared as described in the unit was amplified by PCR for 14 cycles, the resultant DNA was pre-stained by addition of 1/5 (v/v) 6× GRGreen Loading Buffer (Excellgen) to the PCR mixture. The pre-stained amplified DNA solution was loaded onto the gel and the DNA was separated for 3 hr at room temperature under 8 V/cm. 100-bp DNA Ladder (NEB) supplemented with pre-amplified split-adaptor primer dimers was run in parallel as a separation reference. (right) After separation, the gel region containing DNA corresponding to 20 to 300 bp of cDNA insert size was cut out and processed further as described in Basic Protocol 2. **(B)** Confirmation of size-selection of the library obtained from (A) by microfluidic electrophoresis on Bioanalyzer 2100 instrument (Agilent technologies). Sample was run on a dsDNA chip (Agilent High Sensitivity DNA Kit). Shown is a representative electrophoretogram, with dsDNA concentration expressed as fluorescent units (FU). The size distribution of the inserts is inferred by subtracting the length of the flanking regions (~140 nt) from the size distribution seen in the library.

Bioanalyzer with High Sensitivity DNA Kit or other similar setup. In this case, amounts of the unextended split-adaptor primer amplification products (dimers) can be assessed, as well as such analysis can be further used to control efficiency of the subsequent PDD treatment (see Fig. 11.15.2 for an example). The PCR amplification can be performed generally as described in Basic Protocol 2, but Taq DNA polymerase can be substituted for Pfx DNA polymerase for economy reasons for these test amplifications.

ALTERNATE PROTOCOL 2

CIRCULARIZATION OF cDNA IN THE PRESENCE OF OLIGONUCLEOTIDE INHIBITING SPLIT-ADAPTER SELF LIGATION

The combination of the *Exo I* treatment and purification of cDNA molecules with AM-Pure XP beads described in Support Protocol 5 is sufficient to prevent split-adaptor self-ligation products from overpopulating the library during amplification. After these procedures, amplifiable split-adaptor dimers are still present in the library in amounts congruent with the desired cDNA containing molecules, but can be effectively eliminated during a size-selection stage. However, in case of unfavorable design of the split-adaptor primer that cannot be altered or if the cDNA-containing molecules do not differ in length enough from empty split-adaptor primer amplicons to allow size selection, self ligation and amplification of the split-adaptor primer can overwhelm other molecules in the final library. In this case, a different approach to ligation can be applied that is based on strict dependency of type 1 RNA ligases on single-stranded, 5'-phosphorylated substrates with accessible 3' hydroxyls (Gumport and Uhlenbeck, 1981; Romaniuk and Uhlenbeck, 1983; Walker et al., 1975). A 5' unphosphorylated oligonucleotide with blocked 3' end complementary to the split-adaptor primer can be used to anneal with the primer molecules and create a 5' overhang that masks 3' hydroxyls of the free, unextended split-adaptors and thus precludes them from ligation (explained in Fig. 11.15.3A). At the same time, lacking 5' phosphates and 3' hydroxyls, the blocking oligonucleotide is inert for CircLi-gase and its excess does not exhaust the reaction components. We have used a blocking

**Probe-Directed
Degradation
(PDD) for
Removal of
Unwanted cDNA
from RNA-Seq
Libraries**

11.15.18

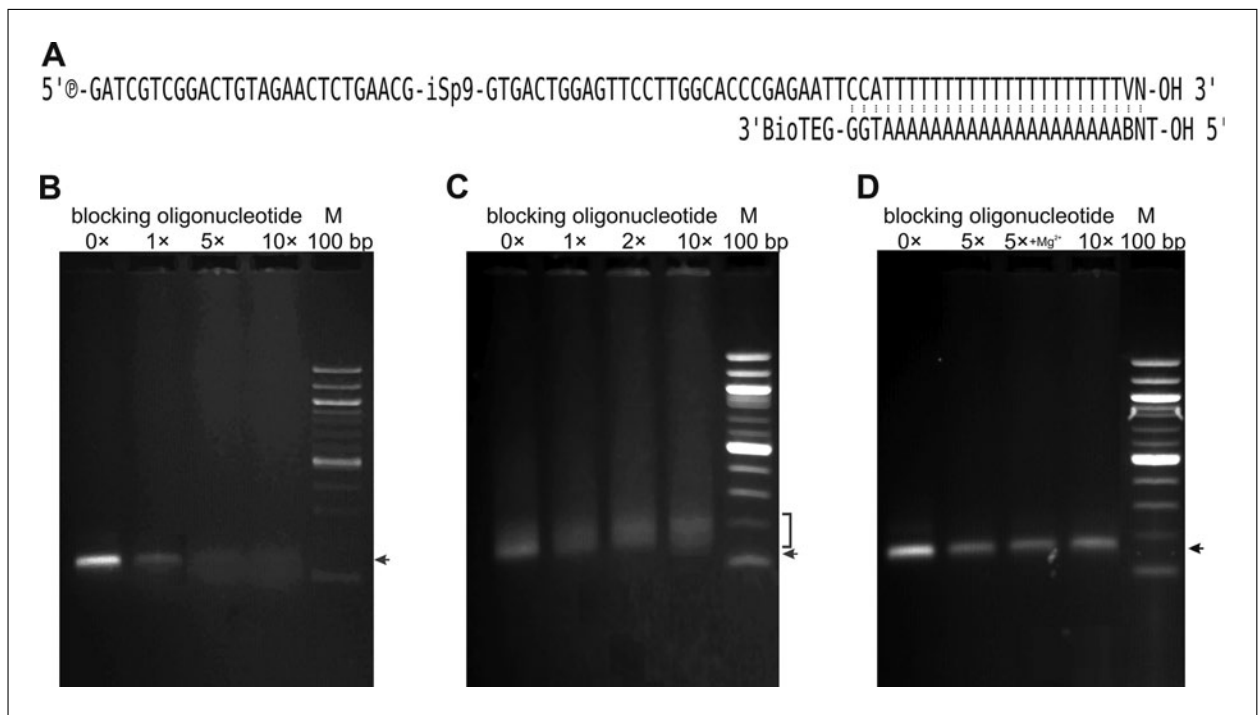


Figure 11.15.3 Suppression of the unextended split-adapter primer self ligation by the use of complementary blocking oligonucleotide. **(A)** Schematic representation of the complex inhibiting split-adapter self ligation, where complementary blocking 3' biotinylated oligonucleotide is annealed with the 3' portion of the split-adapter primer, creating a 5' overhang. **(B)** Inhibition of the split-adapter primer self ligation by annealing it with the increasing amounts of the complementary blocking oligonucleotide. 1 pmol of the split-adapter primer was annealed with 0 (control), 1, 5 and 10 pmol(s) (indicated on top of the corresponding lanes) of the blocking oligonucleotide in 10 μ l of 1 \times annealing buffer and further ligated as described in Alternate Protocol 2, using 3 hr incubation time. **(C)** Same as (B), but 0.5 pmols of the cDNA library constructed generally as described in Support Protocols 2 to 5, with 3 \times split-adapter primer excess and without SPRI beads purification stage, were used as a template for circularization. **(D)** Same as (B), but with ligation reaction temperature elevated to 50°C. +Mg²⁺ denotes additional 5 mM and 2.5 mM MgCl₂ in the annealing and circularization reactions. **(B)** through **(D)** 1/10 of the ligation reactions were used as templates for 15-cycle amplification with Taq DNA polymerase in 20 μ l PCR mixtures. Shown are results of electrophoretic separation were 5 μ l of the amplified material were pre-stained with 1 μ l of 6 \times GRGreen Loading Buffer (Excellgen) and then loaded onto nondenaturing 1.5% agarose gel.

oligonucleotide with the sequence TNBAAAAAAAAAAAAAAAAAAAAAAAAATGG/3BioTEG/ and found that as low as equimolar to the split-adapter primer amount almost completely prevents formation of amplifiable ligation products (Fig. 11.15.3B) and does not affect amplification of useful library molecules with cDNA inserts at even higher ratios (Fig. 11.15.3C). Use of the oligonucleotide with the suggested design requires lower ligation temperature (Fig. 11.15.3D); however, BNA- or PNA-based oligonucleotides can be considered to increase stability of the blocking duplexes. We did not test this approach in a complete PDD method because measures implemented in Support Protocol 5 were sufficient, but we suggest it as a possible alternative in case difficulties with split-adapter self-ligation persist. Due to their low thermal stability, the duplexes of blocking oligonucleotides with the library molecules should not become effective targets for DSN in the subsequent stages. Additionally, another purification step could be applied, or a complementary decoy oligonucleotide (such as CCATTTTTTTTTTTTTTTTTTTTTTVNA) could be added in excess to the library after circularization, to prevent adapter cleavage by DSN.

