Full-length poly(A) and mRNA sequencing (FLAM-seq)

Ivano Legnini

Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrück Center for Molecular Medicine (MDC)

Jonathan Alles

Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrück Center for Molecular Medicine (MDC)

Nikos Karaiskos

Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrück Center for Molecular Medicine (MDC)

Salah Ayoub

Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrück Center for Molecular Medicine (MDC)

Claudia Quedenau

Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrück Center for Molecular Medicine (MDC)

Nikolaus Rajewsky (✉ rajewsky@mdc-berlin.de)

Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrück Center for Molecular Medicine (MDC)

Method Article

Keywords: RNA, long reads, full-length mRNA sequencing, poly(A) tails

Posted Date: August 5th, 2019

DOI: <https://doi.org/10.21203/rs.2.10045/v1>

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

[Read Full License](#)

Abstract

We developed FLAM-seq, a fast and simple method for generating cDNA libraries of full-length mRNAs, including the poly(A) tail. By combining a new strategy for cDNA preparation with PacBio long-read sequencing, FLAM-seq enables to generate hundreds of thousands of reads per sample in an easy and short procedure, with starting material ranging from 500 ng to 10 µg of total RNA. Besides the quantification of gene expression and the information about the mRNA isoform, the data generated with FLAM-seq allow accurate measurement of poly(A) tail length for thousands of genes in the sequenced samples. Here we describe the protocol step by step from the RNA preparation to the first steps of data analysis.

Reagents

Truseq mRNA preparation kit (cat. RS-122-2102, Illumina)

USB poly(A) length assay kit (cat. 764551KT, Thermo Fisher)

RNAClean XP Beads (cat. A63987, Beckman Coulter)

SMARTScribe Reverse Transcriptase kit (cat. 639537, Clontech)

XP DNA beads (cat. A63881, Beckman Coulter)

Advantage 2 DNA polymerase mix (cat 639201, Clontech)

RNA 6000 Pico Kit (cat. 5067-1513, Agilent Technologies, optional but recommended)

High Sensitivity NGS Fragment Analysis Kit (cat. DNF-474, Advanced Analytical Technologies GmbH)

Agilent DNA 12000 Kit (cat 5067-1508, Agilent Technologies)

Buffer EB (cat 19086, Qiagen)

Sequel™ SMRT® Cell 1M v3 Tray (101-531-000, Pacific Biosciences)

SMRTbell™ Template Prep Kit 1.0-SPv3 (100-991-900, Pacific Biosciences)

Sequel Sequencing Kit 3.0 (101-597-800, Pacific Biosciences)

Sequel Binding and Internal Ctrl Kit 3.0 (101-626-600, Pacific Biosciences)

Oligonucleotides

RT primer 1:

GGTAATACGACTCACTATAGCGAGANNNNNNNNNNNCCCCCCCCTTT

RT primer 2 (to use in alternative to 1):

TGAGTCGGCAGAGAACTGGCGAANNNNNNNNNNCCCCCCCCTTT,

Template switch oligo (where “i” indicates stereoisomers of dC and dG as in the associated publication):

iCiGiCAAGCAGTGGTATCAACGCAGAGTACATrGrGrG

PCR primer 1 (to use in combination with RT primer 1):

GGTAATACGACTCACTATAGCGAG

PCR primer 2 (to use in combination with RT primer 2):

TGAGTCGGCAGAGAACTGGCGAA

Equipment

Agilent Fragment Analyzer

Agilent 2100 Bioanalyzer

Magnetic rack for 96 well

Qubit Fluorometer

PacBio Sequel system

Ordinary lab equipment (vortex, thermocycler, benchtop centrifuge ...)

Procedure

Poly(A)⁺ RNA preparation (using Illumina Truseq mRNA preparation kit).

1. Dilute 2-10 µg of total RNA in 50 µl of RNase-free H₂O.
2. Vortex RNA Purification Beads and add 50 µl to RNA sample. Pipet up and down 6 times to mix.
3. Incubate in thermocycler at 65° C for 5 min., on ice for 5 min. and at room temperature for 5 min.
4. Place sample in magnetic separator for 5 min. Remove and discard all the supernatant.

