MAPCap: A fast and quantitative transcription start site profiling protocol

Vivek Bhardwaj
Max Planck Institute of Immunobiology and Epigenetics https://orcid.org/0000-0002-5570-9338
Giuseppe Semplicio
Max Planck Institute of Immunobiology and Epigenetics
Niyazi Umut Erdogdu
Max Planck Institute of Immunobiology and Epigenetics
Asifa Akhtar (akhtar@ie-freiburg.mpg.de)
Max Planck Institute of Immunobiology and Epigenetics

Keywords: CAGE, Gene expression analysis, Transcription start sites, High-throughput sequencing

Posted Date: July 30th, 2019

DOI: https://doi.org/10.21203/rs.2.9396/v1

Abstract

Below we present a simple and quick TSS quantification protocol, **MAPCap** (**M**ultiplexed **A**ffinity **P**urification of **Cap**ped RNA) that enables users to combine high-resolution detection of transcription start-sites and differential expression analysis. MAPCap can be used to profile TSS from dozens of samples in a multiplexed way, in 16-18 hours. MAPCap data can be analyzed using our easy-to-use software **icetea** (https://bioconductor.org/packages/icetea), which allows users to detect robust TSS using replicates, and perform differential TSS analysis.

Introduction

The MAPCap protocol described below is discussed in detail in our manuscript: **MAPCap allows high-resolution detection and differential expression analysis of transcription start sites** (Bhardwaj and Semplicio et al., Nature Communications. 2019).

Please find the full protocol, along with the sequence of oligos, PCR primers and ERCC controls as the attached file (pdf).

Reagents

Reagents

Unless indicated otherwise, all buffers were prepared as a stock solution with RNase-free water in freshly autoclaved bottles. Buffers are decanted into Falcon tubes for immediate use and only pipetted out of them. Unless indicated otherwise, all buffers are kept at 4ºC and Falcons with aliquots of buffers are kept on ice.

1) **10x Depletion Buffer**, 200 mM Tris-HCl (pH 7.4), 400 mM KCl.

2) **TE Buffer**, 10 mM Tris-HCl (pH 8), 1 mM EDTA.

3) **2x IPP Buffer**, 100 mM Tris-HCl (pH 7.4), 300 mM NaCl, 0.2% NP-40.

4) **2x PNK-MES Buffer**, 50 mM MES (pH 6.0), 100 mM NaCl, 20 mM MgCl2, 0.2% Tween-20. Store at room temperature.

5) **2x ProtK Buffer**, 200 mM Tris-HCl (pH 7.4), 100 mM NaCl, 20 mM EDTA, 0.2% Tween-20.

6) anti-m7G antibody (SYSY, 201011)

7) Dynabeads ProteinG slurry magnetic beads (Invitrogen, 10003D)

8) Quick-RNA Miniprep Kit (Zymo Research, R1054)

9) Direct-zol RNA Microprep Kit (Zymo Research, R2061)

10) RNA Clean & Concentrator Kit (RCC, Zymo Research, R1013)

11) Oligo Clean & Concentrator Kit (OCC, Zymo Research, D4060)

12) RNase-free water (Invitrogen, AM9906)

13) Ampure XP beads (Beckman Coulter, A63880)

14) SuperScript III First-Strand Synthesis System (Invitrogen, 18080051)

15) ERCC RNA Spike-In Mix (Invitrogen, 4456740)

16) T7-FlashScribe Transcription Kit (CellScript, C-ASF3507)

17) MEGAclear Transcription Clean-Up Kit (Invitrogen, AM1908)

18) RNA 5’ Polyphosphatase (Epicentre (ordered through Lucigen), RP8092H)

19) RNase H (Invitrogen, 18021014)

20) RNasin Plus RNase inhibitor (Promega, N261A)

21) Terminator 5’-Phosphate-Dependent Exonuclease (Epicentre (ordered through Lucigen), TER51020)

22) Phusion High-Fidelity DNA Polymerase (NEB, M0530S)

23) Vaccinia Capping System (NEB, M2080S)

24) rSAP (NEB, M0371L)

25) Proteinase K (Invitrogen, 100005393)

26) CircLigase II ssDNA Ligase (Epicentre (ordered through Lucigen), CL9021K)

27) NEBNext High-Fidelity 2x PCR Master Mix (NEB, M0541S)

28) NEBNext Multiplex Oligos for Illumina (Primer Set 1-4)

