

Library Preparation for small RNA sequencing using 4N adapters: In house 4N Protocol C

MD Giraldez (✉ rogerpalexander@gmail.com)

Department of Internal Medicine, University of Michigan, Ann Arbor, MI

M Tewari (✉ mtewari@med.umich.edu)

Departments of Internal Medicine and Biomedical Engineering, Center for Computational Medicine and Bioinformatics, and Biointerfaces Institute, University of Michigan, Ann Arbor, MI

DJ Galas (✉ dgalas@pnri.org)

Pacific Northwest Research Institute, Seattle, WA

TAP Driedonks

Department of Biochemistry and Cell Biology, Faculty of Veterinary Medicine, Utrecht University, The Netherlands

LC Laurent

Department of Reproductive Medicine and Sanford Consortium for Regenerative Medicine, University of California San Diego

PG Woodruff

Cardiovascular Research Institute and the Department of Medicine, Division of Pulmonary, Critical Care, Sleep, and Allergy, University of California San Francisco

JE Freedman

Department of Medicine, Division of Cardiovascular Medicine, University of Massachusetts Medical School, Worcester, MA

K Van Keuren-Jensen

Neurogenomics, The Translational Genomics Research Institute (TGen), Phoenix, AZ

Y Wang

Center for Cancer Computational Biology, Dana-Farber Cancer Institute, Boston, MA

I Ghiran

Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

K Wang

Institute for Systems Biology, Seattle, WA

H Jiang

Center for Computational Medicine and Bioinformatics and Department of Biostatistics, University of Michigan, Ann Arbor, MI

ENM Nolte-t Hoen

Department of Biochemistry and Cell Biology, Faculty of Veterinary Medicine, Utrecht University, The Netherlands

HPJ Buermans

14 Leiden Genome Technology Center, Department of Human Genetics, Leiden University Medical Center, The Netherlands

R Rubio

Center for Cancer Computational Biology, Dana-Farber Cancer Institute, Boston, MA

D Baxter

Institute for Systems Biology, Seattle, WA

J Khoory

Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

S Lu

Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

A Courtright

Neurogenomics, The Translational Genomics Research Institute (TGen), Phoenix, AZ

K Tanriverdi

Department of Medicine, Division of Cardiovascular Medicine, University of Massachusetts Medical School, Worcester, MA

PL De Hoff

Department of Reproductive Medicine and Sanford Consortium for Regenerative Medicine, University of California San Diego

S Srinivasan

Department of Reproductive Medicine and Sanford Consortium for Regenerative Medicine, University of California San Diego

AJ Barczak

Lung Biology Center, Department of Medicine, University of California San Francisco

PM Godoy

Lung Biology Center, Department of Medicine, University of California San Francisco

A Etheridge

Pacific Northwest Research Institute, Seattle, WA

RM Spengler

Department of Internal Medicine, University of Michigan, Ann Arbor, MI

DJ Erle

Lung Biology Center, Department of Medicine, University of California San Francisco

Method Article

Keywords: small RNA sequencing, 4N adapters, small RNA library preparation, extracellular RNA, exRNA, ERCC

Posted Date: May 15th, 2018

DOI: <https://doi.org/10.1038/protex.2018.081>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

This protocol describes a library preparation method for sequencing small RNA. The method uses degenerate adapters to alleviate the problem of bias in the ligation steps during small RNA library preparation, and it optimizes several other parameters to make it appropriate for use specifically with plasma RNA.

Introduction

Extracellular RNAs (exRNAs) have been identified in every biofluid that has been tested. In biofluids, they have been found in extracellular vesicles, ribonucleoprotein complexes, and lipoprotein complexes. ExRNAs are of considerable interest because they can serve as signaling molecules between cells, they have the potential to serve as biomarkers for prediction and diagnosis of disease, and exRNAs or the extracellular particles that carry them might be used for therapeutic purposes. The Extracellular RNA Communication Consortium (ERCC) is a group of laboratories funded by the U.S. National Institutes of Health. One goal of the ERCC is to develop robust and standardized methods for collecting and processing of biofluids, separating different types of exRNA-containing particles, and isolating and analyzing exRNAs. The "Reference Profiles group":<http://exrna.org/referenceprofiles/> within the consortium is tasked with collecting reliable profiles of the spectrum of extracellular RNAs found in healthy individuals. These reference profiles will serve as benchmarks for comparison with the exRNA profiles from patients with disease. The present protocol for small RNA library preparation was developed by the ERCC Reference Profiles group for use with RNA isolated from human plasma. Key elements of this small RNA library preparation method include the use of 4 random nucleotides on the end of the adapters ligated to the small RNAs of interest, the use of higher than usual adapter concentrations, and the use of increased amounts of polyethylene glycol (PEG) in the ligation steps. Small RNA library preparation methods that lack such degenerate adapters have been found to exhibit significant bias in the representation of different RNA sequences.¹⁻⁴ The use of random adapters is designed to alleviate that problem. The use of high adapter concentrations and macromolecular crowding agents such as PEG reduces bias by driving the ligation reactions toward completion. It should be noted that the use of improved ligation conditions and increased adapter concentrations also results in the formation of more adapter dimers, so purification of desired ligation products from such unwanted side products by size fractionation is essential in this protocol. In some cases, two sequential PCR and gel purification steps might be required to remove adapter dimer products efficiently. Specific notes on these issues are included below. **PCR amplification** Small RNA libraries may require 10-20 cycles of PCR amplification depending on the type and amount of input. It has been shown that increased amplification does not significantly affect library bias.^{1,5} Increased number of PCR cycles will, however, increase the amount of adapter dimers that must be separated from the library. Typically, a single PCR and gel purification step is sufficient to remove most of the adapter dimer products from insert-containing PCR products. However, for low input samples, the adapters are present in such great excess over the input RNA that it can be very helpful to perform two consecutive PCR and gel purification steps. The first PCR consists of only a few