5. Remove sample from rack. Add 200 µl of Bead Washing Buffer and pipet up and down 6 times to mix sample.
6. Place the sample back in the magnetic separator for 5 min. Remove and discard all the supernatant.
7. Add 50 µl of Elution Buffer and pipet up and down 6 times.
8. Incubate in thermocycler at 80° C for 2 min.
9. Remove sample from thermocycler and keep at room temperature for 5 min.
10. Add 50 µl of Bead Binding Buffer and pipet up and down 6 times. Incubate at room temperature for 5 min.
11. Place sample in magnetic separator for 5 min. Remove and discard all supernatant.
12. Remove sample from rack. Add 200 µl of Bead Washing Buffer and pipet up and down 6 times.
13. Place sample in magnetic separator for 5 min. Remove and discard all supernatant.
14. Add 16 µl of RNase-free H₂O and pipet up and down 6 times.
15. Incubate in thermocycler at 70° C for 2 min., then on ice for 1 min.
16. Place sample in magnetic rack for 5 min. Transfer 16 µl of the supernatant to a new 0.2 ml PCR tube. Keep 1 µl for Bioanalyzer (optional).

GI tailing of RNA (using USB poly(A) tail length assay, Thermo Fisher).

1. Prepare the following mix, keeping the reagents on ice.

poly(A)⁺ RNA: 14 µl

5X tail buffer mix: 4 µl

10X tail enzyme mix: 2 µl

Tot. vol.: 20 µl

Incubate at 37°C for 60 min.

2. Add 1.5 µl of tail stop solution, keep on ice for 2 min.
3. Add 1.8X XP RNA beads, incubate at room temperature for 5 min.

4. Put on magnetic rack, keep for 3 min. Remove supernatant.
5. Wash twice with 50 µl Ethanol 80% for 30 sec.
6. Carefully remove supernatant, leave tubes open for 10 min in order to dry.
7. Resuspend beads in 17 µl H₂O, put on magnetic rack for 5 min. and collect supernatant. Keep 1 µl for Bioanalyzer (optional).

cDNA synthesis (using SMARTscribe reverse transcriptase kit, Clontech).

1. Prepare the following mix:

5X First strand buffer: 8 µl

DTT 20mM: 1.5 µl

dNTP mix 10 mM: 4 µl

RNase Inhibitor: 2 µl

isoTSO 12 µM: 2 µl

SMARTScribe RT: 2 µl

H₂O: 2.5 µl

Tot. vol.: 22 µl

Keep the master mix at room temperature.

2. In a second PCR tube, mix 16 µl of purified GI-tailed RNA and 2 µl of RT primer (10µM).
3. Mix contents and spin briefly. Incubate in thermocycler starting the following program: 72°C for 3 min., 42°C for 60 min., 70 °C for 10 min. and at 4°C hold.
4. When the RNA/primer sample has done completed the first step at 72°C, wait for 1-2 min. until it reaches the temperature of 42°C and add the RT master mix, previously equilibrated at room temperature, reaching a final volume 40µl.
5. When the reaction is finished, add 0.6X XP DNA beads, incubate at room temperature for 5 min.
6. Put on magnetic rack, keep for 3 min. Remove supernatant.
7. Wash twice with 50 µl Ethanol 80% for 30 sec.

8. Carefully remove supernatant, leave tubes open for 10 min.

9. Resuspend beads in 42 µl of H₂O, put on magnetic rack for 5 min. and collect supernatant. Keep 1 µl for Bioanalyzer (optional).

cDNA library amplification (using Advantage 2 PCR enzyme system, Clontech).

1. Prepare the following mix:

10X Advantage 2SA PCR buffer: 10 µl

Diluted first-strand cDNA: 40 µl

dNTP Mix (10 mM each): 2 µl

5' PCR Primer II A (10 µM): 2 µl

Univ. RV Primer (10 µM): 2 µl

H₂O: 42 µl

50X Advantage 2 Polymerase Mix: 2 µl

Tot. vol.: 100 µl

Keep the master mix on ice.

2. Start the thermocycler with the following program: 98°C for 1 min., N. cycles x [98°C for 10 sec., 63°C for 15 sec., 68°C for 3 min.], 68°C for 7 min. Put the tube into the thermocycler when it has reached the temperature of 98°C. Note: if the sample has never been processed before, do PCR optimization splitting the 100 µl reaction into 4 tubes and do 18, 20, 22, 24 cycles, then compare the amplified libraries in terms of size profile and yield.

3. When the reaction is finished, add 0.6X XP DNA beads, incubate at room temperature for 5 min.

4. Put on magnetic rack, keep for 3 min. Remove supernatant.

5. Wash twice with 200 µl Ethanol 80% for 30 sec.

6. Carefully remove supernatant, leave tubes open for 10 min.

7. Resuspend beads in 42 µl of H₂O, put on magnetic rack for 5 min. and collect supernatant. Keep 1 µl for Fragment Analyzer. Quantify the library with a Qubit.