Equipment

* Eppendorf pipettes: 1ul to 1ml

* Eppendorf tubes

* Falcon tubes: 15 and 50ml

* Ice bucket with ice

* Vortex mixer

* Tube rotator

* Magnetic rack

* ThermoMixer

* Thermocycler

* Covaris E220 Focused Ultrasonicator

* Covaris micro snap tubes

* Mini centrifuge

* Eppendorf Bench-top centrifuge

* Qubit RNA/DNA HS

* Fragment analyser/Bioanalyser

* Real Time PCR instrument & equipment

* Illumina HiSeq/MiSeq/NextSeq

Procedure

Preparation

Generation of capped ERCC spikes

Ten spike sequences were chosen from the ERCC RNA Spike-In Mix and PCR primers were designed to produce ~500 bp long DNA fragments (Table 1). At the 5’ end, we inserted a T7 class II promoter 2.5 which creates more homogenous 5’ end transcription promoter sequences. The general PCR forward

primer was designed to anneal to the T7 transcription site while the individual reverse primers anneal to the respective ERCC spike sequence.

1) Reverse transcribe 1 µg of ERCC RNA Spike-In Mix with random hexamers using SSIII according to manufacturer's instruction.

2) Top-up to 50 µl volume with RNase-free water.

3) Perform PCR with ~10 ng of reverse transcribed ERCC spike mix as template.

4) Clean-up PCR reaction using OCC according to manufacturer's instruction. Elute in 10 µl of RNase-free water.

5) Perform in vitro transcription with 500 ng of template DNA using the T7 FashScribe Transcription Kit according to manufacturer's instruction.

6) Clean up using MEGAclear Transcription Clean-up Kit according to manufacturer's instruction.

7) Perform 5' end capping using the Vaccinia Capping System according to manufacturer's instruction.

8) Clean-up using OCC. Elute in 17.5 µl RNase-free water.

9) Measure RNA either by Qubit or Nanodrop.

10) Prepare the Polyphosphatase master mix as follows:

- 10x Reaction Buffer 2 µl
- RNasin+ 0.5 µl
- RNA 5' Polyphosphatase (20 U/µl) 1µl
- RNA in water 16.5 µl (not more then 10 µg)

11) Incubate at 37ºC for 30 min.

12) Clean-up using OCC. Elute in 16.5 µl RNase-free water.

13) Prepare the Terminator exonuclease master-mix as follows:

- 10x Reaction Buffer 2 µl
- RNasin+ 0.5 µl
- Terminator Exonuclease (1 U/µl) 1µl

● RNA in water 16.5 µl

14) Incubate at 30ºC for 60 min.

15) Clean-up using OCC. Elute in 10 µl.

16) Measure the concentration of each spike with Qubit.

17) Create a master-mix of all 10 spikes as a 2-fold serial dilution

● Example: take from one spike 64 ng, from the second one, 32 ng…and from the last spike, 0.125 ng. This results in a spike master mix with a concentration of 2.5 ng/ul.

PROCEDURE

RNA Isolation

RNA can be isolated from any source of interest. We use either the Quick-RNA Microprep Kit for cells in suspension or the Direct-zol Microprep Kit for tissues and whole larvae according to manufacturer's instruction.

CRITICAL: *To avoid any contamination with DNA, treat the RNA samples with DNAse. Elute in RNase-free water.*

RNA Preparation

Preparing the beads for IP

1. Wash 25 µl of Dynabeads ProteinG slurry magnetic beads per sample with IPP buffer and resuspend in at least 300 µl IPP buffer. (e.g. For 12 samples we used 300 µl of beads, which after washing were resuspended in 1 ml of IPP buffer.)

2. For each IP, we use 2.5 µg of anti-m7G antibody.

3. Add the antibody to the beads in IPP and incubate at 4ºC rotating for at least 1 hour.

4. Wash the beads three times with IPP buffer, resuspend in IPP buffer and make aliquots of 150 µl per sample

Preparing total RNA for IP

Capped RNAs within the cell mainly consist of snRNAs and snoRNAs. To avoid losing sequencing power and increase detection sensitivity, we have therefore developed a targeted depletion approach using DNA antisense oligos followed by RNase H treatment. These oligos were designed to bind within the first 40 nt of small abundant capped RNAs. The most abundant sn- and snoRNAs based on a first run without depletion were chosen as targets (Table 2). The thereby resulting 20-40 nt long capped RNA species are then removed through column purification using the RNA Clean and Concentrator Kit.

rRNA species do not contain a cap and are therefore excluded from the IP. However, due to their abundance, it is possible that a substantial portion of rRNAs can stick to the beads. To reduce the number of reads coming from rRNAs, we implemented a similar approach for rRNA removal as for the abundant RNAs (Table 3). In addition (also possible as a stand alone solution) we treat the RNA sample with Terminator Exonuclease, which degrades RNA species with a single phosphate group at their 5’ end.