Additional Materials (also see Support Protocol 5)

Annealing buffer (see recipe)

10 pmol/ μ l blocking oligonucleotide 5' TNB AAA AAA AAA AAA AAA AAA
AAT GG /3BioTEG/ 3' (ordered from IDT and dissolved in nuclease-free
deionized water)

***Anneal the blocking oligonucleotide with the molecules containing split-adapter
primer sequences***

1. Acquire 10 μ l of the 21 μ l sample from Support Protocol 5, step 28 (~2 pmol cDNA) and freeze the rest.
2. Transfer the sample into low-binding thin-walled nuclease-free 200- μ l PCR microcentrifuge tube.
3. Add the following components of the annealing reaction (total volume 30 μ l) into the tube and mix by vortexing three times, each time for 3 sec:
 - 10 μ l of 200 fmol/ μ l ss cDNA and split-adapter (67 fmol/ μ l final)
 - 3 μ l of 10 \times annealing buffer (1 \times final)
 - 6 μ l of 10 pmol/ μ l Blocking oligonucleotide (2 pmol/ μ l final)
 - 11 μ l deionized nuclease-free water.
4. Place the tube into the thermal cycler and heat for 3 min at 95°C with the heated lid on.
5. Slowly cool the tube to 25°C (5°C per min).

Perform ligation with CircLigase II at lower temperature

6. Use the following template to calculate the amounts of the additional components required to complete a 50- μ l ligation reaction:
 - 5 μ l of 10 \times CircLigase II Buffer (1 \times final)
 - 2.5 μ l of 50 mM MnCl₂ (2.5 mM final)
 - 10 μ l of 5 M Betaine (1 M final)
 - 1.5 μ l deionized nuclease-free water
 - 1 μ l of 100 U/ μ l CircLigase II (2 U/ μ l final).
7. First add deionized water, CircLigase II Buffer, manganese chloride, and betaine to the tube with annealed blocking oligonucleotide from step 5 and carefully mix the resulting solution by pipetting or vortexing twice, each time for 3 sec.
8. Supplement the reaction with CircLigase II and mix the solution by gentle vortexing three times, for 3 sec each time. Bring the contents of the tube to the bottom by short centrifugation at 14,000 \times g, room temperature.
9. Incubate the tube with circularization reaction in the thermal cycler for 5 hr at 40°C using lid heated to 105°C.
10. Stop the reaction by mixing the solution with 2.5 μ l of 500 mM EDTA, pH 8.0, by pipetting.
11. Inactivate CircLigase by heating the tube with the reaction mixture to 80°C for 10 min.
12. Complete the protocol by following steps 9 to 22 of Support Protocol 6, but adjust the initial amount of Ampure XP beads such that 20 μ l of original beads suspension is resuspended in 105 μ l PN binding buffer (not 42 μ l) prior to addition of the circularization reaction. The dilution of PN binding buffer with sample in the binding reaction is always 2:1.

PROBE-DIRECTED DEGRADATION OF cDNA

This protocol provides instructions on how to anneal the circularized DNA library molecules with the PDD degradation oligonucleotide probes, cleave the duplexes on targeted DNA circles with Duplex-Specific Nuclease (DSN) enzyme and clean up the sample to prepare it for amplification. The DSN DNase (purified from Kamchatka crab and marketed by Evrogen) exhibits high specificity for cleavage of DNA strands complementarily paired within at least 10 bp DNA duplexes or 15 bp DNA:RNA hybrids (Anisimova et al., 2008). The key advantages of using this enzyme are high temperature optimum for activity relative to the thermal stability of its dsDNA targets, and relatively long perfect duplexes required for cleavage (also discussed in Commentary) (Anisimova et al., 2008). The high reaction temperature guards against nonspecific duplex formation on unintended targets during PDD (Itaya and Kondo, 1991). DSN requires long perfectly matched duplexes, enabling design of probes targeting a very specific subset of sequences (see Support Protocol 1 and Commentary).

We have not tested PDD with library generation methods other than the split-adapter approach described above. In theory, however, there should be no reason that the single-stranded, linker-ligated cDNA stage from other library preparation protocols cannot be substituted here, provided that the molar concentration is similar to what is given here for the split-adapter library. However, some protocols may yield constructs with linkers having some double-stranded DNA character even if the insert itself is single-stranded, which would be inadvertently degraded during PDD. In addition, note that the DNA strand of DNA:RNA hybrids is similarly a target for DSN.

We have tested a mixture of 50 probes, but if significantly more than this is needed, there may be a requirement for re-optimization of the amount added to the PDD reaction. It is possible that the hybridization temperature can be changed to suit the desired stringency threshold of the library. Here we use 48°C, as this is normally above the calculated T_m of most off-targets in the probe set. DSN can be effectively inactivated by phenol extraction. Dissolved phenol precipitates upon mixing with PN binding buffer in the Ampure XP bead purification protocol; we therefore use gel filtration to subsequently purify the library.

Materials

- Solution containing circular ss cDNA (10 μ l) (see Support Protocol 6, step 22)
- Stock mixture of all DSN probes at 2 μ M each (ordered from IDT in a 96-well plate format and dissolved at high individual concentrations in nuclease-free deionized water, subsequently mixed together equally)
- 10 \times DSN buffer (Evrogen)
- Deionized nuclease-free water
- 1 U/ μ l DSN (Evrogen)
- Iced water
- 500 mM EDTA, pH 8.0
- HE buffer (see recipe)
- 25:24:1 Phenol:chloroform:isoamyl alcohol mixture, pH 8.0

- Low-binding nuclease-free thin-walled 200- μ l PCR microcentrifuge tubes
- Vortex
- Centrifuge
- Low-binding nuclease-free 1.5-ml microcentrifuge tubes
- Thermal cycler with heated lid
- Micro Bio-Spin P-6 columns (Bio-Rad, cat. no. 732-6221)

Anneal the PDD oligonucleotide probes with the circular ssDNA library molecules

1. Estimate the amount of the cDNA sample in terms of equivalent molecules of RNA used to generate it. We approximate the amount (in pmol) as $3.11 \times \text{mass (ng)/length (nt)}$ where mass (ng) is the original mass of RNA that is represented by the sample.

In this example it is one quarter (due to the halving of the sample at the start of Support Protocols 4 and 6) of the original input of 500 ng, i.e., 125 ng. This approximates 2 pmol if the median fragment length was ~180 nt.

Processing more than four samples at a time is not recommended.

2. Transfer the entire 10 μl sample containing circular ssDNA (see Support Protocol 6, step 22) into a new low-binding thin-walled nuclease-free 200- μl PCR tube.
3. Supplement the solution with sufficient DSN probes mixture stock such that each probe is equivalent in concentration to the amount of RNA represented in the library (calculated as above, in this example, 1 μl of 2 μM stock probe mix to give 2 pmol of each probe).

Use amount of the DSN probes providing equimolar or slightly higher ratio of each probe to the calculated (expected) molar amount of the initial RNA fragments obtained in Support Protocol 2. Insufficient depletion probes will lead to incomplete elimination of the targets, however, excessive abundance of the DSN probes may interfere with the reaction. We have tested a 50-probe mixture under these conditions.

4. Add 1.5 μl of 10 \times DSN buffer and make up to 15 μl with nuclease-free water.

This is the smallest scale of reaction size possible unless mineral oil is used to prevent evaporation.

5. Carefully mix the resulting solution by vortexing the tube three times, each time for 3 sec, and spin the contents of the tube down completely by centrifuging it for 5 sec at 14,000 $\times g$, 25°C.
6. Make up the following DSN mixture on ice in a new nuclease-free 1.5-ml microcentrifuge tube, where n is the number of samples:

$n \times 0.55 \mu\text{l}$ of 10 \times DSN buffer
 $n \times 0.55 \mu\text{l}$ of 1 U/ μl DSN
 $n \times 4.4 \mu\text{l}$ deionized nuclease-free water.

7. Transfer the DSN mixture into 5- μl aliquots (1 per library) in fresh 200- μl thin-walled nuclease-free PCR tubes.

Pre-aliquoting avoids the need to pipet hot liquids accurately in step 12.

8. Place the tube containing the library in a thermal cycler and heat it for 3 min at 95°C with heated lid set at 105°C. Leave the DSN mixture on ice.
9. Set the incubation temperature to 75°C, and then slowly ramp the temperature (~3°C per min) to the hybridization temperature, 48°C, in the thermal cycler.

Perform the probe-directed cleavage of the annealed duplexes with duplex-specific nuclease

10. When the thermal cycler reaches the hybridization temperature, incubate for another 5 min.
11. Increase the block temperature by 5°C (e.g., from 48°C to 53°C) and open the lid. Warm the DSN mix aliquots for ~30 sec on the block of the thermal cycler.

The block temperature is temporarily increased to counteract the cooling of liquids or tubes if they are temporarily removed for pipetting.

12. Rapidly pipet a 5 μ l aliquot of DSN mix to each library sample and mix well by pipetting.

Do not keep tube lids open longer than necessary.

13. Set the temperature back to the desired annealing temperature of 48°C and incubate the mixture for a further 20 min in the thermal cycler.