Sequencing library preparation (using the SMRTbell™ Template Prep Kit)

1. Prepare the following mix, keeping the reagents on ice.

ds DNA: up to 37 µl

DNA Damage Repair Buffer: 5µl

NAD+: 0.5 µl

ATP high: 0.5 µl

dNTP: 0,5 µl

DNA Damage Repair Mix: 2 µl

water: ___ µl to adjust

Tot. vol.: 50µl

2. Incubate at 37°C for 60 min.

3. Add 2.5 µl End Repair Mix, incubate at 25°C for 5 min.

4. Add 0.6X AMPure beads, incubate at room temperature for 10 min in a vortex mixer at 2000 rpm.

5. Spin down (1 sec.), put on magnetic rack to collect the beads and pipette off cleared supernatant.

6. Wash twice with 500 µL Ethanol 80% for 30 sec.

7. Spin down (1 sec.), carefully remove supernatant, leave tubes open for 10 sec in order to dry.

8. Resuspend beads in 30 µl EB, incubate at room temperature for 5 min. in a vortex mixer at 2000 rpm.

9. Spin down (1 sec.), put on magnetic rack for 1 min. and collect cleared supernatant (optional: keep 1 µl for Bioanalyzer).

10. For adapter ligation, prepare the following mix, keeping the reagents on ice.

DNA (End Repaired): 29 µl

Annealed Blunt Adapter: 2 µl

→ mix before processing

Template Prep Buffer: 4 µl

ATP low: 2 μ l

→ mix before processing

Ligase: 1 μ l

Water: 2 μ l

Tot. vol.: 40 μ l

11. Incubate at 25°C for 15 min., 65°C for 10 min.

12. For exonuclease treatment, add to 40 μ l of ligated DNA:

Exo III: 0.5 μ l

Exo VII: 0.5 μ l

13. Incubate at 37°C for 60 min.

14. Add 0.6X AMPure beads, incubate at room temperature for 10 min in a vortex mixer at 2000 rpm.

15. Spin down (1 sec), put on magnetic rack to collect the beads and pipette off cleared supernatant.

16. Wash twice with 500 μ l Ethanol 80% for 30 sec.

17. Spin down (1 sec), carefully remove supernatant, leave tubes open for 10 sec in order to dry.

18. Resuspend beads in 30 μ l EB, incubate at room temperature for 5 min in a vortex mixer at 2000 rpm.

19. Spin down (1 sec), put on magnetic rack for 1 min. and collect cleared supernatant.

20. Repeat steps 14.-17.

21. Resuspend beads in 11 μ l EB, incubate at room temperature for 5 min. in a vortex mixer at 2000 rpm.

22. Spin down (1 sec.), put on magnetic rack for 1 min. and collect cleared supernatant.

23. Check library quality on Bioanalyzer and Qubit.

Sample setup and sequencing

1. Use the Sample Setup module in SMRT Link to generate a customized pipetting protocol for primer conditioning, primer annealing and polymerase binding.

2. prepare the following mix:

Sequencing Primer v3: 1 μ l

1x Elution buffer: 29.1 μ l

3. Incubate at 80°C for 2 min., hold at 4°C.

Annealing Primer:

4. Mix (according to the customized pipetting protocol) the following:

sample

10x Primer buffer v2

conditioned Sequencing primer v3

water

5. Incubate at 20°C for 60 min, put on ice for immediate use or store at 20°C for long-term use.

6. Mix (according to the customized pipetting protocol) the following:

sample

Sequel dNTP

DTT

Sequel binding buffer

diluted Sequel Polymerase 3.0

water

7. Incubate at 30°C for 60 min., use bound complex right away or store at 4°C for up to 7 days.

8. Dilute the complex with Magbead binding buffer v2 (according to the customized pipetting protocol) and add AMPure beads; incubate at room temperature for 5 min.

9. Spin down (1 sec), put on magnetic rack to collect the beads and pipette off cleared supernatant. Important: do not wash with ethanol.

10. Resuspend beads in Magbead binding buffer v2, incubate at room temperature for 15 min.

11. Spin down (1 sec), put on magnetic rack for 1 min. and collect cleared supernatant.

12. Measure recovered DNA with Qubit and enter concentration and volume into the Sample Setup Module.

13. Mix the following (85 µl total volume per SMRT cell)

prepared sample

DTT

Diluted internal control

Sequel Additive

Sequel Complex Dilution buffer

14. Load 85 µl and store up to 24 h at 4°C.

15. Use the Run Design module in SMRTLink to create a run.

16. Refer to the table for loading and pre-extension recommendations for your insert size:

<https://www.pacb.com/wp-content/uploads/Quick-Reference-Card-Loading-and-Pre-Extension-Recommendations-for-the-Sequel-System.pdf>

17. Set movie recording time to 360 - 600 min.; loading method: diffusion mode.

Data collection and analysis

Use the SMRT Analysis module in SMRTLink to create CCS fastq files (Minimum Number of Passes = 3, default parameters).