Optional: polyA RNA isolation

If the desired RNA species contains a polyA tail, we recommend a polyA RNA isolation prior to performing MAPCap. This is the most effective way to remove unwanted capped and/or abundant RNA species (sn-, sno- and rRNAs). Starting from polyA RNA, you can directly continue from the fragmentation step of the procedure.

Optional: Adding capped RNA spike-ins

Add 0.05% of capped RNA spike master mix to each sample. E.g. for an RNA starting amount of 5 µg, add 2.5 ng of capped RNA spike master mix.

Abundant RNA removal

Starting volume: 26 µl

1. Prepare a master mix containing 4 µl of 10x Terminator buffer (from the Terminator Kit), 4 µl AS-oligo mix against sn/snoRNAs (6 µM) and 4 µl AS-oligo mix against rRNAs (4 µM).

2. Add 12 µl of AS-master mix to each sample, incubate the RNA at 70ºC for 2 min followed by gradual cooling to 37ºC (for this, we changed the temperature of the thermocycler from 70ºC to 37ºC. This takes around 10 min.)

3. Once the samples have reached 37ºC, add 1 µl of RNase H (2 U/µl) and incubate at 37ºC for 30 min at 1100 rpm interval shaking (30 s on, 120 s off).

4. Incubate the samples at 70 ºC for 5 min, put on ice immediately for at least 1 min, and add 2 µl of 5' Terminator Exonuclease (1 U/µl). Incubate for 60 min at 30 ºC at 1100 rpm interval shaking (30 s on, 120 s off).

5. Clean-up using the RCC Kit according to manufacturer's instruction with slight modifications:

a. Prepare an adjusted RNA Binding Buffer as follows: 100 µl Binding Buffer + 50 µl EtOH (>95%).

b. Elute in 100 µl TE buffer.

6. Transfer the sample to Micro Snap Cap Tubes (Covaris) for fragmentation.

7. We used the Covaris E220 Focused Ultrasonicator with the following settings: Duty cycle: 10%, Intensity: 5, Power: 175 W, Cycles/Burst: 200, time: 180s).

CRITICAL: *It is important to obtain RNA of a uniform size of around 200-400 nt in order to achieve ideal conditions for sequencing. We recommend testing the fragmentation in advance with several conditions (Figure 1).*

IP

1. Transfer the sonicated RNA in 100 µl TE buffer to the 150 µl aliquots of antibody coupled beads.

2. Resuspend the mix and incubate at 4ºC for 1-2 hours.

General remarks regarding washing beads:

● To properly remove all liquid from the beads for washing and resuspension in a new solution, we perform a short spin to remove any liquid that has accumulated in the lid of the tube.

● We then place the tube on a magnet and wait until the beads are properly bound before removing the solution with a vacuum pump or manual pipetting.

● After the last removal of the washing buffer, we take the tubes from the magnet and perform another quick spin on a table centrifuge before placing the tubes back on the magnet. This allows us to collect the remaining liquid at the bottom of the tube with a 10 µl pipet.

General remarks regarding resuspension of beads in a new reaction mix:

● Most reactions are carried out in 20 µl volume. It is therefore important to remove all excess volume from the wash step before continuing with the protocol (see general remarks on washing, point 3).

● We elute beads by dispensing the 20 µl from the wall of the tube where the beads were fixed to the side. Repeat this step until all beads are off the wall and gently resuspend until the beads are homogeneously in solution.