In case a different set of probes is used, the incubation temperature of the DSN reaction may be optimized according to the recommendations provided in Support Protocol 1.

Stop the reaction and clean-up the DNA by gel filtration

14. While keeping the tube heated at 48°C in the thermal cycler, quickly add 1 μ l of 500 mM EDTA, pH 8.0, to the reaction mixture.

Do not keep tube lids open longer than necessary.

15. Mix the contents of the tube by pipetting while keeping the tube in the thermal cycler at 48°C.

16. Remove the tube from the thermal cycler and supplement the solution with 54 μ l HE buffer.

17. Add 75 μ l of 25:24:1 phenol:chloroform:isoamyl alcohol mixture, pH 8.0, to the solution and vortex at high speed for 2 min.

Make sure the phases mix well by forming milky suspension.

18. Centrifuge the tube for 15 min at 14,000 \times g, room temperature.

19. Transfer 50 μ l of the upper aqueous phase into a fresh low-binding nuclease-free 1.5-ml microcentrifuge tube.

20. Acquire a new Micro Bio-Spin P-6 column or equivalent.

21. Equilibrate the column with HE buffer according to manufacturer's instructions as follows:

- a. Centrifuge the column for 1 min at 1000 \times g, room temperature, and discard the flowthrough solution.
- b. Add 500 μ l HE buffer to the column and centrifuge again for 1 min at 1000 \times g, room temperature.
- c. Repeat the washes three times more.

22. Slowly apply the entire sample from step 19 in the center of the Micro Bio-Spin P-6 column gel.

23. Recover the gel-filtered sample by centrifuging the Micro Bio-Spin P-6 column for 3 min at 1000 \times g, room temperature (yields about 50 μ l). Use the obtained material immediately to perform Basic Protocol 2 or store for 1 year at or below -20°C .

AMPLIFICATION AND SIZE-SELECTION OF THE PDD-TREATED cDNA

Use the procedures described in this protocol to amplify double-stranded DNA molecules from single-stranded circular DNA precursors treated with PDD, add sequence tags compatible with Illumina's TruSeq Small RNA Sample Prep kit sequencing to these molecules during amplification and select the molecules of the amplified library according to the desired cDNA insert length by their electrophoretic mobility. The size-selected libraries each possessing a unique sequencing tag then can be finally purified, analyzed by on-chip electrophoresis and mixed with each other in a desired molar ratio. The structures of the forward primers set with the different Illumina tags (the XXXXXX hexanucleotide

**BASIC
PROTOCOL 2**

**Transcriptional
Profiling**

11.15.23

region) and reverse primer in relation to the split-adapter primer of the library DNA circles are explained in Figure 11.15.1C. Each library can be primed with one of the forward primers containing the unique hexanucleotide index to be discriminated from other libraries in a single multiplexed sequencing run. Note that the primers overhang from the initial circular DNA molecules of the library to allow addition of the portions required for amplification during Illumina sequencing runs as well as incorporation of the hexanucleotide tags. Therefore, in the course of the first two PCR cycles when templates without these extensions dominate, a lower annealing temperature is used. The desired annealing temperature for the runs can be estimated using standard tools (such as those available on the IDT Web site) for calculation of DNA duplexes thermal stability in regard to the reaction buffer used. Generally, a lower number of cycles for amplification is preferred; however, the amount of the generated library material must be sufficient to reliably visualize the amplicons in a size-selection gel run. For the described protocols, we suggest using amplification for about 10 to 20 cycles, and a pilot test amplification of a small amount of the PDD-treated DNA is recommended to define the optimal amplification depth. Usage of high-fidelity DNA polymerases with proofreading activity is recommended together with addition of single-stranded DNA binding protein to the PCR mixture to ensure specific priming of the starting single-stranded DNA molecules. We have found polyacrylamide or agarose gels both equally suitable for separation of the amplified fragments within our range of interest, however, the described separation technique may need to be adjusted to a different insert size.

Materials

- 12 μ l of DSN-treated circular ssDNA solution (see Basic Protocol 1, step 23)
- Deionized nuclease-free water
- 10 mM dNTPs solution (each), PCR grade
- 10 \times *Pfx* Buffer (Invitrogen/Life Technologies, cat. no. 11708-013)
- 50 mM MgSO₄ (Invitrogen/Life Technologies, cat. no. 11708-013)
- 10 μ M forward primer: 5' CAA GCA GAA GAC GGC ATA CGA GAT XXX
XXX GTG ACT GGA GTT CCT TGG CAC CCG AGA ATT CCA 3' (where
XXX XXX is an Illumina indexing hexanucleotide sequence), ordered from IDT
and dissolved in nuclease-free deionized water.
- 10 μ M reverse primer: 5' AAT GAT ACG GCG ACC ACC GAG ATC TAC ACG
TTC AGA GTT CTA CAG TCC GA 3' (ordered from IDT and dissolved in
nuclease-free deionized water)
- 0.5 μ g/ μ l Extreme Thermostable Single-Stranded DNA Binding Protein (ET SSB
protein; NEB, cat. no. M2401S)
- 2.5 U/ μ l Platinum *Pfx* DNA Polymerase (Invitrogen/Life Technologies, cat. no.
11708-013)
- UltraPure Agarose (Invitrogen, cat. no. 16500-500)
- 0.5 \times TBE buffer (see recipe)
- 6 \times GRGreen Loading Buffer (Excellgen, cat. no.-EG-1008)
- 0.5 μ g/ μ l 100 bp DNA Ladder (NEB, cat. no. N3231)
- HE buffer (see recipe)
- 3 M NaOAc, pH 8.0
- 5 μ g/ μ l nuclease-free glycogen
- Stop solution (see recipe)
- 25:24:1 Phenol:chloroform:isoamyl alcohol mixture, pH 8.0
- Absolute ethanol
- 80% ethanol, 20% (v/v) deionized water solution
- High Sensitivity DNA Kit (Agilent Technologies, cat. no. 5067-4626)
- Low-binding thin-walled nuclease-free 200- μ l PCR microcentrifuge tubes
- Vortex mixer

Thermal cycler with heated lid
 500-ml Schott bottles
 Microwave oven
 Low-binding nuclease-free 1.5-ml microcentrifuge tubes
 Apparatus for horizontal electrophoresis allowing at least 100 × 7 (L×H) mm gels
 Blue light transilluminator with optical filter
 Sterile syringe with needle
 Quantum Prep Freeze 'N Squeeze DNA Gel Extraction Spin Columns (Bio-Rad, cat. no. 732-6165)
 Thermoshaker for 1.5-ml microcentrifuge tubes
 2100 Bioanalyzer (Agilent Technologies)

NOTE: Primer sequences are adapted from Illumina primer sequences. All rights reserved. Derivative works created by Illumina customers are authorized for use with Illumina instruments and products only. All other uses are strictly prohibited.

Amplify dsDNA library with unique Illumina sequencing tags from the circular PDD-treated ssDNA

1. Acquire 12 µl of the circular ssDNA solution from Basic Protocol 1 (step 23) and transfer the solution into a new low-binding thin-walled nuclease-free 200-µl PCR tube.
2. Use the following template to calculate amounts of the reagents required for the amplification reaction (total volume 30 µl):

12 µl Circular ssDNA template
 10.7 µl deionized nuclease-free water
 0.9 µl of 10 mM dNTPs (each; 0.3 mM final)
 3 µl of 10× *Pfx* buffer (1× final)
 0.75 µl of 50 mM MgSO₄ (1.25 mM final)
 0.9 µl of 10 µM forward primer (0.3 µM final)
 0.9 µl of 10 µM reverse primer (0.3 µM final)
 0.25 µl of 0.5 µg/µl ET SSB protein (4.2 ng/µl final)
 0.6 µl of 2.5 U/µl Platinum *Pfx* DNA Polymerase (0.05 U/µl final).

3. First add fresh deionized water, dNTPs, *Pfx* Buffer, MgSO₄, forward and reverse primers to the circular ssDNA solution in the PCR tube and mix by vortexing three times, each time for 3 sec.
4. Add ET SSB protein and *Pfx* DNA Polymerase to the contents of the tube and thoroughly mix the solution by pipetting and gentle vortexing three times, each time for 3 sec.
5. Bring the contents of the tube to the bottom completely by brief centrifugation at 14,000 × *g*, room temperature, and place the tube into the thermal cycler.
6. Use the following settings to program the thermal cycler:

Step	Temperature	Duration	Cycles
Initial melting	94°C	5 min	1
Melting	94°C	30 sec	2
Annealing	62°C	30 sec	
Polymerization	68°C	45 sec	
Melting	94°C	30 sec	12
Annealing	76°C	30 sec	
Polymerization	68°C	45 sec	
Final polymerization	68°C	1 min	1

7. Run the above program on the thermal cycler using lid heated to 105°C.
8. Upon completion of the amplification, store the tube with the amplified DNA for 96 hr on ice and proceed to the size-selection steps. For longer storage (up to 1 year), keep the solution with the amplified DNA at –20°C or lower.

