RNA-Seq *Panicum* sample preparation for sequencing on the Illumina HiSeq - v4

Following the protocol developed by Meyer & Matz RNA-Seq sample preparation for sequencing on the SOLiD System (Meyer, Aglyamova et al. 2011).

Updated October 16, 2013 to include new Matz lab's changes.

Updated July 8, 2013 to include NucleoFast PCR purification, new RNA oligo, a set of barcode to be used together following a Hamming distance analysis.

Updated July 19, 2011 to include magnetic bead based PCR purification.

Important points before starting

This protocol describes in detail the procedures used to prepare cDNA fragment libraries for quantitative analysis of gene expression (RNA-Seq) by deep sequencing on the Illumina HiSeq System.

About 0.5-1 µg (at least 100 ng) of DNAse-treated total RNA is required per sample, and this starting material should be carefully quantified and analyzed by gel electrophoresis prior to beginning these procedures to verify that the RNA is intact, and free of genomic DNA contamination.

The procedure can be reasonably completed within three days.

Day 1: RNA is fragmented and used to synthesis cDNA (steps 1-2).

Day 2: cDNA is amplified, sample-specific barcodes are incorporated, and size-selection is accomplished by means of gel extraction (steps 3-4).

Day 3: the preparations are evaluated by PCR quality and quantified by qPCR.

The sequences of all oligonucleotide primers used in this protocol and images from gel examples are provided at the end of this document.

Procedure

1. RNA fragmentation

   NOTE: the buffer in which the original RNA is incubated is critically important for the success of fragmentation, as are the volume and concentration of the RNA. Prior to working with the precious experimental samples, we recommend testing a range of different incubation times to identify the duration that produces the appropriate size range in these samples. In Panicum spp we found that 10 minutes are enough.

   a. Aliquot 1 µg of total RNA in 10 µl of 10 mM Tris (pH 8.0). If concentrations are not sufficiently high to allow this loading, RNA samples can be concentrated by drying in Speedvac (without heating) or by standard ethanol or LiCl precipitation. Set aside an additional sample (~100 ng) of the original intact RNA for comparison with the fragmented samples.

   b. Incubate RNA at 95°C for 10 minutes to fragment to the desired size range (400 – 600 bp) (See figure 1). This can be most easily accomplished in a thermocycler.

   c. Analyze 100 ng from fragmented RNA, alongside the intact RNA from the same sample, on a standard 2% agarose gel to evaluate the extent of RNA fragmentation. The smear must extend all the way up into the region where ribosomal RNA bands were, while the bands themselves should be mostly gone.

2. First-strand cDNA synthesis

   NOTE: if RNA quantity is not limiting, first-strand cDNA should be synthesized using 1 µg of fragmented RNA to ensure adequate representation of all transcripts. The reaction shown below

is intended for ~ 1 µg; this can be doubled or halved as needed if the amount of RNA is very different from 1 µg.

a. Measure the volume remaining after fragmentation using a pipette. The following recipe assumes a starting volume of 9 µl (10 µl minus evaporation), so if the volume is lower than this add additional water to achieve 9 µl.

b. Add 1 µl of the 10 µM oligonucleotide 3ILL-30TV to each well. Incubate at 65°C for 3 minutes in a thermocycler, then transfer immediately onto ice.

c. Prepare a cDNA synthesis master mix. The following volumes are intended for a single reaction, so multiply these values by the number of reactions plus a small amount (~10%) to account for pipetting error.

	(Volumes given in µl)
H_2O	1
dNTP (10 mM each)	1
DTT (0.1 M)	2
5X first-strand buffer	4
10 µM S-ILL-swMW (RNA oligonucleotide; stored at -80°C)	1
SuperScript II Reverse Transcriptase (Invitrogen #18064-022)	1
Total volume	10

d. Add 10 µl of this master mix to the RNA from (2b), mix thoroughly, and incubate in a thermocycler for one hour at 42°C.

e. Incubate at 65°C for 15 minutes to inactivate the RT, store on ice or at -20°C until ready to proceed to the next step.