Library preparation

1. Wash the beads 3 times with IPP buffer. Add at least 500 µl of IPP buffer and rotate the beads at 4ºC for 5 min.

2. Prepare PNK-MES master mix as follows:

● 2x PNK-MES Buffer 10 µl

● b-Mercaptoethanol (0.1 M) 1 µl

● T4 PNK (10 U/µl) 1 µl

● RNasin+ (40 U/µl) 0.5 µl

● RNase-free water 7.5 µl

3. Resuspend the beads in 20 µl of PNK-MES master mix.

4. Incubate at 37ºC for 30 minutes in a thermocycler with 1100 rpm interval shaking (30 s on / 120 s off).

5. Wash the beads 3x with cold IPP buffer.

6. Prepare Ligation master mix as follows:

● 10x T4 Ligase Buffer 2 µl

● ATP (1 mM) 2 µl

● RNase-free water 7.5 µl

● Add 6 µl of PEG-8000 to this solution and mix by flicking the tube and pipetting until the PEG is properly dissolved.

● T4 RNA Ligase 1 (10 U/µl) 1 µl

● RNasin+ (40 U/µl) 0.5 µl

7. Resuspend the beads in 19 µl of Ligation master mix.

8. Add 1 µl of s-oligo (10 µM) to each sample, mix the beads by flicking followed by a quick spin.

NOTE: *Use individual s-oligos for each sample. Afterwards, the samples will be pooled for sequencing and computational demultiplexing will be performed based on the s-oligo sample barcode.*

CRITICAL: *Adjust the s-oligo concentration added to the sample based on the starting amount of RNA used. For 5 ug we use 1 µl of a 10 µM solution.*

9. Incubate at 25ºC for 1 hour in a thermocycler with 1100 rpm interval shaking (30 s on / 120 s off).

10. Wash the beads 3x with cold IPP buffer. After 2 washes, resuspend one sample with 900 µl cold IPP buffer and transfer to every subsequent sample to combine all samples. Transfer to a new Eppendorf tube.

11. Prepare the rSAP master mix as follows:

● IPP buffer 35 µl

● RNasin+ (40 U/µl) 1 µl

● rSAP (1 U/µl) 4 µl

12. Resuspend the beads in 40 µl of rSAP master mix and incubate at 37ºC for 30 min with 1100 rpm interval shaking (30s on / 120 s off).

13. Wash beads 3x with IPP buffer.

14. Prepare the ProtK master mix as follows:

● 2x ProtK Buffer 50 µl

● 10% SDS 2 µl

● Proteinase K (20 mg/ml) 10 µl

● RNase-free water 38 µl

15. Resuspend the beads in 100 µl ProtK master mix and incubate at 37ºC for 20 min with 1100 rpm interval shaking (30 s on / 120 s off).

16. Place the tube on the magnet and transfer the supernatant to a new tube. The beads can be discarded at this point.

17. Clean up using OCC according to manufacturer's instruction.

18. Elute in 9 µl of RNase-free water.

19. Prepare Reverse Transcription master mix using the SuperScript III Kit as follows:

● 10x Buffer 2 µl

● dNTP (10 mM) 1 µl

● MgCl2 (25 mM) 4 µl

● DTT (0.1 M) 2 µl

● RNaseOUT (40 U/µl) 1 µl

● SSIII (200 U/µl) 1 µl

20. Add the 11 µl Reverse Transcription master mix to the sample and incubate for 10 min at 42ºC, 10 min at 50 ºC, 10 min at 55 ºC and 10 min at 65 ºC with 1100 rpm interval shaking (30 s on / 120 s off). Increase the temperature in the thermocycler after 10 minutes while keeping the sample in the cycler.

21. Purify the sample using OCC according to manufacturer's instruction.

22. Elute the sample with 6.5 µl RNase-free water.

23. Prepare the Circularization master mix as follows:

● 10x Buffer 1 µl

● MnCl2 (50 mM) 0.5 µl

● Betaine (5 M) 2 µl

● CircLigaseII (100 U/µl) 0.5 µl

24. Add the 4 µl of Circulaization master mix and incubate for 1-16 hours at 60 ºC. (We usually perform this step overnight in an incubator to avoid condensation on the wall and lid.)

Library amplification

In order to perform as few PCR amplification cycles as possible, we recommend performing a qPCR on the circularized DNA template and compare the cycle numbers to a standard library. The standard library is a previously sequenced MAPCap library where the concentration is known. We aim at a library of around 2-4 ng/µl in 10 µl final volume after PCR and Ampure beads clean-up. We use 1 µl of circularized cDNA which is serially diluted 3 times to get different measurements with different concentrations. Typically, MAPCap requires between 14 and 18 cycles of PCR depending on the starting amount used.