Perform size-selection for the 20- to 300-bp insert size in an electrophoresis run using nondenaturing agarose gel

9. Suspend 4 g of UltraPure Agarose in 100 ml of 0.5× TBE buffer in appropriate glassware (a 500-ml Schott bottle with lid left loose will suffice).
10. Heat the bottle in the microwave oven until the liquid starts boiling and then mix the contents of the bottle by intensive orbiting for about 1 min.
11. Leave the bottle for 10 min at room temperature.
12. Repeat steps 10 and 11 for 3 to 4 times more.
13. Transfer the top solution with completely melted agarose into a new Schott bottle (or other suitable glassware) leaving the unmelted portions of the agarose on the bottom intact and let the solution cool to a safe temperature (65° to 75°C, refer to the electrophoresis apparatus manufacturer's instructions).
14. Pour the TBE agarose solution in the electrophoresis cuvette equipped with the comb forming wells able to accommodate 35 µl of solution (although several wells can be used).

Refrain from pouring gels thicker than 7 mm.

15. Cool the poured solution for 20 min at room temperature and for 30 min more at 4°C.
16. Submerge the cast gel in the electrophoresis unit filled with 0.5× TBE according to the electrophoresis chamber manufacturer's instructions.
17. Mix the solution with the amplified DNA from step 8 with 5 µl of 6× GRGreen Loading Buffer by pipetting and vortexing three times, each time for 3 sec and load it into the well(s).

Any sensitive system to visualize dsDNA in a gel without damaging the DNA can substitute for GRGreen Loading Buffer. However, we found it convenient to use this particular stain as it can be loaded in a premix with the sample and does not require staining the entire gel.

18. In a new low-binding nuclease-free 1.5-ml microcentrifuge tube mix 0.3 µl of 0.5 µg/µl 100-bp DNA Ladder with 27 µl deionized water, 3 µl of 10× Pfx Buffer, and 5 µl of 6× GRGreen Loading Buffer by pipetting and vortexing (three times, each time for 3 sec) and load the resulting solution into the adjacent well.
19. Start the electrophoresis by applying 3 V/cm for 15 min; continue with 5 V/cm for another 30 min and use 8 V/cm for the rest of the run. Control movement of the 100 bp band by periodic visualizations of the DNA in the gel under blue light using appropriate optical filters.
20. Upon the end of the electrophoresis, transfer the gel onto a clean (DNA and DNase-free) transparent plastic support and visualize the gel.

Figure 11.15.2A illustrates an example of a successful run where the library molecules of the desired size range are well separated from the split-adapter dimers.

21. Using a clean sharp tool (such as disposable scalpel or a syringe with needle) cut out the gel region containing the required fragments (shown in Fig. 11.15.2A).

Recover the amplified DNA from the gel and clean up the sample

22. Collect the cut gel pieces into a new Quantum Prep Freeze 'N Squeeze DNA Gel Extraction Spin Column upper filter insert and crush it against the insert wall with the syringe needle.
23. Incubate the spin column for 30 min at -20°C .
24. Centrifuge the spin column for 10 min at $12,000 \times g$, room temperature.
25. Measure the volume of the collected liquid and transfer it into a new low-binding nuclease-free 1.5-ml microcentrifuge tube.
26. Add equal volume of HE buffer and $0.3 \times$ volumes of 3 M NaOAc, pH 8.0, into the Spin Column upper filter insert with the gel remnants and incubate it in thermoshaker at 1000 to 1500 rpm for 30 min, room temperature.
27. Repeat steps 24 and 25; merge the collected liquid with that from step 25 and mix the resulting solution by vortexing.
28. Distribute the united solution into new low binding nuclease free 1.5-ml microcentrifuge tubes by equal aliquots of no more than $300 \mu\text{l}$.
29. Add $2 \mu\text{l}$ of $5 \mu\text{g}/\mu\text{l}$ glycogen, $0.1 \times$ volumes ($30 \mu\text{l}$ or less) of stop solution to each tube, and mix by vortexing three times, each time for 3 sec.
30. Add equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol mixture, pH 8.0, to the solution in each tube and vortex at high speed for 2 min.

Make sure the phases mix well forming milky suspension.

31. Centrifuge the tube(s) for 15 min at $14,000 \times g$, room temperature.
32. Transfer the upper aqueous phase into fresh low-binding nuclease-free 1.5-ml microcentrifuge tubes containing 2.5 volumes ($825 \mu\text{l}$ or less) of absolute ethanol and mix the resultant solution.

Avoid transferring phenol phase or molecules from the interface area.

33. Incubate the tube(s) for at least 2 hr at -20°C .
34. Collect the precipitated material by centrifuging for 15 min at $14,000 \times g$, 0° to 4°C .
35. Decant the supernatant and wash the pellet in 1 ml of 80% ethanol, 20% (v/v) water solution by vortexing the tube four times, each time for 3 sec.
36. Centrifuge for 5 min at $14,000 \times g$, room temperature, and completely remove the supernatant.

This can be achieved by first taking away the major portion of the supernatant, brief spinning of the tube at $14,000 \times g$, and next removing the residual liquid with a fine pipette tip. Make sure the pellet resides in the tube.

37. Keep the tube with lid open for 15 min at 40° to 50°C to dry the pellet.

Put a foil cover or similar protection on the tubes to prevent intense airflow over the tubes lifting the dried pellets.