Once the raw data have been processed, the final CCS fastq files can be used directly as input for the FLAM-seq analysis pipeline, available at:

<https://github.com/rajewsky-lab/FLAMAnalysis>.

Troubleshooting

Low-input samples

We applied minor changes to the protocol for limiting samples (0.5-2 µg). Briefly, the RNA after poly(A) selection is eluted in 10 µl, GI tailing is performed in 12.5 µl and RNA eluted in 10 µl, cDNA synthesis is performed in 20 µl and eluted in 20 µl, PCR is performed in 50 µl for 23-25 cycles. Moreover, we found

that fixing cells with methanol before RNA extraction improves RNA quality and yield for samples of neuronal origin.

Poly(A) versus total RNA

FLAM-seq can be performed with either poly(A)-selected RNA or with ribo-depleted RNA. IN an exploratory analysis, we observed that RNAs with very short tails can be depleted with the poly(A) selection, however they represent a minor population in most biological samples (Figure 1). The advantage of poly(A) selection resides in the fact that PacBio sequencing has low depth compared to other sequencing technologies and poly(A) selection limits the fraction of reads mapping to non-mRNA molecules. Using poly(A)-selected RNA can therefore slightly bias the tail length measurement but allows obtaining enough reads to measure tails for thousands of mRNAs. Since tail length per transcript typically have a highly dispersed distribution, what is the required sequencing depth really depends on the biological questions that one wants to address. For poly(A) selection, we used the Illumina Truseq mRNA preparation kit, however any equivalent oligo-dT-based method is suitable in principle.

GI tailing of RNA

We were able to replace the Thermo Fisher USB poly(A) length assay kit with yeast poly(A) polymerase (Thermo Fisher, 74225Z25KU) fed with equimolar G/I ribonucleotides, which is less expensive. However, we found that the buffer provided with this enzyme is very viscous and apparently incompatible with bead-based purification and downstream reactions. If the user decides for this alternative, we recommend to perform phenol-chloroform extraction followed by precipitation before proceeding with cDNA synthesis.

Library profile

After PCR amplification, we recommend to check the library profile on an Agilent Fragment Analyzer. Fragment size of FLAM-seq libraries usually peaks between 1.5 and 2.5 kb, depending on the sample quality and the 5' coverage of the amplified cDNAs (Figure 2a). Transcript coverage can be biased towards the 3' ends, especially for mRNAs larger than ~ 1.5 kb, as reported in the associated publication. We found that increasing RT incubation to 90 minutes and PCR elongation to 5-7 minutes improves library size, but can yield more PCR artifacts depending on the RNA sample quality and amount. When optimizing FLAM-seq, we found that libraries often contained a significant fraction of PCR artifacts, generated by multiple rounds of concatenation of the template switch oligo with the RT primer and resulting in a ladder-like profile of the final library (an extreme example in Figure 2b). To avoid this issue, we introduced a size-selection and purification step after RT and PCR, as well as an anchored dC primer

and a modified template switch oligo (with isonucleotides on its 5' end) for the reverse transcription. If the final library still contains low-size concatemers, we recommend to increase the RNA input and introduce a stricter size-selection for long fragments after cDNA preparation.

Data analysis

The analysis pipeline that we provide discards many reads from the original sequencing output: if the library had a good size profile, still roughly 50% of the reads are typically discarded as they do not contain a poly(A) tail or one of the adapter sequences or they do not uniquely map to the genome. The PacBio Sequel system can provide up to one million reads per SMRT cell, with productively sequenced molecules typically in the range of 200,000 per SMRT cell. In our experience, 2 or 3 SMRT cells per sample are usually required for the basic analyses provided in our associated publication.

Time Taken

Anticipated Results

References

Acknowledgements

Figures

HeLa S3

Figure 1

Figure 1. Poly(A)-selected versus ribo-depleted RNA FLAM-seq libraries. Histograms of poly(A) tail length for HeLa S3 total, ribodepleted RNA (purple) versus poly(A)-selected RNA (blue).

a **b** **Figure 2**

Figure 2. Library profile from Agilent Fragment Analyzer. a. Library profile of a FLAM-seq library obtained from HeLa S3 poly(A)-selected RNA. b. Library profile from a sample of poor quality, where most of the amplified product consists of size-defined concatemers derived from annealing of the RT primer to the template switching oligo and numerous subsequent template switching events. Lower (LM) and upper (UM) markers from the Fragment Analysis kit are labeled in red.