**SMART protocol for RNA-seq – revised June 3, 2016**

**0. Notes on RNA samples**

The following protocol assumes that one is studying eukaryotic gene expression. If the RNA is from bacteria, then step 1 is skipped. RNA may be isolated using any of a number of commercially-available kits or homemade protocols; individual labs will probably have their own preferences, driven largely by the organism being studied. Regardless, RNA quality is important and should be confirmed – preferably, on a Bioanalyzer (RIN scores >8 are aimed for) or on ethidium-stained agarose genes (where clean and abundant rRNA bands are expected, as well as little or no very low MW material).

In the procedures that follow, not much attention is paid to quantifying RNA or cDNA throughout the process. We usually try to start with 1 µg of total RNA, but have had success with as little as 100 ng per samples. Regardless, the quantity of RNA is determined and recorded at the outset, and consistency in recovery of RT and PCR products is assumed thereafter.

**1. Poly(A) enrichment**

1. Aliquot 10 µL of oligo(dT) beads (NEB).

[NOTE – this amount of the NEB beads suffices for quantities of total RNA below 10 µg. We have not explored larger quantities of RNA, as this seems exorbitant. Also, we have used as little as 100 ng of total RNA with good results. Lower quantities should work as well, but one should keep in mind that non-specific binding of RNA to beads may reduce the enrichment of poly(A) RNA. Also, it is probable that the starting amounts of poly(A) RNA may fall below the affinities of RT for template DNA if one starts with less than 100 ng of total RNA. This needs to be accounted for in the following steps (specifically, step 2), possibly by extending RT times and perhaps periodically adding more enzyme.]

1. Wash the beads twice with 100µL of Binding Buffer (20mM Tris-HCl pH 7.5, 1.0M LiCl and 2mM EDTA), and remove the supernatant.
2. Resuspend the beads in 50 µl of Binding Buffer.
3. Bring RNA to 50 µL with RNAse-free water. Heat to 65°C, cool on ice, then add to the washed beads. Incubate at room temperature for 5 min.
4. Collect beads using the magnetic stand, discard supernatant.
5. Wash the beads twice with 100µL of Washing Buffer B (10mM Tris-HCl PH 7.5, 0.15M LiCl, 1mM EDTA).
6. Remove the supernatant from the beads, add 15 µL of 10mM Tris-HCl (this depends on how much RNA is used, and how many samples the polyA RNA is going to be used for) and heat the beads at 800C for 3 minutes to elute mRNA.
7. Remove beads with the magnetic stand, save the supernatant.

[NOTE: it may be advisable to repeat this, to remove as much rRNA as possible. However, this is still an open issue, since organellar and stable RNAs are known to be polyadenylated, which means repeated rounds of poly(A) enrichment yielding reduced returns.

Also, since we use NEB oligo-dT beads, it should be OK to follow NEB’s protocol for poly(A) enrichment instead of the one I described in the preceding. NEB and other companies sells kits as well as just the beads, and some may prefer to use the ready-made solutions and protocol.]

**2. RNA fragmentation, cDNA synthesis, and clean-up**

1. mix 14.5 µL RNA, 1 µL (=100 pmol) RT primer + 5 µL 5X 1st strand buffer.
2. heat to 95 °C for 2 min, chill.

[NOTE – the high temperature in the RT buffer is a suitable and efficient proxy for stand-alone RNA fragmentation kits or systems. By folding fragmentation directly into the RT reactions, the process is streamlined.]

1. **Immediately** add (while cold!):