38. Dissolve the dried pellet(s) in $15 \mu\text{l}$ HE buffer. Store for 1 year at -20°C or below.

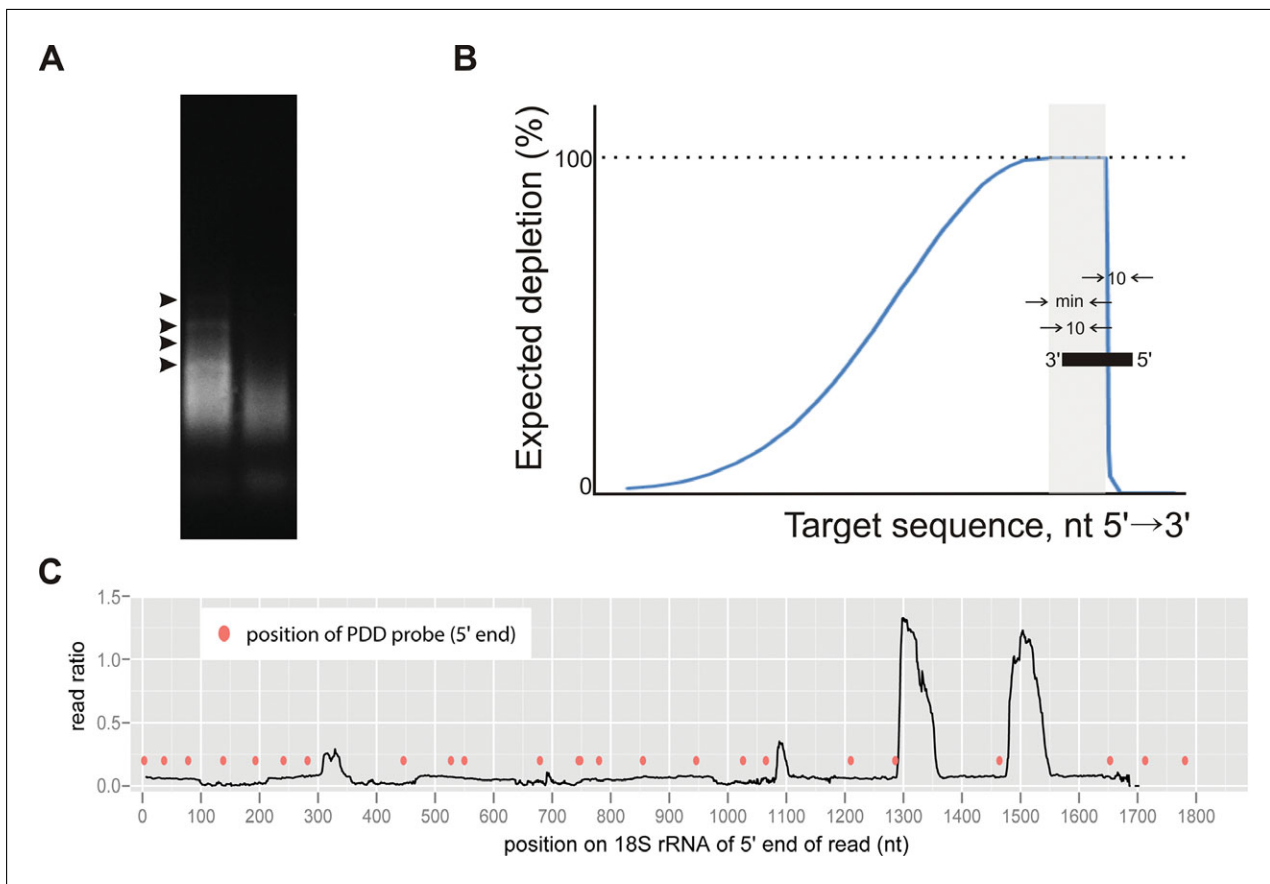


Figure 11.15.4 Typical results from PDD treatment. **(A)** (left lane) An amplified library derived from RNase-treated yeast lysate exhibits distinct bands (arrows) corresponding to RNase-resistant rRNA fragments. (right lane) These are absent in the library aliquot treated with PDD (using a probe-set designed to target rRNA in yeast) prior to amplification. **(B)** Schematic representation of a typical profile of PDD depletion efficiency shown for an isolated degradation probe (filled rectangle). Expected depletion in % is plotted against nucleotide positions of target sequence displayed from 5' to 3' end, position of a complementary degradation probe is indicated from 3' to 5'. Arrows indicate distances in nucleotides interrelated by minimal 10 nt overlaps required for efficient DSN cleavage ("10"), minimal cDNA fragment insert size in the library ("min") and a region where a complete target removal can be expected (highlighted background). **(C)** A randomly fragmented RNA library, half of which was treated with PDD, was sequenced on an Illumina HiSeq 2500 and the density of reads mapping to the 18 S rRNA was calculated from the 5' ends. Shown is a density profile of inserts having length >119 nt, as a ratio of PDD-treated to untreated, after normalizing to the number of mRNA-mapped reads in each library. Positions of 18 S-targeting probes are shown: each probe corresponds to an upstream region of depletion in the PDD-treated sample (adapted from Archer et al., 2014).

It may take several minutes of intensive vortexing to dissolve the pellet(s) completely as they will contain polymers from the agarose gel. These polymers will not affect the subsequent analytical electrophoresis as well as Illumina sequencing. However, note that for the same reason UV absorbance readings at this step will not be accurate and cannot be used to calculate mass of the amplified dsDNA. Much lower contamination with UV absorbing impurities will occur in case of polyacrylamide gel use at the probable cost of lower DNA output.

Assess the quality of the library using analytical electrophoresis

39. Analyze an aliquot of the sample from step 38 in an Expert 2100 Bioanalyzer or similar instrument using a High Sensitivity DNA Kit and Assay (or similar) according to the manufacturer's instructions. Figure 11.15.2B illustrates results of an analysis of successfully constructed and purified library.

It may be convenient to compare the results with those from a library not treated with PDD. Extinction of individual peaks, characteristic of RNase-fragmented rRNA, from the PDD-treated libraries can indicate success of the DSN reaction (Fig. 11.15.4A).

40. Calculate the molarity of the library using the results of the Expert 2100 run and combine it with other libraries containing unique hexanucleotide sequencing tags in the desired proportions if multiplexed library sequencing is intended. The sample now can be used for Illumina HiSeq (or similar) sequencing input run performed with the TruSeq Small RNA Sample Prep Kit.

REAGENTS AND SOLUTIONS

Annealing buffer

Component	Concentration stock	Volume for		
		Final 1×	Final 10×	1 ml 10×
HEPES-KOH, pH 7.6 (25°C)	1.81 M	25 mM	250 mM	138.12 μl
NaCl	5 M	100 mM	1000 mM	200 μl
MgCl ₂	1 M	5 mM	50 mM	50 μl
Deionized nuclease-free water				611.88 μl

Store for 6 months at 0° to 4°C or for 2 years at –20°C

HE buffer

Component	Concentration stock	Volume for		
		Final 1×	Final 10×	10 ml 1×
HEPES-KOH pH 7.6 (25°C)	1.81 M	10 mM	100 mM	55.2 μl
EDTA pH 8.0	500 mM	0.1 mM	1 mM	2 μl
Deionized nuclease-free water				to 10 ml

Store for 6 months at 0° to 4°C or for 2 years at –20°C

HEPES breakage buffer (HBB)

Component	Concentration Stock	Volume for		
		final 1×	final 10×	500 ml 1×
HEPES-KOH, pH 7.6 (25°C)	1.81 M	20 mM	250 mM	5.52 ml
KCl	solid	100 mM	1000 mM	3.73 g
MgCl ₂	1 M	2 mM	20 mM	1 ml
Deionized nuclease-free water				to 500 ml

Store for 6 months at 0° to 4°C or for 2 years at –20°C

PN binding buffer

Component	Concentration stock	Volume for	
		final 1×	10 ml 1×
PEG 8000	40% (w/v)	20% (w/v)	5 ml
NaCl	5 M	2.5 M	5 ml

Store for 1 year at room temperature

PN washing buffer

Component	Concentration stock	Volume for	
		final 1×	6 ml 1×
PN binding buffer	1×	0.67×	4 ml
HE buffer	1×	0.33×	2 ml

Store for 1 year at room temperature

Stop solution

Component	Concentration stock	Volume for	
		final 1×	40 ml 1×
SDS	20% (w/v)	10% (w/v)	20 ml
EDTA pH 8.0	500 mM	100 mM	8 ml
Deionized nuclease-free water			12 ml

Store for 1 year at room temperature

TBE buffer, 0.5×

Dissolve one TBE 10× POWDER Ready-Pack (Amresco, cat. no. 0478-2PK) in 2 liters of fresh deionized nuclease-free water and filter through 0.2- to 0.45- μ m nitrocellulose membrane (or similar) to obtain 5× TBE solution. Use 5× TBE solution and deionized nuclease-free water to prepare required amounts of 0.5× TBE buffer. Store 5× or 0.5× TBE buffer solutions in tightly closed vessels for 1 year at room temperature.

TE buffer

Component	Concentration stock	Volume for		
		final 1×	final 10×	10 ml 1×
Tris-OAc pH 7.4 (25°C)	1 M	10 mM	100 mM	100 μ l
EDTA pH 8.0	500 mM	0.1 mM	1 mM	2 μ l
Deionized nuclease-free water				to 10 ml

Store for 6 months at 0° to 4°C or for 2 years at –20°C

COMMENTARY

Background Information

The motivations for depletion of particular sequences from cDNA libraries are diverse and dependent on the desired purpose for which the library is constructed. For many applications, it is the highly abundant structural RNAs such as rRNA, tRNA and spliceosomal RNA that massively outnumber mRNA and small regulatory RNA and prevent their intensive study. rRNA in RNA samples derived from eukaryotic cells has traditionally been an encumbrance for cloning of mRNA molecules into cDNA and their further analysis. This problem is not limited to ribosomal RNA, as libraries generated from other tissues may be overwhelmed with other dominating RNA sequences e.g., globin mRNA in red blood cell RNA. To overcome this difficulty in classical approaches, where full-length mRNAs were generally desired, several measures have been devised and are still widely used in methods for transcript analysis. These measures mostly rely on the presence of either 3' poly(A) or 5' cap structures as specific features of most (but not all) cytoplasmic mRNA molecules

in eukaryotes (Harbers, 2008). 3' poly(A) tracts are the most commonly used mRNA markers, allowing their specific purification on affinity resin with immobilized oligo(dT) nucleotide chains. Total RNA can also be treated with 5'-phosphate-dependent 5'→3' exonuclease effectively cleaving non-capped ribosomal RNA (Jager et al., 2009). These techniques can be applied independently or in combination with each other (Edery et al., 1995; Harbers, 2008; Maruyama and Sugano, 1994; Scotto-Lavino et al., 2006).

Since high-throughput sequencing has become available and its power has made it a widespread analytical tool, the classical methods for removal of undesired RNA contaminants can still be used if full-length canonical mRNAs are desired and obtainable. However, with its relatively short reads produced in large numbers, RNA-seq is often conceived as an instrument for screening of entire classes of different RNAs, including naturally uncapped mRNAs or metabolic intermediates of canonical messages, small regulatory RNAs or rare transcripts, or for

characterizing experimentally generated fragments of RNA, e.g., for footprinting studies, nucleotide-resolution detection of RNA modifications by cleavage, and other applications. In this case, cap and 3' poly(A) no longer can serve as selective markers, and the whole focus of the separation technique is shifted towards introduction of the least possible bias into cDNA libraries to enable reliable quantitative measurements. Removal of undesired, but abundant sequences in such scenarios presents a particular challenge when RNA is to be analyzed that is significantly degraded, either by chance, such as in low-quality or aged samples or on purpose, such as in RNA protection assays. Left unremoved, fragments of ribosomal RNA can occupy more than 90% of the sequencing capacity, rendering such experiments unfeasible. Consequently, numerous techniques addressing this issue have emerged, all of them using complementarity to the ribosomal RNA itself as the only possible selective feature. Ribosomal RNA, in most cases, is hybridized to species-specific complementary oligonucleotides (often LNAs), and is then captured using some affinity tag attached to the probes. (Pang et al., 2004) This can deeply deplete RNA-seq libraries of rRNA sequences, however, the high efficiency of these approaches may affect their selectivity. Whilst removal of ribosomal RNA by its binding to affinity-tagged, modified oligonucleotides works to remove uncleaved or somewhat degraded rRNA, highly fragmented RNA is more challenging. High efficiency of rRNA removal can be achieved even for fragmented samples by the use of large numbers of overlapping antisense oligonucleotides, however, this approach creates an even greater risk of off-targets elimination. One established way around this is the introduction of a secondary stage of selection where rRNA:antisense oligonucleotide hybrids are enriched for perfectly matched duplexes by binding to specifically constructed immobilized antibody, thus preventing some of the off-targets from extinction (O'Neil et al., 2013). In any case, it is clear that with its dependence solely on thermal stability of the rRNA:antisense oligonucleotide duplexes, customizing the design of antisense probes for this method would require extensive testing to find an optimal balance between selectivity and efficiency for RNA from each species of interest, without a guarantee of undetectable off-targets loss. The highly modified affinity-tagged LNA probes typically utilized are

prohibitively expensive to manufacture *en masse* as customized sequences, precluding easy customization (e.g., to target non-rRNA sequences or rRNA sequences from different species).