1. Set up a PCR reaction:

● 2x NEBNext PCR master mix 20 µl

● Universal primer P5 (10 µM) 1 µl

● Index primer P3 (10 µM) 1 µl

● Circularized cDNA library 9 µl

● RNase-free water 9 µl

2. PCR program: 1. 98 ºC for 30 sec, 2. 98 ºC for 10 sec, 3. 65 ºC for 30 sec, 4. 72 ºC for 20 sec, 5. Repeat 2-4 for desired amount of cycles, 6. 72 ºC for 5 min, 7. 4 ºC forever.

3. Clean-up the amplified library with two rounds of Ampure beads at a 1x ratio.

● Equilibrate Ampure XP beads at room temperature for 30 min and vortex the beads well before use.

● Add 40 µl (1x ratio) of Ampure beads to the library and mix well.

● Incubate 5 min at room temperature before placing the tube on a magnet.

● Once beads are fixed to the magnet and the solution is clear, remove 75 µl of the liquid. (Caution: Do not disturb the beads.)

● Wash 2x with 200 µl 80% EtOH without removing the beads from the magnet.

● Dry the beads at room temperature while on the magnetic stand for 5 min (do not over-dry).

● Remove the tube from the magnet and resuspend the beads in 20 µl of RNase-free water.

● Reclaim beads on a magnet and collect the sample in a new tube.

● Add 20 µl (1x ratio) of Ampure beads to the library and mix well.

● Incubate 5 min at room temperature before placing the tube on a magnet.

● Once beads are fixed to the magnet and the solution is clear, remove 35 µl of the liquid. (Caution: Do not disturb the beads.)

● Wash 2x with 200 µl 80% EtOH without removing the beads from the magnet.

● Dry the beads at room temperature while on the magnetic stand for 5 min (do not over-dry).

● Remove the tube from the magnet. Resuspend the beads in 13 µl of RNase-free water. Reclaim beads on a magnet and collect the sample in a new tube.

4. Proceed with quantification of the DNA concentration (Qubit DNA HS) and size distribution control using Bioanalyzer/Fragment Analyzer. An example of a final MAPCap library is shown in Figure 2. Libraries should be clean of adapter dimers (peak around 140 bp).

5. Final libraries can be stored at -20ºC.

6. Sequence the library following Illumina guidelines (50-75 bp, paired-end, 6 bp index reads).

Troubleshooting

Time Taken

Total: 16-18 hours (calculated for 12 samples)

● Coupling antibody to the beads: ***1-2 hours***

● Abundant RNA depletion: ***2.5 hours***

● RNA fragmentation: ***1 hour***

- IP: ***1-2 hours***
- 3’ end repair (PNK): ***30 minutes***
- Ligation: ***1 hour***
- S-oligo dephosphorylation (rSAP): ***30 minutes***
- Release RNA from beads (ProtK): ***20 minutes***
- Reverse transcription: ***1.5 hours***
- Circularization: ***2 hours (or over night)***
- qPCR: ***2 hours***
- PCR: ***1 hour***
- Final clean-up: ***30 minutes***

- Coupling antibody to the beads: ***1-2 hours***
- Abundant RNA depletion: ***2.5 hours***
- RNA fragmentation: ***1 hour***
- IP: ***1-2 hours***
- 3’ end repair (PNK): ***30 minutes***
- Ligation: ***1 hour***
- S-oligo dephosphorylation (rSAP): ***30 minutes***
- Release RNA from beads (ProtK): ***20 minutes***
- Reverse transcription: ***1.5 hours***
- Circularization: ***2 hours (or over night)***
- qPCR: ***2 hours***
- PCR: ***1 hour***
- Final clean-up: ***30 minutes***

Anticipated Results

1. Insert size of final sequencing library: should be in the range of 200-500 bp. See the "good library" figure (Fig. 2)

2. Number of mapped and de-duplicated reads: Ideally, a good library should have high mapping efficiency (>80% uniquely mapped) and low number of PCR duplicates. The proportion of PCR duplicates varies depending on the number of PCR cycles (which depends on the amount of starting material), but >90% PCR duplicates (compared to uniquely mapped reads) should be concerning as it may lead to low number of kept reads for TSS detection.

3. Reads within TSS: Generally, >60% of duplicate-free reads should be expected to fall within the detected TSSs. Although this number can be increased by modifying the TSS detection parameters (see the icetea vignette), this should be kept in mind that there's a trade-off between sensitivity and specificity during TSS detection.

