Zamore Lab Random-Primed, Strand-Specific RNA-Seq Protocol with UMI-adapters

By ZZ 2012, updated by Xuan 2016; by Ildar 2021 and 2024

Citations:

Zhang et al. (2012) Strand-specific libraries for high throughput RNA sequencing (RNA-Seq) prepared without poly(A) selection. *Silence* **3**:9.

Fu et al. (2018) Elimination of PCR duplicates in RNA-seq and small RNA-seq using unique molecular identifiers. *BMC Genomics* **19**:531.

Data analysis tools: https://github.com/weng-lab/umitools

DNA oligos

Multiplexing Adapters (IDT, all adapters are ordered with hand mixing random Ns):

```
Set 1:
MP-Ada1.1 5'-/5Phos/CCCNNNNNAGATCGGAAGAGCACACGTCT        RNase-free HPLC
MP-Ada1.2 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNGGGT  Std desalting

Set 2:
MP-Ada2.1 5'-/5Phos/GATNNNNNAGATCGGAAGAGCACACGTCT        RNase-free HPLC
MP-Ada2.2 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNATCT  Std desalting

Set 3:
MP-Ada3.1 5'-/5Phos/TGANNNNNAGATCGGAAGAGCACACGTCT        RNase-free HPLC
MP-Ada3.2 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNTCAT  Std desalting
```

Annealing adaptors:

Anneal each pair of oligos in a separate tube: mix MP-AdaX.1 and MP-AdaX.2 in 1× Rapid Ligation Buffer (Enzymatics #L6030-HC- L) for a final volume of 50 µL (f.c of each adapter 10 µM). Heat the mixture to 95°C for 2 min, then slowly cool to 22°C at −0.1°C/sec in a PCR machine. Incubate at 22°C for an additional 5 min. Place on ice. Make a 10 µM equimolar mix of all three sets of annealed oligos.

Multiplexing PCR Primer 1.0 (MP-primer-1):

```
5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'
```

Multiplexing PCR Primer 2.0 (MP-P2-IdX):

```
5'-CAAGCAGAAGACGGCATACGAGATNNNNNNGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'
where NNNNNN is the multiplexing barcode as follows:
Id1  CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Id2  CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Id3  CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Id4  CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Id5  CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Id6  CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Id7  CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Id8  CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Id9  CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Id10 CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Id11 CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Id12 CAAGCAGAAGAGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
```

PROCEDURE

I. Ribosomal RNA depletion

Mouse: We use a pool of 186 rRNA antisense oligos (100 µM/each) that tile across the 18S, 28S, 5S, and 5.8S rRNAs and the 12S and 16S mitochondrial rRNAs. Final working concentration: 0.5 µM for each oligo. For sequences see “rRNA depletion oligos” Excel file.

Fly: We use a pool of 160 rRNA antisense oligos (100 µM/each) that tile across the 18S, 28S, and 5.8S rRNAs and the L and S mitochondrial rRNAs. Final working concentration: 0.5 µM for each oligo. For sequences see “rRNA depletion oligos” Excel file.

- Use 0.5–2 µg high quality total RNA (strongly recommended) as starting material. Suggested kit for RNA extraction: mirVana (Ambion)
- Add 1–2 µL of pooled rRNA antisense oligos for every 1 µg of total RNA
- Add 1 µL of 1/100 diluted ERCC spike-in mix 1 (~1,035 attomole; Thermofisher #4456740) to enable absolute quantification of RNA-seq data. Record the number of cells or the amount of total RNA used to make the library.
- To the mix add 3–12 µL of rRNA oligo hybridization buffer to make the final volume 14 µL.
- Heat the mixture at 95°C for 2 min, then slowly cool at –0.1°C/sec to 22°C in a PCR machine. Incubate at 22°C for an additional 5 minutes. Place on ice.
- Add the following to RNA (final volume: 20 µL):
    - 2 µL 10× RNase H digestion buffer
    - 2 µL water
    - 2 µL Thermostable RNase H (10 units; Hybridase, Lucigen #H39500)
- Incubate at 45°C for 30 min. Place on ice.
- *DNase treatment*: Add the following to each tube (final volume 50 µL):
    - 23 µL water
    - 5 µL Turbo DNase 10× buffer
    - 2 µL Turbo DNase (Thermofisher #AM2238 2U/µL)
- Incubate at 37°C for 20 min.
- Purify RNA using RNA Clean & Concentrator-5 (Zymo Research #R1015), enriching RNA >200nt. This will remove 5S rRNA and tRNAs.
    - Adjust *RNA Binding Buffer*: combine 50 µL RNA binding buffer and 50 µL 100% ethanol
    - Add 100 µL of adjusted *RNA Binding Buffer* to RNA samples. Mix well.
    - Transfer the mixture to the Zymo-Spin column. Centrifuge for 30 sec at 17,000 × *g*. RNA >200 nt will be retained on the column. Discard the flow-through.
    - Add 400 µL *RNA Prep Buffer*. Centrifuge for 30 sec 17,000 × *g*.
    - Add 700 µL *RNA Wash Buffer*. Centrifuge for 30 sec at 17,000 × *g*.
    - Add 400 µL *RNA Wash Buffer*. Centrifuge for 2 min at 17,000 × *g*.
    - Centrifuge for 30 sec at 17,000 × *g* again to remove all the ethanol.
    - Move the column to a new 1.7 mL tube.
    - Add 6 µL water to the column. Incubate @ room temp for 1 min. Centrifuge for 30 sec at 17,000 × *g* to elute RNA.