3. cDNA amplification

NOTE: it is important not to over-amplify the cDNA at this stage, to avoid artifacts and distortion of expression ratios. The fewer cycles used, the better. If a visible smear is not produced in 15 cycles, try repeating the PCR with additional template (FS-cDNA from step 2e), up to a maximum of 6 µl, and correspondingly less water in the master mix. If no smear is detected with 6 µl template and 15 cycles, this indicates a problem with the first-strand cDNA synthesis.

a. Two test-scale PCRs are prepared for a representative subset (6-8) of samples to verify yield and specificity of the reaction. The control amplification (A), lacking the 5'-specific primer, should be empty. The following volumes are intended for a single reaction each, so multiply these values by the total number of reactions plus a small additional amount to account for pipetting error. This recipe assumes 2 µl of template (from step 2e), so if you use a different amount of template, adjust the water accordingly.

	(Volumes given in µl)	
Reagent	A	B
H_2O	13.2	11.8
dNTP (2.5 mM ea)	2	2
10X PCR buffer	2	2
10 µM 5ILL oligo	0	0.4
10 µM 3ILL-30TV oligo	0.4	0.4
Titanium Taq polymerase (Clontech # 639208)	0.4	0.4
FS-cDNA template from step 2e	2	2

Total volume	20	20

b. Amplify in a thermocycler using the following profile:

95°C 5 min, (95°C 40 sec, 63°C 2 min, 72°C 1 min) X 13 cycles

c. After 13 cycles check 5 µl of the PCR products for all reactions on a standard 2 % agarose gel. A "smear" of cDNA (~200-500 bp) should be faintly visible in reaction B (Figure 2), and nothing should be detected in the reactions A. If nothing is detected in reaction D, you can continue the reaction for additional cycles (up to a maximum of 19), or repeat the reaction with additional template. If product ever appears in reactions A, this indicates too much template, too many cycles, or contamination in one or more reagents. The goal at this stage is to identify the minimum cycle number that produces a visible smear in reaction B while remaining clean in reactions A.

NOTES:

- If you started with large amount (1 µg) of total RNA you might see a carry-over degraded RNA smear on the gel, in both A and B reactions. Do not confuse it with the PCR product! Make sure the product actually accumulates as you are adding more cycles.

- Different samples might require slightly different number of cycles, this is OK since all the potential biases due to PCR amplification will be removed at the data analysis stage by discarding PCR duplicates.

Very important: if a smear is not produced after 19 cycles, the representation of the cDNA is not adequate for RNA-seq; you must optimize previous stages. Ideal RNA-seq results can be obtained for samples that are amplified in 15 cycles or less, 17 cycles is OK.

d. Once the optimum amount of template and number of cycles have been determined (3e), prepare a single large-scale reaction for each cDNA sample as follows. This recipe assumes 5 µl of template, so if you use more template adjust the water accordingly.

	(all volumes given in µl)
H_2O	32
dNTP (2.5 mM ea)	5
10X PCR buffer	5
10 µM 5ILL oligo	1
10 µM 3ILL-30TV oligo	1
Titanium Taq polymerase	1
FS-cDNA template from step 2e	5
Total volume	50

e. Amplify in a thermocycler using the following profile:

95°C 5 min, (95°C 40 sec, 63°C 2 min, 72°C 1 min) X 13 cycles

f. After PCR, check 5 µl of the product on a 2 % agarose gel to verify that the reaction worked as expected before freezing or purifying the product.

g. Purify PCR products using Macherey-Nagel NucleoFast PCR Clean-up protocol (Cat. No. 743500.4) according to the manufacturer's instructions:

   a. Transfer PCR samples (~45 ul each) to NucleoFast 96 PCR plate.

   b. Filter contaminants to waste under vacuum for 12 min.

   c. Wash membrane with 100 ul of RNase-free water under vacuum for 12 min.

d. Recover purified PCR samples add 30 ul of RB, incubate for 5 minutes at room temperature after the addition.

e. Quantify the purified products by OD260 (Nanodrop).

h. Prepare 30 µl of the purified PCR products diluted to 5 ng μl^{-1} (in 10 mM tris HCl pH 8, or the RB elution buffer from the PCR-cleanup kit).