References

Kodzius, R. et al. CAGE: cap analysis of gene expression. Nat. Methods 3, 211–222 (2006).

Aktaş, T. et al. DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. Nature 544, 115–119 (2017).

Acknowledgements

The authors acknowledge the Deep-Sequencing Unit at MPI-IE for data production. Ibrahim Illik for sharing the s-oligo and the FLASH protocol, and Plamen Georgeiv for help with MLE mutant and wild-type flies.

Figures

RNA concentration [ng/μl]	50	50	50	50	50	50	100	200
fragmentation time [s]	0	30	60	90	120	180	180	180

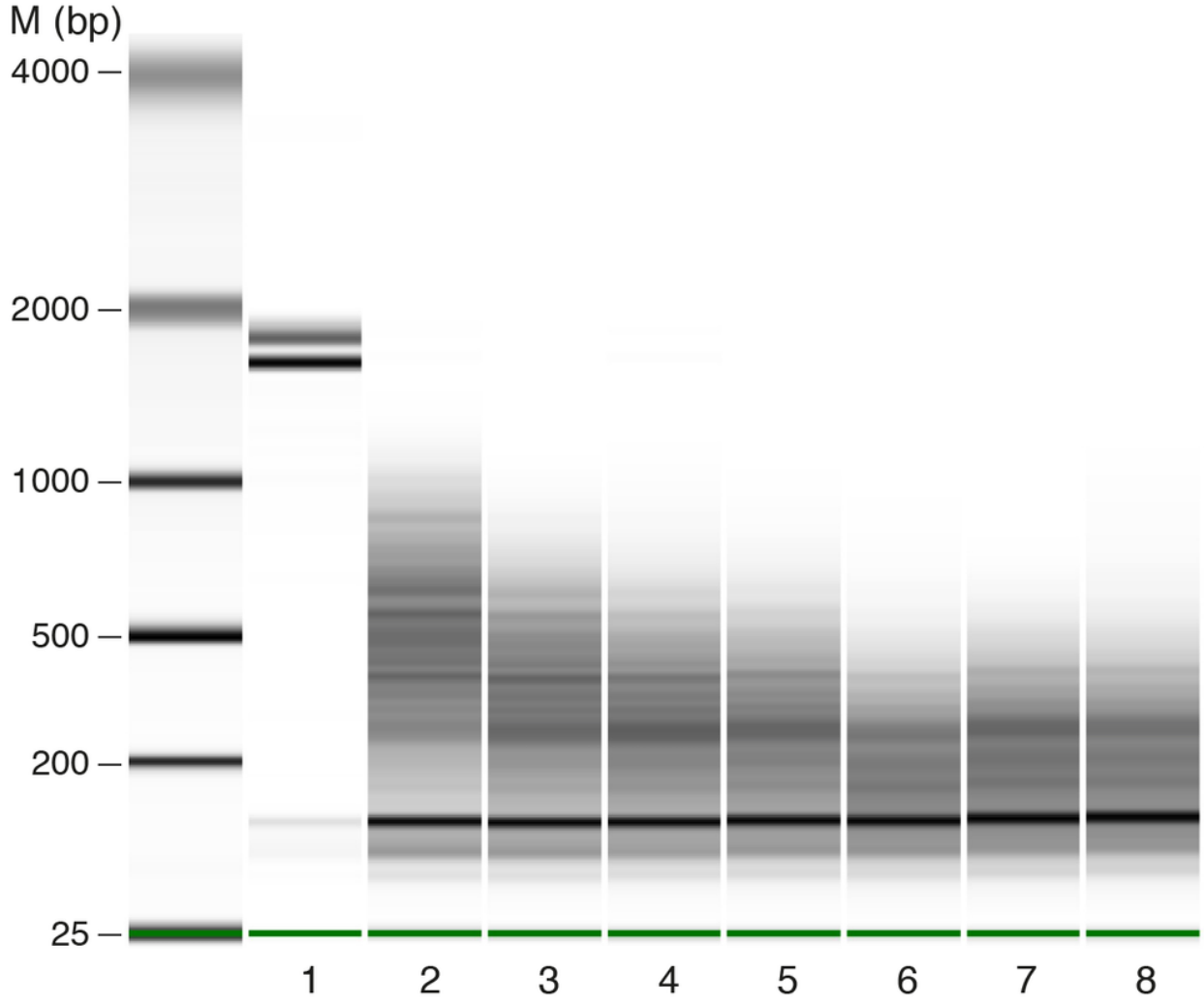

Figure 1

Fig. 1: Various fragmentation times and RNA concentrations have been tested to determine the ideal conditions to receive RNA fragments of 200-400 nt in length.