II. Library preparation

After rRNA depletion continue to RNA fragmentation and priming with NEBNext Ultra II Directional RNA Library Prep Kit for Illumina(E7760) as described in kit protocol Section 4 with the following adaptations:

- Optimize fragmentation time and PCR cycles for the desired insert length and your specific organism, tissue, and/or amount of starting total RNA.
- Use in-house multiplexing UMI adapters and multiplexing PCR primers instead of those that come with the kit.

1) Fragmentation and Priming

Assemble the fragmentation and priming reaction **on ice** in a nuclease-free tube by adding the following components:

FRAGMENTATION AND PRIMING MIX	VOLUME
Purified mRNA or rRNA Depleted RNA	5 µl
• (lilac) NEBNext First Strand Synthesis Reaction Buffer	4 µl
• (lilac) Random Primers	1 µl
Total Volume	**10 µl**

Mix thoroughly by pipetting up and down ten times.

Place the tube in a thermal cycler and incubate as follows:

- Incubate for 12 min at 94°C with the heated lid set to 105°C.
- Hold at 4°C

2) First Strand cDNA Synthesis Reaction

Assemble the first strand synthesis reaction **on ice** by adding the following components to the fragmented and primed RNA:

FIRST STRAND SYNTHESIS REACTION	VOLUME
Fragmented and Primed RNA	10 µl
• (brown) NEBNext Strand Specificity Reagent	8 µl
• (lilac) NEBNext First Strand Synthesis Enzyme Mix	2 µl
Total Volume	**20 µl**

Mix thoroughly by pipetting up and down ten times.

Incubate the sample in a preheated thermal cycler with the heated lid set at ≥ 80°C as follows:
Step 1: 10 minutes at 25°C
Step 2: 15 minutes at 42°C
Step 3: 15 minutes at 70°C
Step 4: Hold at 4°C

3) Second-Strand cDNA Synthesis

Assemble the second-strand cDNA synthesis reaction **on ice** by adding the following components to the first-strand reaction product

SECOND STRAND SYNTHESIS REACTION	VOLUME
First-Strand Synthesis Product	20 µl
• (orange) NEBNext Second Strand Synthesis Reaction Buffer with dUTP Mix	8 µl
• (orange) NEBNext Second Strand Synthesis Enzyme Mix	4 µl
Nuclease-free Water	48 µl
Total Volume	**80 µl**

- Keeping the tube on ice, mix thoroughly by pipetting the reaction up and down ten times.
- Incubate in a thermal cycler for 1 h at 16°C with the heated lid set to ≤ 40°C (or off)
- Purify with 144 µl Ampure beads and elute with 53 µl 0.1× TE buffer

4) End-Prep of cDNA Library

Assemble the end prep reaction **on ice** by adding the following components to the second strand synthesis product.

END-PREP REACTION	VOLUME
Second Strand Synthesis Product	50 µl
• (green) NEBNext Ultra II End Prep Reaction Buffer	7 µl
• (green) NEBNext Ultra II End Prep Enzyme Mix	3 µl
Total Volume	**60 µl**

- For multiple reactions, a master mix of the reaction components can be prepared on ice before addition to the sample. If a master mix is made, add 10 µl of master mix to 50 µl of cDNA for the End-Prep reaction.
- Set a 100 µl or 200 µl pipette to 50 µl and then pipette the entire volume up and down at least 10 times to mix thoroughly. Perform a quick spin to collect all liquid from the sides of the tube.