A somewhat more affordable and flexible approach utilizing cleavage of rRNA:antisense oligonucleotide duplexes with RNase H has recently been demonstrated to be of use if ready-made modified probes are not available for the species of interest or the particular sequence to be targeted (Morlan et al., 2012; Rosenow et al., 2001). It was shown that rRNA removal with this method was comparable to methods with oligonucleotide immobilization while requiring less optimization of the setup. However, 100% coverage of the target RNAs by oligonucleotide probes is required (Morlan et al., 2012). Further, the relatively short length of perfect duplexes targeted by RNase H, combined with the high thermal stability of RNA:DNA duplexes means that isolated mismatches can still be potential targets for cleavage. Because the RNase H cleavage in this method is performed prior to ligation of 5' and 3' adaptors, a true representation of input fragment extremities is not guaranteed in the final library [although gross biases in terms of the number of fragments mapping to transcripts were not observed with this method (Morlan et al., 2012)]. More exotic methods include RNA separation with denaturing high-performance liquid chromatography (Castro et al., 2013) or use of LNAzymes to break the complementary matched chain (Dolinsek et al., 2013); however, unless more comparative studies are available it is impossible to assume higher specificity and efficiency of these methods over traditional immobilized oligonucleotide capture.

These considerations prompted us to look for an accessible method providing easy customization and high selectivity of abundant RNA removal from deeply fragmented samples while keeping the starting material unaffected as much as possible. In classical approaches when removal of abundant molecules is required, nonlinear dependence of the re-annealing reaction speed on the concentration of the hybridizing strands is used, resulting in the preferential re-association of highly abundant duplexes when reaction conditions are tuned appropriately [the so-called C₀T hybridization, (Ko, 1990)]. The preferentially re-annealed duplexes with abundant sequences are then either specifically removed from the library leaving rare single-stranded molecules,

such as in cDNA normalization methods (Vandernoot et al., 2012), or their subsequent amplification is inhibited, such as in comparative studies using suppression subtractive hybridization (Diatchenko et al., 1996). It was recently demonstrated that a combination of C₀T hybridization with Duplex-Specific Nuclease (DSN) from Kamchatka crab can be a superior approach for the full-length cDNA normalization (Peng et al., 2008; Bogdanova et al., 2009; Bogdanova et al., 2011; Christodoulou et al., 2011; Gijavanekar et al., 2012). Moreover, this enzyme has been successfully used to deplete RNA-seq libraries of ribosomal RNA after C₀T hybridization (Yi et al., 2011; Miller et al., 2013). DNA oligonucleotide probes have been used to cleave unwanted sequences with DSN in a cDNA library preparation method, while targeting of overrepresented rRNA sequences in RNA-seq libraries with DNA oligonucleotides designed specifically against them in a probe-directed degradation (PDD) approach have also been proved effective (Li et al., 2013; Wickersheim and Blumenstiel, 2013).

We therefore wished to design a PDD depletion method that would employ DSN activity at the cDNA stage, using probes designed to target rRNA sequences, and to embed this technique in the polyadenylation/split-adapter ligation approach for RNA-seq library construction, which has been demonstrated to produce less bias than other approaches. The key advantages of DSN are its high activity at temperatures up to 70°C and selectivity for cleavage of relatively long, perfectly matched DNA strands in duplexes with DNA (from 10 bp) and RNA (from 15 bp) (Anisimova et al., 2008). Owing to the high thermal stability of DSN, a set of degradation probes can be designed to anneal with targeted RNA at temperatures not allowing their stable interactions with the rest of the transcriptome. The possibility of binding with isolated mismatches resulting in degradation is further diminished by the inability of DSN to cleave strands containing a mismatch within its 10-bp duplex requirement. Minimal oligonucleotide probes for DSN cleavage can therefore completely avoid any possible cross-targeting of the entire sequence heterogeneity of smaller bacterial genomes and provide at least two orders of advantage in sequence selectivity for larger eukaryotic genomes in comparison with RNase H depletion. We have demonstrated that targets cleavage with DSN generally follows an easily predictable probability pattern within a selected size-range of library inserts. If the desired size-range of RNA fragments is

known, a comprehensive probe design can be performed where abundant rRNA fragments are first identified and degradation probes are designed specifically against them with spacing between the probes determined by the library insert range. Therefore the need of a complete or excessive rRNA sequence coverage by the degradation probes is avoided, granting the freedom to add, remove, or shift the probes to achieve better efficiency and selectivity. We implemented the PDD depletion using polyadenylation/split-adapter ligation method because currently it is considered as the approach producing least insert bias (Jackson et al., 2014). However, because the PDD depletion works at the cDNA stage, the method can be used with virtually any other cloning technique, as well as it can be applied as an auxiliary supplement for existing libraries if additional or deeper removal of unwanted sequences is necessitated (for example, if some proportion of the LNA probes in a depletion kit designed for one species do not bind the desired target in the species of interest). PDD might also be applied to remove adapter-dimers in other library preparation techniques by targeting the junction sequence. Here however, the polyadenylation strategy necessitates an oligo-dT stretch and redundant 3' terminal nucleotides in the split-adapter primer, rendering the self-ligated primer junction difficult to specifically target. PDD does not introduce any detectable bias into the library and can be used in quantity-sensitive applications where high (up to nucleotide) resolution of the initial RNA fragments is expected. Use of a limited number of unmodified DNA probes, design of which can be bioinformatically streamlined, allows PDD to be easily tailored to fit any source species and fragmentation level of RNA with a desired depth and selectivity of depletion.

Critical Parameters and Troubleshooting

The library preparation method described here contains multiple successive stages, including important enzymatic reactions and sensitive purification procedures, which may potentially fail and render all of the subsequent manipulations ineffective. As a general precaution and especially in case if such library construction is performed for the first time, we recommend verifying performance of each enzyme using simple assays with test molecules often supplied with these reagents by the manufacturers before working with a

real library. Other general considerations include careful handling of library while it is in the RNA form (i.e., starting from Support Protocol 3 and until Support Protocol 6), because additional fragmentation at these stages may introduce undesirable bias into the distribution of the initial RNA fragments or simply impair cloning efficiency. More specific problems that may potentially occur are discussed below.

Probe design is very important for the success of PDD. A procedure to streamline the design of efficient probes is described in Support Protocol 1 and can also be derived from the original research paper (Archer et al., 2014). Note that the discrimination temperature at which the DSN reaction is performed will need to be re-adjusted in this case. Even completely following the instructions presented here it may happen that individual probes will occur not absolutely efficient. This perhaps is due to the specific structures developed by the unwanted RNA fragments in solution which are hard to reliably predict. Therefore, if the probes are designed with minimal redundancy to achieve high selectivity, it may be expected that after initial testing re-design or addition of several probes may be required.

Because long 3' poly(A) homopolynucleotides complementary attached to the cDNA are later used to separate library molecules from empty split-adapter primers, it is important to make sure that the RNA is effectively polyadenylated. Although a pyrophosphate precipitate formed in the polyadenylation reaction can be an indirect indicator of its success, it is recommended to perform a test polyadenylation reaction with homogeneous RNA taken in molar amount similar to the expected library molarity to simulate the reaction prior construction of the library. Upon polyadenylation, the test reaction products can be concentrated and assayed using denaturing gel electrophoresis.

Split-adapter primer described in this unit and the original article can generally be used without changes in the PDD method applied to any source RNA. However if there are some particular reasons for a new design (such as changes in sequencing platform necessitating other sequences for annealing of amplification primers etc.), terminal nucleotides preferred for the ligase used for circularization must be considered. Different type I RNA ligases can have varying sequence preferences. There is not much information about CircLigase II available; however, expecting its similarity to the CircLigase enzyme, purines, and

especially G is a preferred residue for the 5' terminus of the split-adapter primer. Note that there are also 3' end nucleotide preferences of the ligation reaction that introduce fragment insert bias natural for this cloning approach.