4. Adaptor extension and size selection

NOTE: Because the size distribution of templates is a critical factor for successful emulsion PCR, any templates intended for sequencing on the Illumina System should be carefully size-selected prior to emPCR. The directions below outline a simple procedure for selecting fragments ranging from 350-500 bp in size that does not require any special equipment. Other methods of size selection could be substituted provided they achieve this same size range.

a. First, four test-scale PCRs are prepared for a representative subset (6-8) of samples to verify yield and specificity of the reaction, each using 10 ng (2 µl) PCR product (step 3h above) as template.

b. Prepare four separate master mixes for small-scale test PCR. The following volumes are for a single reaction, so multiply these values by the total number of samples plus a small additional amount to account for pipetting error. The values shown here assume the use of 10 ng (2 µl from step 3h above) clean PCR product as template, so if you change this be sure to change the volume of water accordingly. Be sure to write down which barcode and multiplex is assigned to each sample at this stage, since this cannot be easily determined later in the process.

	(volumes given in µl)			
	A	B	C	D
H2O	5.8	5.6	5.6	5.4
dNTP (2.5 mM ea)	1	1	1	1
10X PCR buffer	1	1	1	1
Multiplex oligo (10 µM)	0	0.2	0	0.2
Barcode oligo (10 µM)	0	0	0.2	0.2
Titanium Taq polymerase	0.2	0.2	0.2	0.2
cDNA template from step 3h	2	2	2	2
Total volume	10	10	10	10

c. Amplify in a thermocycler using the following profile:

95°C 5 min, (95°C 40 sec, 63°C 2 min, 72°C 1 min) X 4 cycles

d. After 4 cycles check 5 µl of the PCR products for all reactions on a gel. The ideal result is a faint smear in reaction D, with no visible product in reactions A-C (Figure 3). If products appear in these control reactions, consider using less template. If nothing is detected in reaction D, add 1-2 more cycles and check the results on a gel. If no product is visible before 6 cycles, repeat the reaction with a larger volume of template (in our experience this has never been required). A small amount of product in the controls can be tolerated, but if reactions A-C are comparable in intensity to reaction D something is wrong.

e. When the optimum number of cycles and volume of template have been determined, prepare a large-scale reaction based on those parameters with 50 ng (10 µl of the diluted purified cDNA, step 3i) template in 50 µl total volume. The following master mix assumes the use of 10 µl of template per 50 µl reaction; if you adjust this template volume be sure to adjust the volume of water accordingly. This recipe is for a single reaction, so multiple these values by the number of samples to be prepared plus a small additional amount for pipetting error.

(volumes given in µl)	
H2O	27
dNTP (2.5 mM ea)	5
10X PCR buffer	5
Multiplex oligo (10 µM)	1
Titanium Taq polymerase	1
PCR template from step 3h	10
Barcode oligo (10 µM)	1
Total volume	50

f. Amplify these reactions using the same profile and cycle number as determined above.

95°C 5 min, (95°C 40 sec, 63°C 2 min, 72°C 1 min) X 4 cycles

g. Prepare a gel for size selection. This preparative gel should consist of 2% agarose in 1X TBE buffer, with SYBR Safe DNA staining dye (Invitrogen # S33102) added according to the manufacturers' instructions (1:10,000 dilution). A low-molecular weight ladder is required for accurate selection of the appropriate sizes; we recommend pBR322 DNA-MspI Digest (New England Biolabs # N3032S) or Low MW Ladder (NEB, #N3233S). Be sure to use large volume combs to allow loading of the entire 50 µl reaction + 10 µl loading dye into a single well.

h. Load samples and run the gel until marker bands in the 100-500 bp size range are well separated. Illuminate the gel very briefly (< 30 seconds total exposure time) on a UV-transilluminator set at low intensity, for just long enough to mark the appropriate region (400-500 bp) with a clean razor blade. Cut only the middle of the lane, leave the edges (see picture).Turn off the UV light and carefully cut out the marked region, transferring it into a microcentrifuge tube (2 ml tube are suggested).

i. Extract the cDNA from this gel slice by adding 35 µl of nuclease-free water to the tube, spinning down briefly to bring water and gel slice into contact, and incubating overnight at 4°C. No further purification procedures are necessary; simply use the water elute in the subsequent steps.