cycles to prevent over-amplifying the excess of adapters. The PCR products are separated by gel electrophoresis and the eluate, now enriched for insert-containing PCR products, is then amplified further in a second PCR using universal (not index-specific) PCR primers. A second gel purification is useful for removing residual adapter contaminants. ****Size selection**** Because of the large excess of adapter dimers in low-input small RNA libraries, electrophoretic purification of PCR products is often necessary. This can be done using either acrylamide or agarose gels, as long as the gel can sufficiently resolve the insert-containing fragments (≈150bp) from the adapter dimer fragments (≈125bp). Use of an automated size selection instrument like the Pippin Prep, BluePippin or PippinHT can reduce variability introduced during gel excision. A full list of the protocols developed by the ERCC is available at the "exRNA Portal": <http://exrna.org/resources/protocols/>, the ERCC's website. This protocol is one of four protocols relating to a 2018 Nature Biotechnology paper. Library Preparation for small RNA sequencing using 4N adapters: In house 4N Protocol A "10.1038/protex.2018.049": <http://dx.doi.org/10.1038/protex.2018.049> Library Preparation for small RNA sequencing using 4N adapters: In house 4N Protocol B "10.1038/protex.2018.050": <http://dx.doi.org/10.1038/protex.2018.050> Library Preparation for small RNA sequencing using 4N adapters: In house 4N Protocol C "10.1038/protex.2018.081": <http://dx.doi.org/10.1038/protex.2018.081> Modified TruSeq Small RNA Library Prep using Randomized 4N Adapters: In house 4N Protocol D "10.1038/protex.2018.052": <http://dx.doi.org/10.1038/protex.2018.052>

Reagents

****Oligonucleotide sequences****¹ - 5' adapter (desalted)² - 5'/5InvddT/A**C**A**C**rGrUrUrCrArGrArGrUrUrCrUrArCrArGrUrCrCrGrArCrGrArUrCr(N:25252525)r(N)r(N)r(N) - 3' adapter (HPLC purification) - 5' /5rApp/(N:25252525)(N)(N)(N)TGGAATTCTCGGGTGCCAAGG/3ddC/ - RT primer (desalted) - 5' GCCTTGACCCGAGAATTCCA - RP1 PCR primer (HPLC Purification) - same as Illumina RP1 PCR primer - Indexed PCR primers RPI1-RPI48 (HPLC Purification) - same as Illumina RPI1-RPI48 primers - Universal PCR primer F (desalted) - 5' AATGATACGGCGACCACCGAG - Universal PCR primer R (desalted) - 5' CAAGCAGAAGACGGCATACGA ****Reagents for 3' Ligation**** - T4 RNA ligase 2 truncated KQ (NEB M0373) - 10X T4 RNA ligase reaction buffer (included with T4 RNA ligase) - 50% PEG 8000 (included with T4 RNA ligase) - RNaseOut RNase inhibitor (Invitrogen 10777-019) - Strip tubes (Axygen PCR-0208-CP-C or equivalent) ****Reagents for adapter depletion**** - *E. coli* single-stranded DNA binding protein (SSB) (Promega M3011) - 5' deadenylase (NEB M0331) - RecJf (NEB M0264) ****Reagents for 5' Ligation**** - 10mM ATP (NEB P0756) - T4 RNA ligase 1 (NEB M0204) ****Reagents for Reverse Transcription**** - Strip tubes (Axygen PCR-0208-CP-C or equivalent) - Superscript III (Invitrogen 18080-044) - 5x First strand buffer (included with Superscript III) - 0.1M DTT (included with Superscript III) - 25mM dNTP mix, diluted 1:2 before use (Thermo Scientific R1121) - RNaseOut RNase inhibitor (Invitrogen 10777-019) - RiboShredder RNase blend (Epicentre RS12500) ****Reagents for PCR Amplification #1**** - LongAmp Taq 2X Master Mix (NEB M0278) - Ampure XP beads Beckman Coulter (A63880) ****Reagents for Gel Purification**** - 3% agarose cassettes for PippinHT (Sage Science HTC3010) ****Reagents for PCR Amplification #2**** - KAPA 2X real-