It is very important to strictly adhere to the described procedures while performing purification of the cDNA:poly(A)+ RNA hybrids on SPRI resin. The process includes several steps at elevated temperatures at which RNA integrity must be maintained, requiring RNase-free handling of the sample material and appropriate composition of the buffers. Temperature, ionic strength, concentration of polyethylene glycol and length of nucleic acids all will affect binding of the cDNA:poly(A)+ RNA hybrids and free split-adapter primer to SPRI resin and the separation efficiency. It should also be noted that different batches of SPRI beads preparations (even from the same supplier) can have slightly different affinity properties and to perform library protocols with confidence it is recommended to test try SPRI purification of material from the pilot polyadenylation reaction described above (which should be retained on the resin) and the split-adapter primer (which should not bind).

We generally did not find any difficulties performing the circularization reaction itself and according to our experience the reaction is usually complete in less than an hour. However, in advance of a first library preparation efficiency of circularization can be confirmed by including only the split-adapter primer into the reaction mixture. This will generate the split-adapter self-ligation products that can be then amplified with the primers for library PCR and further used as a mobility marker to identify the region of the empty amplifiable split-adapters during library size selection. It is important not to contaminate pre-amplification libraries with this amplified self-ligation product (or any other amplified library), utilizing the typical precautionary measures such as dedicated pre-PCR and post-PCR areas.

It is critical to completely remove (degrade) the RNA, either enzymatically (as described here) or chemically, from the circularized DNA library molecules before the assembly of the DSN reaction because complementary RNA will cause cleavage of the cDNA strands as DSN recognizes hybrid DNA:RNA duplexes. Note also that bivalent cations other than magnesium must be removed from the sample prior to the addition of the DSN since its fidelity might be adversely affected. For a customized probe set, the annealing and

incubation temperature of the DSN reaction should be re-calculated accordingly, following considerations outlined in Support Protocol 1 and (Archer et al., 2014). It is usually not necessary and potentially undesired to prolong the incubation of the DSN cleavage stage as it can result in off-targets degradation to accumulate. However the incubation time or amount of the DSN added may need to be adjusted as we found some variation in the DSN activity between different batches of this enzyme. To analyze the PDD efficiency qPCR with primers designed to targetable sequences and off-targets can be performed with treated library and its control counterpart.

Anticipated Results

If the instructions provided in this unit are followed closely at least 94% of the rRNA reads (or other targeted unwanted abundant sequences) will be eliminated from the constructed cDNA library. If higher depletion efficiency is required, conducting two successive rounds of PDD depletion with an intervening purification step may be preferable to trying to alter any of the PDD reaction conditions or probe densities (although we have not attempted this). The resulting library can be sequenced with Illumina MySeq/HiSeq and similar instruments using Illumina's TruSeq Small RNA Sample Prep Kit or another set of reagents with compatible primers and hexanucleotide indexes. With the amounts of the components scaled down accordingly, it will be possible to confidently generate a depleted cDNA library from about 100 ng of the total unfractionated starting cellular RNA. In respect to the initial RNA fragments, there will be no distortion of size and sequence expectance rates and positions of 5' and 3' extremities detectable in the library inserts within the mode of size-selection if compared to the library obtained without the PDD step. The typical outcome of the PDD application demonstrating extinction of the rRNA sequences from the library cDNA size-selected for the 20 to 300 nt range is exemplified in Figure 11.15.4A. Figure 11.15.4B schematically illustrates the expected depletion rates introduced by an isolated PDD probe. Note that the slope on the left of the depletion expectance curve depends on the modal insert length present in the cDNA library and on stringency of the size selection (particularly the minimum size cutoff). Shorter insert sizes and narrower fragment ranges will result in a steeper curve; therefore if depletion specificity is of an extreme importance, even greater spacing of the

probes can be employed at the cost of the depletion depth, especially when the library has a strict minimum insert length. We did not, however, observe more depletion for regions targeted by multiple probes (Fig. 11.15.4C). Observable from these results is the need to keep probe spacing regular, as the rRNA read representation (insert length > 119 nt) in the PDD-treated library rapidly increased to that of the untreated library when a gap of more than ~100 nt exists between PDD probes.

Time Considerations

The complete experiment, including the anticipated preparation of the RNA sample and all steps of Support Protocols 1 to 6 and Basic Protocols 1 and 2, would normally take between 1 and 2 weeks depending on the intensity of the work. Basic Protocols 1 and 2 can be completed in 2 days together. Support Protocol 1 will require from 1 to 8 hr and can be completed in advance of other steps, as oligonucleotide manufacture and delivery will require some waiting time. Support Protocol 2 can be completed in 1 hr up to the ethanol precipitation stage (step 6), which will require minimum extra 3 hr, or work can be stopped at this point. Support Protocol 3 will require 5 hr (including 2 hr hands-on) and can be stopped at the step 17, steps 17 and 18 will take at least 3 hr and can be a stopping point. Support Protocol 4 can be completed in 2 hr if four or less independent samples are processed in parallel. Support Protocol 5 can be completed in 2 hr plus 15 min for each individual sample if processing several samples in parallel. We do not recommend processing more than four samples simultaneously for Support Protocols 4 and 5. Support Protocol 5 can be a stopping point at step 28. Support Protocol 6 will require 6 hr total with about 1 hr of hands-on experience, and can be stopped at step 23. Basic Protocol 1 will take about 2 hr plus 15 min per each individual sample in processing several samples in parallel. We do not recommend to process more than four samples simultaneously. Basic Protocol 2 will require about 2 hr for sample amplification and can be stopped at step 8. It will further take about 2 to 3 hr for size-selection with additional 15 min per sample if processing several samples in parallel.

Acknowledgements

This work was supported by an Australian Research Council Discovery Grant (DP1300101928) and a NHMRC Senior Research Fellowship (514904) awarded to TP. NES was supported by a Go8 European

Fellowship. We acknowledge technical support from the Australian Cancer Research Foundation Biomolecular Resource Facility.