5. qPCR quantification for mixing on the same HiSeq lane

NOTE: Illumina preps other than RNA-seq can be similarly processed using the same primers and scripts; quantify all the preps that are to be mixed on the same lane simultaneously. For checking quality and quantity of gel eluted DNA we do two pcr – one –to check the product size on gel – it should be the same as the band we cut out – no additional products. For mixing samples together in equal proportions we perform Qpcr with P5 and P7 primers and mix samples according to Ct analysis.

a. For quality check prepare a PCR master mix. The following volumes are for a single reaction, so multiply these values by the total number of reactions plus a small additional amount to account for pipetting error.

(volumes given in µl)	
H_2O	6.4
dNTP (2.5 mM ea)	1
10X PCR buffer	1
IC2-P7 primer (10 µM)	0.2
IC1-P5 primer (10 µM)	0.2
Titanium Taq polymerase	0.2
DNA template from step 4i	1
Total volume	10

b. Amplify in a PCR-thermocycler using the following profile:

95°C 5 min, (95°C 40 sec, 63°C 1 min, 72°C 1 min) X 12 cycles

Run 5 (or less) µl on gel. The size of the product should match the size you aiming when cut a band for gel-extraction.

c. For qPCR quantification, for mixing samples in equal proportions into the library prepare two dilutions of each sample (clear elute from step 4i), 1/10 and 1/50 in 10 mM Tris-HCl. Arrange the dilutions in a 96-well plate for easy pipetting. For each sample, make two identical dilution series – these will be your technical replicates. Do not simply split the same dilution series into two – we aim to replicate the process of making those dilutions.

d. Mix SYBR-based qPCR Master mix appropriate for your qPCR instrument with water and two primers, according to the recipe below, and aliquot 14 µl per reaction. Plan for two no-template-control (NTC) reactions.

(Volumes given in µl)	
H_2O	6.1
SYBR Green mix	7.5
IC2-P7 primer (10 µM)	0.2
anti-T30 primer (10 µM)	0.2
Dilution template from step 5c	1
Total volume	15

e. Amplify in a qPCR machine with some NTC (no-template control) using the following profile:

95°C 5 min, (95°C 40 sec, 63°C 1 min, 72°C 1 min) X 20 cycles

(The product should amplify around 8-12 cycles, the NTC controls should be empty).

f. Arrange the data in Excel in the form of a table with four columns: sam (sample name), lane (intended HiSeq lane), conc (DNA dilution; use 0.1 for 1/10, 0.02 for 1/50), and ct (qPCR result for this sam-conc combination). There must be at least two technical replicates for each combination of sam-conc (i.e. two rows with the same sam and conc and different ct values). If all samples are to be mixed on the same lane, enter '1' throughout the lane column. The order of columns and rows does not matter, but the names of the columns do matter (they are case sensitive).

g. Export the data from Excel as comma-separated values (.csv). Open script mix_illumina_qpcr.R in R, follow the instructions given in the comments within the script.

h. Mix samples (elutes from the gel slices, step 3m) according to final mixing table produced by the script. This material is in principle ready for Illumina sequencing, except you might need to concentrate the resulting sample 2-3 fold to meet the sequencing facility requirements. In that case, we recommend mixing a larger volume of the all-barcodes mixture and concentrating it using SpeedVac.

Sequences 5’ - 3’of oligonucleotide primers used in this protocol

3ILL-30TV ACGTGTGCTCTTCCGATCTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTV

5ILL CTA CAC GAC GCT CTT CCG ATC T

S-Ill-swMW ACCCCATGGGGCTACACGACGCTCTTCCGATCTNNMWGGG

(*Note: This custom RNA oligo should be stored in multiple aliquots at -80°C to prevent degradation of this labile and expensive reagent*).