time PCR master mix (KAPA KK2701) or other high fidelity PCR master mix - Ampure XP beads Beckman Coulter (A63880) **Reagents for Library Validation** - DNA 1000 chip (Agilent, 5067-1504) or High sensitivity DNA chip (Agilent, 5067-4626)

Equipment

Equipment for Ligations and Reverse Transcription - Vacuum concentrator - Thermal cycler
Equipment for PCR Amplifications - Thermal cycler or qPCR machine
Equipment for Gel Purifications - PippinHT (Sage Science, HTP0001) size selection instrument - Centrifuge - Vacuum concentrator
Equipment for Library Validation - Agilent 2100 Bioanalyzer

Procedure

This protocol is optimized for plasma libraries starting with ~100 μL of plasma.³ - Resuspend the 5' adapter at 25 μM . - Resuspend the 3' adapter at 10 μM . - Resuspend the RT primer at 10 μM . - Resuspend the RP1 PCR primer and the indexed PCR primers (RPI1-48) at 20 μM . - Combine the universal forward and reverse PCR primers into a primer cocktail containing 20 μM of each primer. - Make dried PEG strip tubes. These create a high PEG concentration in a small volume for the 3' ligation. Batches of dried PEG PCR tubes can be made as follows. Add 3 μL of 50% PEG 8000 (supplied with NEB RNA ligases) into each strip tube and dry in the speedvac using low heat (37°C). This may take 1-2 hours depending on the speedvac. Once dry, the PEG will appear as a white, flaky pellet in the bottom of the tube. Cap the tubes and store desiccated at room temperature. Note that the 50% PEG solution is very viscous, so care must be taken when pipetting. Diluting the 50% PEG to 25% and aliquoting larger volumes may improve pipetting accuracy. Positive displacement pipettes may also be helpful.

****3' Ligation**** To each strip tube, add: $x \mu\text{L}$ RNA 1 μL Adenylated 3' adapter (10 μM stock concentration) 5-x μL Water 6 μL Total The dried PEG in the strip tube will take up about 1 μL in the reaction after resuspension. 1. Heat tube containing RNA and adapter to 70°C for 2 minutes, then snap cool on ice. To each tube containing 7 μL of denatured RNA and adapter, add: 1 μL 10X T4 RNA ligase reaction buffer 1 μL RNaseOut RNase inhibitor 1 μL T4 RNA ligase 2 truncated KQ 3 μL Total 2. Incubate at 25°C for 2 hours.

****Adapter Depletion**** 1. Add 1 μg *E. coli* SSB (diluted in 1X ligase buffer) and incubate at 25°C for 10 minutes. 2. Add 1 μL 5' deadenylase and incubate at 30°C for 1 hour. 3. Add 1 μL of RecJf and incubate at 37°C for 1 hour.

****5' ligation**** In a separate tube, add 1 μL of 5' adapter (25 μM stock concentration) per ligation and denature at 70°C for 2 minutes, then snap cool on ice. To the tube of denatured 5' adapter, add (per ligation): 1 μL 10mM ATP 1 μL T4 RNA ligase 1. Add 3 μL of adapter, ATP and ligase mix to the completed 3' ligation for a total volume of 16 μL . 2. Incubate at 25°C for 1 hour.

****Reverse transcription (RT)**** To a new tube, add: 1 μL RT primer (10 μM stock concentration) 6 μL Ligated RNA (Store leftover ligated RNA at -80°C.) 7 μL Total 1. Heat to 70°C for 2 min, then snap cool on ice. Add: 2 μL 5x First strand buffer 0.5 μL 12.5 mM dNTP mix 1 μL DTT 1 μL RNaseOut 1 μL Superscript III 5.5 μL Total Add to 7 μL denatured RNA/primer for 12.5 μL total RT reaction. 2. Incubate at 55°C for 1 hour, then at 70°C for 15 minutes. 3. Add 1 μL RiboShredder RNase blend and