Incubate the sample in a thermal cycler with the heated lid set at ≥ 75°C as follows:

- 30 min at 20°C
- 30 min at 65°C
- Hold at 4°C

Proceed immediately to adaptor ligation.

5) Adaptor Ligation

- Dilute 1 µl of 10 µM UMI adaptors with 1.5 µl adapter dilution buffer

Assemble the ligation reaction **on ice** by adding the following components, in the order given, to the end-prep reaction product from previous step:

LIGATION REACTION	VOLUME
End-Prepped DNA	60 µl
Diluted Adaptor	2.5 µl
• (red) NEBNext Ligation Enhancer	1 µl
• (red) NEBNext Ultra II Ligation Master Mix	30 µl
Total Volume	**93.5 µl**

- The Ligation Master Mix and Ligation Enhancer can be mixed ahead of time; the mixture is stable for at least 8 h at 4°C. Do not pre-mix the Ligation Master Mix, Ligation Enhancer and adaptor prior to use in the Adaptor Ligation Step.
- Set a 100 µl or 200 µl pipette to 80 µl and then pipette the entire volume up and down at least 10 times to mix thoroughly. Perform a quick spin to collect all liquid from the sides of the tube.
- Incubate 15 min at 20°C in a thermal cycler with the heated lid off.
- Add 3 µl (blue or red) USER Enzyme to the ligation mixture from previous step. The total volume should now be 96.5 µl.
- Mix well and incubate at 37°C for 15 min with the heated lid set to ≥ 45°C, then hold at 4°C.
- Purify with 87 µl Ampure beads and elute with 22 µl 0.1× TE buffer.

6) PCR Enrichment of Adaptor Ligated DNA

COMPONENT	VOLUME PER ONE LIBRARY
Adaptor Ligated DNA	20 µl
• (blue) NEBNext Ultra II Q5 Master Mix	25 µl
10 µM MP-P1 primer	2.5 µl
10 µM MP-P2-IdX primer	2.5 µl
Total Volume	**50 µl**

CYCLE STEP	TEMP	TIME	CYCLES
Initial Denaturation	98°C	30 sec	1
Denaturation	98°C	10 sec	6–13
Annealing/Extension	65°C	75 sec	
Final Extension	65°C	5 min	1
Hold	4°C	∞	

- The number of PCR cycles should be adjusted based on RNA input. 9 -10 cycles are enough when pre-rRNA depletion starting amount is 2 µg RNA.
- Purify with 45 µl Ampure beads and elute with 23 µl 0.1x TE buffer.

7) Purification with non-denaturing PAGE:

- This step removes any unreacted adapters left after bead purification
- Prepare a 10% non-denaturing polyacrylamide gel (20 cm × 16 cm × 1.5 mm, length × width × thickness), with 20 small wells (each well: 17 mm × 5 mm × 1.5 mm, depth × width × thickness). The gel recipe is as follows:

Water	32 ml
5x TBE	5 ml
40% Acrylamide:Bis-acrylamide 29:1 (Sigma-Aldrich # A7802)	12.5 ml
10% ammonium persulfate	500 µl
TEMED	50 µl