TruSeq-Mpx-2n (similar to the **Illumina_Universal**)
AATGATACGGCGACCACCGAAAAATACACTCTTTCCCTACACGACGCTCTTCCGAT

ILL -BC (BC=*NNNNNN*)
CAAGCAGAAGACGGCATACGAGAT*NNNNNN*GTGACTGGAGTTCAGACGTGTGCTCTTCCGAT

BC(*NNNNNN*)

AATGCT	TCTATA	ACTTGA	TGAAGG
GACACA	TGCAAA	CCGTCC	GGTGTG
GAGTGG	TGGCAC	TAATCG	AGCGAG
CACCGG	TGTTAG	TATAAT	AGCTTT
GAAGTT	AAGGGA	TCATTC	TGGTCT
GCAGGA	GCACCC	TCCCGA	ACCGGC
GTATTA	TACGTG	TCGAAG	AAAAGT
TCACAT	TTAGGC	AAACAC	ACATCT

The barcoding oligos given in the table are a selection of standard Illumina TrueSeq barcodes that have The best “Hamming distance” barcode to complete a line (32 samples*5 millions of reads each = 160 millions of reads); more barcode sequences can be found elsewhere.

If ordering from IDT, order them as “ultrameres” with no purification; this seems to be the best quality-cost balance. Remember that the barcode will be read in a reverse-complement orientation compared to the sequence in the table.

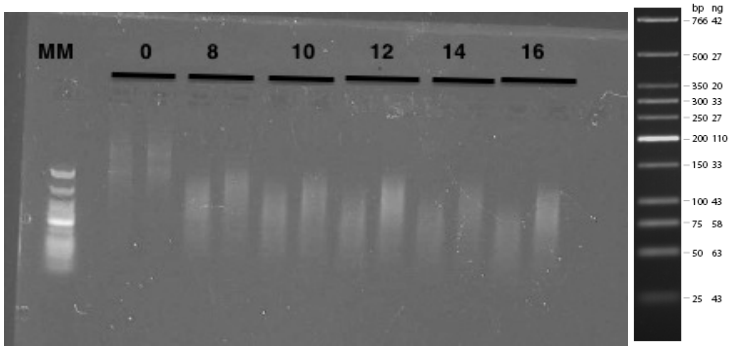
Illmn-Check1-P5 (IC1-P5)(Ill-Libp1) AATGATACGGCGACCACCGA

Illmn-Check2-P7 (IC2-P7)(Ill-Libp2) CAAGCAGAAGACGGCATACGA

anti-T30 AAATTAGATCGGAAGAGCACAC

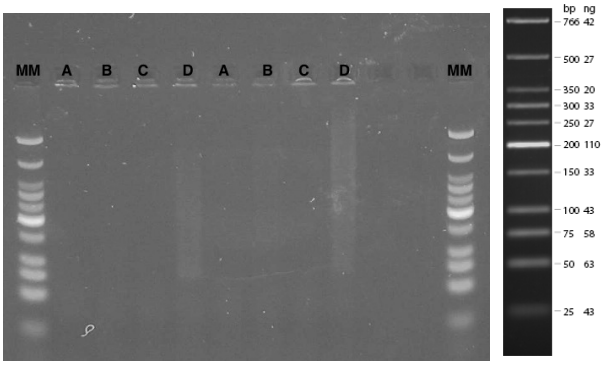
Figures

Figure 1. Degradation time (Step 1b). 0 in wells 1 and 2 mean samples without degradation. 8 mean 8 minutes of degradation, and so far.



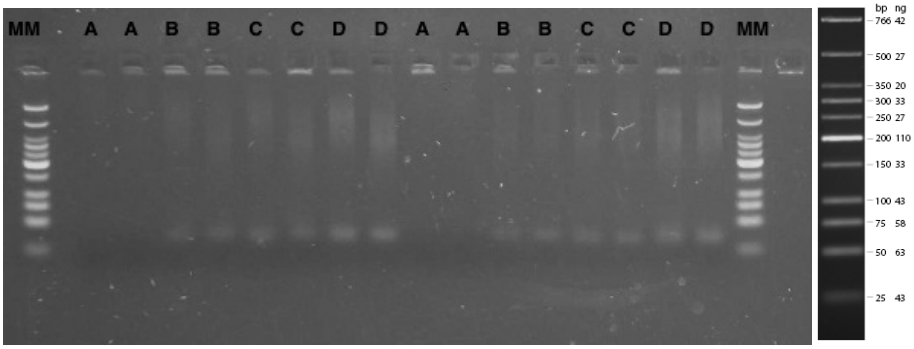
MM: Low molecular weight DNA ladder, New England BioLabs, Catalogue number N3233S. 2 % agarose gel, 100 ng RNA, 20 min 115 V.