2.5 µL 10X dNTPs for reverse transcription reactions

1 µL 100 mM DTT

0.5 µL RNase Inhibitor

1 µL SMARTSCRIBE

1. Mix and incubate for 120 min at 42°C.
2. Add 1 µL of the strand-switching primer (SMART7.5) and an additional 1 µL of SMARTSCRIBE, mix and incubate for an additional 120 min at 42°C.
3. Heat to 70 °C for 5 min.
4. Remove 10 µL of the reaction to a new 500 µL tube and add 15 µL of 10 mM Tris HCl pH7.5. Cap and save the remaining cDNA reaction (this is kept in reserve in case something bad happens in the subsequent parts of the procedure).
5. To the diluted cDNA, add **16.25** **µL** SPRI beads that have been completely mixed and brought to room temperature. Incubate for 8 min at room temperature.
6. Separate beads using the magnet stand (this may take a minute or two, because of the viscosity of the bead solution), remove and discard the supernatant.
7. While the tube is on the magnet stand, add 100 µL **fresh** 80% ethanol. After 5-30 seconds (it doesn't matter, the point here is to wash the pellet; no need to resuspend the beads), remove and discard the supernatant. Repeat the wash.
8. Air dry the washed beads for 10 min at room temperature.
9. Add 100 µL 10 mM Tris HCl, pH 7.5. Mix, and then collect the beads with the magnetic stand. Remove the supernatant to a new 500 µL thin-walled microcentrifuge tube.
10. Add **55** **µL** SPRI beads that have been completely mixed and brought to room temperature. Incubate for 8 min at room temperature.
11. Separate beads using the magnet stand (this may take a minute or two, because of the viscosity of the bead solution), **remove and save the supernatant** to a new 500 µL tube.
12. Add **10** **µL** SPRI beads to the reserved supernatant from step m. Incubate for 10 min at room temperature.
13. Separate beads using the magnet stand, remove and discard the supernatant.
14. While the tube is on the magnet stand, add 100 µL **fresh** 80% ethanol. After 5-30 seconds (it doesn't matter, the point here is to wash the pellet; no need to resuspend the beads), remove and discard the supernatant. Repeat the wash. **(Note that at this point the amount of beads is small. Care should be taken to avoid disturbing the pellet or otherwise losing the beads.)**
15. Air dry the washed beads for 10 min at room temperature.
16. Add 25 µL 10 mM Tris HCl, pH 7.5. Mix, and then collect the beads with the magnetic stand. Remove the supernatant to a new 500 µL thin-walled microcentrifuge tube. **This is the final library to be used for PCR amplifications.**

[The two-step process for RT and strand-switching was arrived at empirically in my lab, with guidance from some prior research publications. For us, the goal has been to develop procedures that can be used with sub-microgram quantities of RNA. I reasoned that a limiting factor would be the performance of RT, which is likely working below the Km’s for template. The easiest way to address this is to extend incubation times, hence the long reaction times.

Also, while we have incubated the reverse transcription reactions at 42°C routinely for our libraries, it is probably OK to use 37°C as well. Lowering the temperature may reduce possible GC biases in terms of where cDNA synthesis is initiated on the template. We have not looked into this experimentally, but this is something to keep in mind.

The elaborate SPRI bead process used here has been empirically determined to yield libraries with a size range of 300-600 bp, with most of the library in the 500 bp range. This process can be fine-tuned, and probably needs to be checked occasionally to idemntify lot-to-lot variability in the SPRI beads.]

**3. PCR amplifications**

Set up Phire reactions using the PE-PCR1 and PE-PCR2 primers. The reaction volumes should be 25 µL, and 1 µL of the library should be used as a template. The annealing temp is 60°C (**for 15 sec**), and the extension time is 60 sec. *(In principle, the extension time can be decreased, perhaps to as little as 30 sec; I use 60 sec to make sure every amplicon is completed, so that no odd PCR strand-switching artifacts arise later due to the buildup of truncated products.)*

As far as cycle # is concern, a bit of range finding is usually needed. With 1 µg of total Arabidopsis RNA, 15 cycles usually works fine. Some members of my lab have done as few as 9 cycles here. At the other end of the spectrum, we have made libraries after as many as 24 cycles of amplification. As a rule of thumb, we try to avoid libraries involving more than 18 or 21 cycles, but sometimes this is unavoidable. Cycle numbers greater than 18 should not be needed if starting with 100 ng or more of total RNA.

Regardless of these considerations, the goal is a smear between 300 and 600 bp; if successful, amplifications can be scaled up and tags purified using an abbreviated SPRI bead protocol:

1. To each 25 µL PCR reaction, add **16.25** **µL** SPRI beads that have been completely mixed and brought to room temperature. Incubate for 8 min at room temperature.
2. Separate beads using the magnet stand (this may take a minute or two, because of the viscosity of the bead solution), remove and discard the supernatant.
3. While the tube is on the magnet stand, add 100 µL **fresh** 80% ethanol. After 5-30 seconds (it doesn't matter, the point here is to wash the pellet; no need to resuspend the beads), remove and discard the supernatant. Repeat the wash.
4. Air dry the washed beads for 10 min at room temperature.
5. Add 25 µL 10 mM Tris HCl, pH 7.5. Mix, and then collect the beads with the magnetic stand. Remove the supernatant to a new 500 µL thin-walled microcentrifuge tube. **This is the final library that will be submitted for Bioanalyzer analysis and sequencing.**

[NOTES: 1. We use Phire routinely for this step, because of a combination of the excellent and reliable performance of the enzyme and its relatively low cost (compared with some other hot-start enzymes). On occasion, we have used other Taq polymerases successfully, including “home-made” enzyme. The choice to include Phire here is intended to provide some measure of consistency; different enzymes can probably be used, but amplification conditions need to be optimized.