incubate at 37°C for 15 minutes.⁴ **PCR amplification #1**⁵ 25 µL PCR master mix (LongAmp Taq 2x Master Mix) 2 µL Illumina RP1 2 µL Illumina RPI1-48 reverse index primer 8 µL water 13 µL cDNA 50 µL Total 1. Amplify for 4 cycles using the following program: 98°C 30 sec. 1X
 98°C 10 sec. 60°C 30 sec. 4X 65°C 35 sec. 65°C 2 min. 1X 2. To purify PCR product, add 100 µL of resuspended room temperature AMPure XP beads to each 50 µL PCR reaction and mix well. Allow the tube to stand at room temperature for 5 minutes and then place on magnetic stand to collect beads (approximately 2 minutes). Remove supernatant and wash beads with 150 µL of 80% ethanol. Repeat wash one time. Remove ethanol and allow beads to dry for 5 minutes. Bound material should be eluted from beads using 23 µL of TE buffer or H₂O. Usually 21 µL can be recovered. Try to minimize bead carry-over. **Gel purification #1**⁶ 1. Add 5 µL of marker 30A for PippinHT to each sample. Mix well. Use 127-156bp for size selection. At the end of the run, collect eluate in elution well. 2. Concentrate eluate in speedvac to 22.5 µL. **PCR amplification #2**⁵ 25 µL KAPA 2X real-time master mix⁷ 2.5 µL Universal primer cocktail (mix containing 20 µM of each primer) 22.5 µL cDNA 50 µL Total 1. Amplify for 11-16 cycles using the following program: 98°C 45 sec. 1X
 98°C 15 sec. 60°C 30 sec. 11-16X 72°C 20 sec. 72°C 2 min. 1X To purify PCR product, add 100 µL of resuspended room temperature AMPure XP beads to each 50 µL PCR reaction and mix well. Allow the tube to stand at room temperature for 5 minutes and then place on magnetic stand to collect beads (approximately 2 minutes). Remove supernatant and wash beads with 150 µL of 80% ethanol. Repeat wash one time. Remove ethanol and allow beads to dry for 5 minutes. Bound material should be eluted from beads using 23 µL of TE buffer or H₂O. Usually 21 µL can be recovered. Try to minimize bead carry-over. **Gel purification #2** 1. Add 5 µL of marker 30A for PippinHT to each sample. Mix well. Use 127-156bp for size selection. At the end of the run, collect eluate in elution well. 2. Concentrate eluate to ~10 µL and run 1 µL on a Bioanalyzer DNA 1000 or dilute 1:10 and run on a High Sensitivity DNA chip to check library size and purity. A band at 149bp corresponds to inserts of 23bp. Adapter dimers are 126bp.

Troubleshooting

1) Oligonucleotide sequences © 2016 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. 2) The 5' end of this adapter has an inverted dideoxy-T followed by 4 DNA bases with phosphorothioate bonds (indicated by asterisks) to prevent nuclease degradation. 3) This protocol is optimized for plasma libraries starting with ~100 µL of plasma. Libraries with more starting material can probably tolerate higher adapter concentration and higher PEG concentration. Libraries with less starting material may need reduced adapter or PEG concentration. **Caution:** Differential expression analysis, i.e. comparison of RNA levels across samples, should not be done between data sets from libraries prepared using different PEG or adaptor concentrations. Relative read numbers are affected by these parameters. 4) Incubating the reverse transcription reaction with RiboShredder RNase blend might not be necessary. 5) A single PCR of 15-20 cycles and gel purification may be sufficient for most plasma samples. If adapter ligation products are still a problem, the PCR and size selection can be split into 2

phases, as shown here. First, do 4 PCR cycles followed by a gel purification. Then use the recovered DNA for a second PCR of 11-16 cycles followed by a second gel purification. 6) If an automated size selection instrument like a BluePippin or PippinHT is not available, size select a band of 149bp with any suitable gel purification system. 7) Use of a real-time master mix is not necessary, but can be helpful for determining the number of amplification cycles for each sample. Some qPCR machines (for example the Bio-Rad CFX series) can be paused and opened to remove individual samples at different cycle numbers, which is helpful when amplifying different sample types at the same time.

References

1) Jayaprakash AD, Jabado O, Brown BD, Sachidanandam R. 2011. Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. *Nucleic Acids Research* **39**: e141. 2) Sorefan K, Pais H, Hall AE, Kozomara A, Griffiths-Jones S, Moulton V, Dalmay T. 2012. Reducing ligation bias of small RNAs in libraries for next generation sequencing. *Silence* **3**: 4. 3) Zhang Z, Lee JE, Riemondy K, Anderson EM, Yi R. 2013. High-efficiency RNA cloning enables accurate quantification of miRNA expression by deep sequencing. *Genome Biology* **14**: R109. 4) Zhuang F, Fuchs RT, Sun Z, Zheng Y, Robb GB. 2012. Structural bias in T4 RNA ligase-mediated 3'-adapter ligation. *Nucleic Acids Research* **40**: e54. 5) Hafner M, Renwick N, Brown M, et al. 2011. RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. *RNA* **17**: 1697-1712.