Figure 2. cDNA PCR test



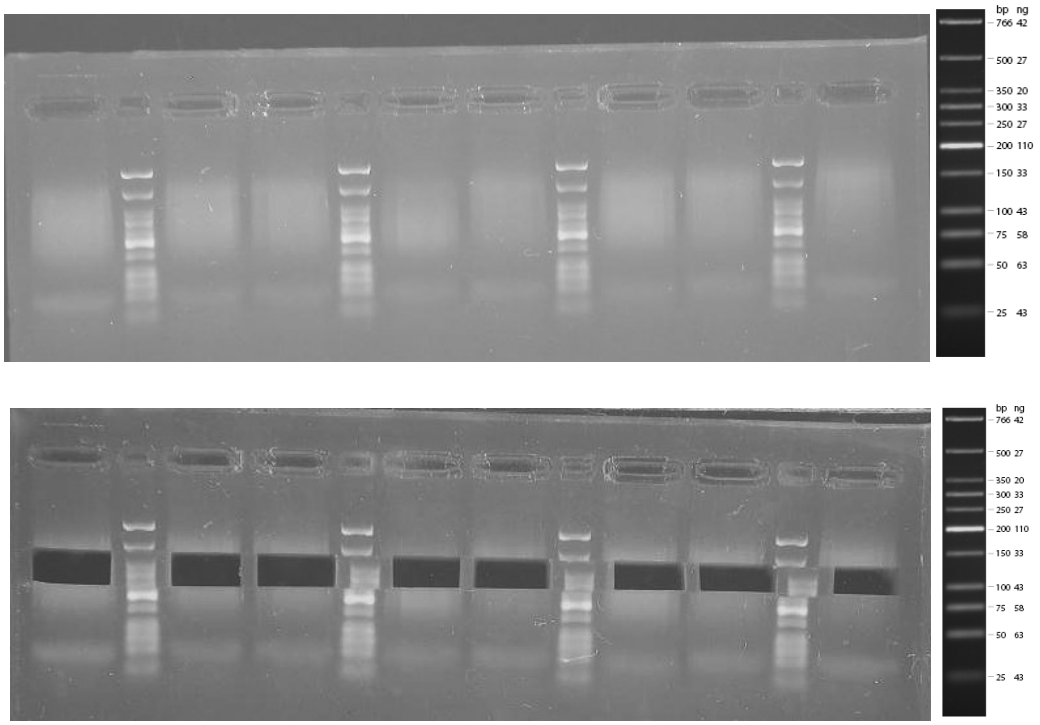
2% agarose, 10 ul PCR product loading, 35 minutes, 100 V. MM: Low molecular weight DNA ladder, New England BioLabs, Catalogue number N3233S. 2 % agarose gel, 100 ng RNA, 20 min 115 V.

Figure 3. Adaptors extension



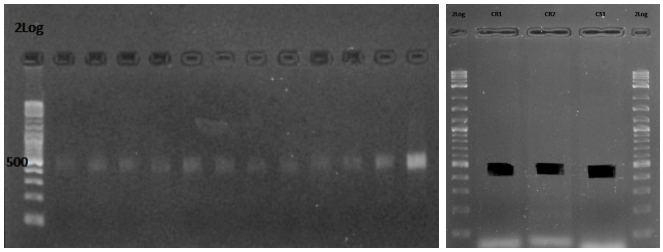
2 % agarose gel, 10 ul product, 30 minutes at 100 V.

Figure 4. Size selection



2 % agarose gel, 40 ul product, 45 minutes at 130 V.

Figure 5: Quality test



Bibliography

Meyer, E., et al. (2011). "Profiling gene expression responses of coral larvae (Acropora millepora) to elevated temperature and settlement inducers using a novel RNA-Seq procedure." Molecular ecology 20: 3599-3616.