Literature Cited

- Allawi, H.T. and SantaLucia, J. Jr. 1997. Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry* 36:10581-10594.
- Anisimova, V.E., Rebrikov, D.V., Shagin, D.A., Kozhemyako, V.B., Menzorova, N.I., Staroverov, D.B., Ziganshin, R., Vagner, L.L., Rasskazov, V.A., Lukyanov, S.A., and Scheglov, A.S. 2008. Isolation, characterization and molecular cloning of duplex-specific nuclease from the hepatopancreas of the Kamchatka crab. *BMC Biochem.* 9:14.
- Archer, S.K., Shirokikh, N.E., and Preiss, T. 2014. Selective and flexible depletion of problematic sequences from RNA-seq libraries at the cDNA stage. *BMC Genomics* 15:401.
- Blondal, T., Thorisdottir, A., Unnsteinsdottir, U., Hjorleifsdottir, S., Aevarrsson, A., Ernstsson, S., Fridjonsson, O.H., Skirmisdottir, S., Wheat, J.O., Hermansdottir, A.G., Sigurdsson, S.T., Hreggvidsson, G.O., Smith, A.V., and Kristjansson, J.K. 2005. Isolation and characterization of a thermostable RNA ligase 1 from a *Thermus scotoductus* bacteriophage TS2126 with good single-stranded DNA ligation properties. *Nucleic Acids Res.* 33:135-142.
- Bogdanova, E.A., Shagina, I.A., Mudrik, E., Ivanov, I., Amon, P., Vagner, L.L., Lukyanov, S.A., and Shagin, D.A. 2009. DSN depletion is a simple method to remove selected transcripts from cDNA populations. *Mol. Biotechnol.* 41:247-253.
- Bogdanova, E.A., Barsova, E.V., Shagina, I.A., Scheglov, A., Anisimova, V., Vagner, L.L., Lukyanov, S.A., and Shagin, D.A. 2011. Normalization of full-length-enriched cDNA. *Methods Mol. Biol.* 729:85-98.
- Castro, T.L., Seyffert, N., Ramos, R.T., Barbosa, S., Carvalho, R.D., Pinto, A.C., Carneiro, A.R., Silva, W.M., Pacheco, L.G., Downson, C., Schneider, M.P., Miyoshi, A., Azevedo, V., and Silva, A. 2013. Ion Torrent-based transcriptional assessment of a *Corynebacterium pseudotuberculosis* equi strain reveals denaturing high-performance liquid chromatography a promising rRNA depletion method. *Microb. Biotechnol.* 6:168-177.
- Christodoulou, D.C., Gorham, J.M., Herman, D.S., and Seidman, J.G. 2011. Construction of normalized RNA-seq libraries for next-generation sequencing using the crab duplex-specific nuclease. *Curr. Protoc. Mol. Biol.* 94:4.12.1-4.12.11.
- Diatchenko, L., Lau, Y.F., Campbell, A.P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E.D., and Siebert, P.D. 1996. Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci. U.S.A.* 93:6025-6030.
- Dolinsek, J., Dorninger, C., Lagkouvardos, I., Wagner, M., and Daims, H. 2013. Depletion of unwanted nucleic acid templates by selective cleavage: LNAzymes, catalytically active oligonucleotides containing locked nucleic acids, open a new window for detecting rare microbial community members. *Appl. Environ. Microbiol.* 79:1534-1544.
- Ederly, I., Chu, L.L., Sonenberg, N., and Pelletier, J. 1995. An efficient strategy to isolate full-length cDNAs based on an mRNA cap retention procedure (CAPture). *Mol. Cell Biol.* 15:3363-3371.
- Gijvanekar, C., Strych, U., Fofanov, Y., Fox, G.E., and Willson, R.C. 2012. Rare target enrichment for ultrasensitive PCR detection using cot-rehybridization and duplex-specific nuclease. *Anal. Biochem.* 421:81-85.
- Gumport, R.I. and Uhlenbeck, O.C. 1981. T4 RNA ligase as a nucleic acid synthesis and modification reagent. *Gene Amplif. Anal.* 2:313-345.
- Harbers, M. 2008. The current status of cDNA cloning. *Genomics* 91:232-242.
- Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324:218-223.
- Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M., and Weissman, J.S. 2012. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat. Protoc.* 7:1534-1550.
- Itaya, M. and Kondo, K. 1991. Molecular cloning of a ribonuclease H (RNase HI) gene from an extreme thermophile *Thermus thermophilus* HB8: A thermostable RNase H can functionally replace the *Escherichia coli* enzyme in vivo. *Nucleic Acids Res.* 19:4443-4449.
- Jackson, T.J., Spriggs, R.V., Burgoyne, N.J., Jones, C., and Willis, A.E. 2014. Evaluating bias-reducing protocols for RNA sequencing library preparation. *BMC Genomics* 15:569.
- Jager, D., Sharma, C.M., Thomsen, J., Ehlers, C., Vogel, J., and Schmitz, R.A. 2009. Deep sequencing analysis of the *Methanosarcina mazei* Go1 transcriptome in response to nitrogen availability. *Proc. Natl. Acad. Sci. U.S.A.* 106:21878-21882.
- Ko, M.S. 1990. An 'equalized cDNA library' by the reassociation of short double-stranded cDNAs. *Nucleic Acids Res.* 18:5705-5711.
- Lehman, I.R. and Nussbaum, A.L. 1964. The deoxyribonucleases of *Escherichia coli*. V. On the specificity of exonuclease I (phosphodiesterase). *J. Biol. Chem.* 239:2628-2636.
- Levin, J.Z., Yassour, M., Adiconis, X., Nusbaum, C., Thompson, D.A., Friedman, N., Gnirke, A., and Regev, A. 2010. Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nat. Methods* 7:709-715.
- Li, S.K., Zhou, J.W., Yim, A.K., Leung, A.K., Tsui, S.K., Chan, T.F., and Lau, T.C. 2013.

- Organism-specific rRNA capture system for application in next-generation sequencing. *PLoS One* 8:e74286.
- Maruyama, K. and Sugano, S. 1994. Oligo-capping: A simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* 138:171-174.
- Meador, J. 3rd, Cannon, B., Cannistraro, V.J., and Kennell, D. 1990. Purification and characterization of *Escherichia coli* RNase I. Comparisons with RNase M. *Eur. J. Biochem.* 187:549-553.
- Miller, D.F., Yan, P.S., Buechlein, A., Rodriguez, B.A., Yilmaz, A.S., Goel, S., Lin, H., Collins-Burow, B., Rhodes, L.V., Braun, C., Pradeep, S., Rupaimoole, R., Dalkilic, M., Sood, A.K., Burow, M.E., Tang, H., Huang, T.H., Liu, Y., Rusch, D.B., and Nephew, K.P. 2013. A new method for stranded whole transcriptome RNA-seq. *Methods* 63:126-134.
- Morlan, J.D., Qu, K., and Sinicropi, D.V. 2012. Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue. *PLoS One* 7:e42882.
- O'Neil, D., Glowatz, H., and Schlumpberger, M. 2013. Ribosomal RNA depletion for efficient use of RNA-seq capacity. *Curr. Protoc. Mol. Biol.* 103:4.19.1-4.19.8.
- Pang, X., Zhou, D., Song, Y., Pei, D., Wang, J., Guo, Z., and Yang, R. 2004. Bacterial mRNA purification by magnetic capture-hybridization method. *Microbiol. Immunol.* 48:91-96.
- Peng, R.H., Xiong, A.S., Xue, Y., Li, X., Liu, J.G., Cai, B., and Yao, Q.H. 2008. Kamchatka crab duplex-specific nuclease-mediated transcriptome subtraction method for identifying long cDNAs of differentially expressed genes. *Anal. Biochem.* 372:148-155.
- Romaniuk, P.J. and Uhlenbeck, O.C. 1983. Joining of RNA molecules with RNA ligase. *Methods Enzymol.* 100:52-59.
- Rosenow, C., Saxena, R.M., Durst, M., and Gingeras, T.R. 2001. Prokaryotic RNA preparation methods useful for high density array analysis: Comparison of two approaches. *Nucleic Acids Res.* 29:E112.
- Scotto-Lavino, E., Du, G., and Frohman, M.A. 2006. 5' end cDNA amplification using classic RACE. *Nat. Protoc.* 1:2555-2562.
- Shagin, D.A., Rebrikov, D.V., Kozhemyako, V.B., Altshuler, I.M., Shcheglov, A.S., Zhulidov, P.A., Bogdanova, E.A., Staroverov, D.B., Rasskazov, V.A., and Lukyanov, S. 2002. A novel method for SNP detection using a new duplex-specific nuclease from crab hepatopancreas. *Genome Res.* 12:1935-1942.
- Tate, C.M., Nunez, A.N., Goldstein, C.A., Gomes, I., Robertson, J.M., Kavlick, M.F., and Budowle, B. 2012. Evaluation of circular DNA substrates for whole genome amplification prior to forensic analysis. *Forensic Sci Int. Genet.* 6:185-190.
- Vandernoot, V.A., Langevin, S.A., Solberg, O.D., Lane, P.D., Curtis, D.J., Bent, Z.W., Williams, K.P., Patel, K.D., Schoeniger, J.S., Branda, S.S., and Lane, T.W. 2012. cDNA normalization by hydroxyapatite chromatography to enrich transcriptome diversity in RNA-seq applications. *Biotechniques* 53:373-380.
- Walker, G.C., Uhlenbeck, O.C., Bedows, E., and Gumport, R.I. 1975. T4-induced RNA ligase joins single-stranded oligoribonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 72:122-126.
- Wang, L.K., Lima, C.D., and Shuman, S. 2002. Structure and mechanism of T4 polynucleotide kinase: An RNA repair enzyme. *EMBO J.* 21:3873-3880.
- Wickersheim, M.L. and Blumenstiel, J.P. 2013. Terminator oligo blocking efficiently eliminates rRNA from *Drosophila* small RNA sequencing libraries. *Biotechniques* 55:269-272.
- Wu, H., Lima, W.F., and Croke, S.T. 1999. Properties of cloned and expressed human RNase H1. *J. Biol. Chem.* 274:28270-28278.
- Yi, H., Cho, Y.J., Won, S., Lee, J.E., Jin Yu, H., Kim, S., Schroth, G.P., Luo, S., and Chun, J. 2011. Duplex-specific nuclease efficiently removes rRNA for prokaryotic RNA-seq. *Nucleic Acids Res.* 39:e140.
- Zhenodarova, S.M., Kliagina, V.P., Maistrenko, F.G., Pustoshilova, N.M., and Smolianinova, O.A. 1989. Substrate specificity of T4 RNA-ligase. The effect of the nucleotide composition of substrates and the size of phosphate donor on the effectiveness of intermolecular ligation. *Bioorg. Khim.* 15:478-483.

Key Reference

Archer et al., 2014. See above.

The original research article outlining the PDD method and analyzing its applicability for construction of RNA-seq libraries.

Internet Resources

<http://github.com/stu2/blast2ape>

GitHub repository for Blast2Ape scripts.

<http://www.biomedcentral.com/content/supplementary/1471-2164-15-401-s3.zip>

Location of original paper supplementary files including alternative Perl-based script.

[http://biologylabs.utah.edu/jorgensen/wayned/ape/ApE software.](http://biologylabs.utah.edu/jorgensen/wayned/ape/ApE software)

<http://www.biopython.org>

Biopython software.

<http://www.ncbi.nlm.nih.gov/blast>

NCBI BLAST sequence homology search.

<http://primer3plus.com/cgi-bin/dev/primer3plus.cgi>

Primer3 software.

<http://www.idtdna.com>

Web site of Integrated DNA Technologies (IDT) containing instruments for oligonucleotide ordering and some useful computational tools for nucleic acids analysis.