2. Perhaps the most unsettling aspect of this is the fact that, often, one cannot know about library quality until the sequencing results are returned – this means one runs the risk of spending money on poor libraries. In my lab’s experience, size-selection and re-amplification is a good quality-control check for libraries. The most persistent problem with RNA-seq libraries is the presence of short amplification products that arise due to the direct copying of the strand-switching primer by RT, using the RT primer as a primer for reverse transcription. These can easily over-amplify and give smears that can be mistaken for good libraries. If these artifacts pass through the size selection and are re-amplified, the results are not a narrow band corresponding to the region expected based on the size selection process, but a long smear extending from 100 bp or so through the 1000+ bp size range. Some examples are shown at the end of this document.]

Primers

>RT-PE6 series (underlined bases “xxx” are where bar codes may be added)

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNxxxNNNNNN

[Note – this is the primer design that we have used to obtain libraries with good coverage and reasonably random distributions of tags. It should be pointed out that, with at least two (of more than 20) sequencing runs we have done, the first base of the run was called “N” on almost all sequences. For this reason, when we assign bar codes and design pooling strategies, we work with the assumption that the second and third “x”’s will be the informative positions. This is why we include an “N” at the very beginning of the sequencing region.

The optimal outcome is to have perfectly random priming via the NNNNNN sequence – this will be reflected in a random distribution of reads through a transcriptome. Of course, with this as well as other methods, perfect randomization is not likely; for this procedure, the nature of the bar code will probably bias priming position. We have not explored this to any extent, but I am confident that the distribution and coverage we get is comparable to other library methods. The best hedge against possible bar-code biases is replication with different bar codes.]

>RT-PE6a

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNAGGNNNNNN

>RT-PE6b

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNCGGNNNNNN

>RT-PE6c

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNAACNNNNNN

>RT-PE6d

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNAGCNNNNNN

>RT-PE6e

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNACGNNNNNN

>RT-PE6f

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNAGANNNNNN

>RT-PE6g

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNCCGNNNNNN

>RT-PE6h

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNCAANNNNNN

>RT-PE6j

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNCAGNNNNNN

>RT-PE6k

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNGGANNNNNN

>RT-PE6l

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNCCANNNNNN

>RT-PE6m

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNCACNNNNNN

>RT-PE6n

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNGAGNNNNNN

>RT-PE6o

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNGCANNNNNN

>RT-PE6p

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNGAANNNNNN

>RT-PE6q

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNGGCNNNNNN

>RT-PE6r

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNACCNNNNNN

>RT-PE6s

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNGACNNNNNN

>RT-PE6t

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNCCCNNNNNN

>RT-PE6u

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNCGANNNNNN

>RT-PE6v

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNCGCNNNNNN

>RT-PE6w

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNGCGNNNNNN

>RT-PE6x

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNGCCNNNNNN

>RT-PE6y

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNAAGNNNNNN

>RT-PE6z

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNAAANNNNNN

>SMART7.5 (“+G” denotes a locked nucleic acid)

CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTGG+G

[This primer is designed to anneal to the expected CCC overhang added by the Clontech RT enzyme to the 3’ end of the first strand cDNA, and thereafter serve as a template for extension of the cDNA. The Zhu paper describes this, as do various Clontech manuals.]

>PE-PCR1

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

>PE-PCR2

CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT

[These are the adapters needed for Illumina sequencing. Note that some of these sequences are present in the RT and SMART primers.]

References (some of the papers that guide development of the method)

Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD (2001) Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. BioTechniques 30 (4):892-897.

Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G, Sandberg R (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nature methods 10 (11):1096-1098. doi:10.1038/nmeth.2639.

Picelli S, Faridani OR, Bjorklund AK, Winberg G, Sagasser S, Sandberg R (2014) Full-length RNA-seq from single cells using Smart-seq2. Nature protocols 9 (1):171-181. doi:10.1038/nprot.2014.006.

Pinto FL, Lindblad P (2010) A guide for in-house design of template-switch-based 5' rapid amplification of cDNA ends systems. Analytical biochemistry 397 (2):227-232. doi:10.1016/j.ab.2009.10.022.



Legend

The gel on the left shows the amplifications of the size-selected cDNAs, amplified for 12, 15, and 18 cycles. Note the range of the smear. The gel on the right shows an amplification of failed libraries. Note the small size range for the smear. with the cycle numbers shown.

Library quality is deemed excellent if the primary reamplification product is in the range shown in the gel on the left. Acceptable quality is reflected in reamplification results in which products are largely of the same size as the excised gel, but also of smaller size. Unacceptable quality (library failure) is reflected in reamplification products with little or no DNA of the expected size.