- Pre-run the gel in 0.5× TBE buffer at 35 W, constant power, for 30 min (the surface temperature of the aluminum plate on the gel glass will reach 45–55°C).
- To 25 µl of the library from the previous step add 5 µl. of 6× Loading Dye (e.g., Tritrack, Thermofisher #R1161) and mix thoroughly.
- Rinse wells with running buffer. Load 30 µl of libraries into each well, keeping at least one well empty between samples. Remember to load 100 bp DNA Ladder (e.g., NEB #N3231S) in the rightmost and the leftmost wells.
- Run the gel in 0.5× TBE at 20 W, constant power, for 75 min or until the xylene cyanol bands are just above the bottom of the gel.
- Open the glass plates and cut off one bottom corner of the gel to mark the loading sequence. Gently peel the gel off the glass plate and transfer it into the glass tray containing 1× SYBR Gold in 100 ml 0.5× TBE buffer. Stain the gel for 30 sec.
- Cut off margins of an Avery Diamond Clear Standard Weight sheet protector, separate the two pieces, and lay one piece on the bench with the inside facing up. Gently transfer the gel to the sheet protector. Always keep the gel on the sheet protector to avoid contamination.
- Scan the gel with Typhoon FLA 9000 or similar. Print the gel image
- Transfer the gel on the sheet protector and place it on top of the printed image. Cut the gel between the 250 bp and 550 bp with a brand-new razor blade. **Do not use the same blade to cut different samples**. Transfer the gel slice into a 2 ml tube and crush it with either a single-use plastic pestle or by centrifuging it through a pierced 0.5 ml tube.
- Add 0.7 ml of 0.4 M NaCl and 25 mM EDTA pH8.0 to each tube and rotate end-over-end at 4°C (i.e., in a cold room) overnight to elute libraries.
- The following day, use a P1000 tip with the end cut off to transfer most of the gel elution mix from each 2.0 ml tube to a Spin-X column (Sigma #CLS8162) and centrifuge at 10,000 × *g* for 5 min to filter out gel pieces.
- Split eluate from each 1.7 ml tube into two fresh tubes, ~350 µl each, add 1 µl (15 µg) of GlycoBlue (Thermofisher # AM9515) to each tube and vortex briefly to mix.
- Add 3 vol (1,050 µl) of 100% ethanol to each tube and vortex briefly to mix. Incubate on ice for 1 h.
- Centrifuge at 20,000 × *g* for 30 min at 4°C.

- Wash the pellet with 900 µl of 75% ethanol. Centrifuge at 20,000 × *g* for 5 min at 4°C, then gently remove supernatant without touching the pellet.
- Centrifuge again at 20,000 × *g* for 1 min at 4°C and gently remove the residual ethanol with a P10 tip.
- Air dry for 2 min and dissolve the pellet in 5–10 µl for each tube and combine the library from the two tubes.
- Quantify the libraries with Kapa (Kapa Biosystems, Inc. # KK4844) or Qubit (Life Technologies # Q33231) and pool equal moles of each library to sequence on an Illumina NextSeq or Element Biosciences AVITI instrument.

III. Calculating absolute quantity of the transcripts using the ERCC spike-in transcripts:

As described in Gainetdinov et al. (*Nature* **619**:394-402. 2023):

1) Reformat the sequences to extract UMIs:

```
python reformat_umi_fastq.py -l R1.fastq -r R2.fastq -L R1.reformated  -
   R R2.reformated
```

2) In piPipes, align reformatted sequences to rRNA using bowtie2 (v.2.2.0). Then map unaligned reads to genome.

```
piPipes rna -l R1.reformated  -r R2.reformated -g mm10 -o output
```

3) Remove PCR duplicates from the bam file.

```
python2 umi_mark_duplicates.py -f file.bam > deumi.bam
samtools view -@ 8 -b -F 0x400 deumi.bam > dedup.bam
```

4) In parallel, align reformatted reads from Step 1 to an index of ERCC spike-in transcripts (ThermoFisher, 4456740) using bowtie (v.1.0.0). Mark UMIs and remove PCR duplicates as described above.
5) Calculate transcript and ERCC read counts using StringTie or Htseq
6) Normalize transcript read counts to ERCCs to calculate absolute quantity:
   - In the beginning of library prep we added 1.035×10^{-15} moles of ERCC. Multiplying this number by Avogadro's number gives us the number of ERCC molecules in the library. Since we know the ERCC read counts from Step 5, we can now calculate the relationship of read counts to molecules and apply that to transcript counts.

```
Normalization_ratio = total_ercc_count/(1.035e-15 * 6.023e23)
Transcript_quantity= transcript_count/Normalization_ratio
```

   - Additionally, we can express each transcript quantity with respect to total input RNA or total number of cells that were used to make the library, e.g.: 300 molecules of Transcript X per cell.

Recipes (Morlan et al. *PLOS ONE 7(8): e42882 (*2012) & Adiconis et al. *Nat Methods 10, 623–629 (2013)*)

rRNA oligo Hybridization buffer

100 mM Tris-Cl (pH 7.4)
200 mM NaCl

10× RNase H digestion buffer

500 mM Tris-Cl (pH 7.4)
1M NaCl
200 mM $MgCl_2$