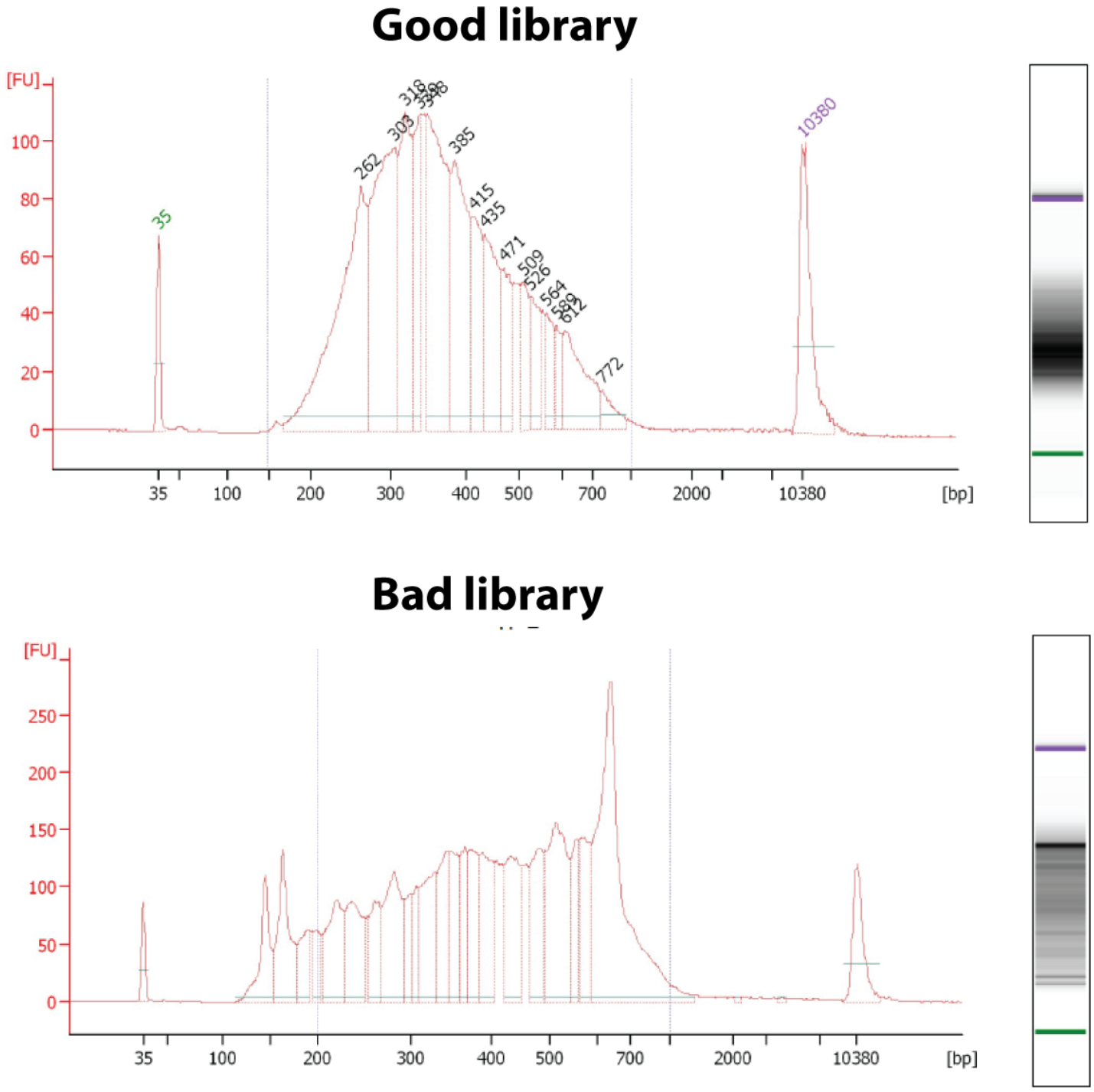

Figure 2

Fig. 2: Example of a good and bad MAPCap library. A good library has the insert size within 200-600 bp and no abundant RNAs.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplement1.xlsx
- supplement2.pdf