Journal of Visualized Experiments

Transcription start site mapping using Super-Low Input Carrier-CAGE

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE59805R1
Full Title:	Transcription start site mapping using Super-Low Input Carrier-CAGE
Keywords:	Cap Analysis of Gene Expression; low-input; carrier; transcription start sites; transcription initiation; degradable carrier; homing endonuclease; promoters; enhancers
Corresponding Author:	Nevena Cvetesic Imperial College London London, UNITED KINGDOM
Corresponding Author's Institution:	Imperial College London
Corresponding Author E-Mail:	nevena.cvetesic@lms.mrc.ac.uk
Order of Authors:	Nevena Cvetesic
	Elena Pahita
	Boris Lenhard
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Open Access (US$4,200)
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	London, United Kingdom

TITLE:

Transcription Start Site Mapping Using Super-Low Input Carrier-CAGE

AUTHORS & AFFILIATIONS:

Nevena Cvetesic[1, 2], Elena Pahita[1,2] and Boris Lenhard[1,2,3]

[1]Institute of Clinical Sciences, Faculty of Medicine, Imperial College London, London, United Kingdom

[2]MRC London Institute of Medical Sciences, London, United Kingdom

[3]Sars International Centre for Marine Molecular Biology, University of Bergen, Bergen, Norway

Corresponding Authors:

Nevena Cvetesic (ncvetesi@ic.ac.uk)

Boris Lenhard (b.lenhard@imperial.ac.uk)

Email Address of Co-author:

Elena Pahita (elena.pahita16@lms.mrc.ac.uk)

KEYWORDS:

Cap Analysis of Gene Expression, low-input, carrier, transcription start sites, transcription initiation, degradable carrier, homing endonuclease, promoters, enhancers

SUMMARY:

Cap Analysis of Gene Expression (CAGE) is a method for genome-wide quantitative mapping of mRNA 5'ends to capture RNA polymerase II transcription start sites at a single-nucleotide resolution. This work describes a low-input (SLIC-CAGE) protocol for generation of high-quality libraries using nanogram-amounts of total RNA.

ABSTRACT:

Cap analysis of gene expression (CAGE) is a method used for single-nucleotide resolution detection of RNA polymerase II transcription start sites (TSSs). Accurate detection of TSSs enhances identification and discovery of core promoters. In addition, active enhancers can be detected through signatures of bidirectional transcription initiation. Described here is a protocol for performing super-low input carrier-CAGE (SLIC-CAGE). This SLIC adaptation of the CAGE protocol minimizes RNA losses by artificially increasing the RNA amount through use of an in vitro transcribed RNA carrier mix that is added to the sample of interest, thus enabling library preparation from nanogram-amounts of total RNA (i.e., thousands of cells). The carrier mimics the expected DNA library fragment length distribution, thereby eliminating biases that could be caused by the abundance of a homogenous carrier. In the last stages of the protocol, the carrier is removed through degradation with homing endonucleases and the target library is amplified. The target sample library is protected from degradation, as the homing endonuclease recognition sites are long (between 18 and 27 bp), making the probability of their existence in the eukaryotic genomes very low. The end result is a DNA library ready for next-generation sequencing. All steps in the protocol, up to sequencing, can be completed within 6

days. The carrier preparation requires a full working day; however, it can be prepared in large quantities and kept frozen at -80 °C. Once sequenced, the reads can be processed to obtain genome-wide single-nucleotide resolution TSSs. TSSs can be used for core promoter or enhancer discovery, providing insight into gene regulation. Once aggregated to promoters, the data can also be used for 5'-centric expression profiling.

INTRODUCTION:

Cap analysis of gene expression (CAGE) is a method used for single-nucleotide resolution genome-wide mapping of RNA polymerase II transcription start sites (TSSs)[1]. Its quantitative nature also allows 5'-end centric expression profiling. Regions surrounding the TSSs (about 40 bp upstream and downstream) are core promoters and represent the physical location where RNA polymerase II and general transcription factors bind (reviewed previously[2,3]). Information on exact locations of TSSs can be used for core promoter discovery and for monitoring promoter dynamics. In addition, as active enhancers exhibit signatures of bidirectional transcription, CAGE data can also be used for enhancer discovery and monitoring of enhancer dynamics[4]. CAGE methodology has recently increased in popularity due to its broad application and use in high-profile research projects such as ENCODE[5], modENCODE[6], and FANTOM projects[7]. In addition, TSS information is also proving to be important for distinguishing healthy and diseased tissue, as disease-specific TSSs can be used for diagnostic purposes[8].

Even though several methods for TSS mapping are available (CAGE, RAMPAGE, STRT, nanoCAGE, nanoCAGE-XL, oligo-capping), we and others have recently shown that CAGE is the most unbiased method to capture true TSSs with the least number of false positives[9,10]. The recent CAGE protocol, nAnT-iCAGE[11], is the most unbiased protocol for TSS profiling, as it avoids cutting the fragments to short tags using restriction enzymes and does not use PCR amplification. A limitation of the nAnT-iCAGE protocol is the requirement for a large amount of starting material (e.g., 5 micrograms of total RNA for each sample). To answer specific, biologically relevant questions, it is often impossible to obtain such high amounts of starting material (e.g., for FACS-sorted cells or early embryonic stages). Finally, if nAnT-iCAGE is successful, only 1–2 ng of DNA library material is available from each sample, thereby limiting the achievable sequencing depth.

To enable TSS profiling using only nanograms of total RNA, we have recently developed Super-low Input Carrier-CAGE[10] (SLIC-CAGE, **Figure 1**). SLIC-CAGE requires only 10 ng of total RNA to obtain high complexity libraries. Our protocol relies on the carefully designed synthetic RNA carrier added to the RNA of interest to achieve a total of 5 µg of RNA material. The synthetic carrier mimics the target DNA library in length distribution to avoid potential biases that could be caused by homogenous molecules in excess. The sequence of the carrier is based on the sequence of the *Escherichia coli* leucyl-tRNA synthetase gene (**Table 1**) for two reasons. First, any leftover of the carrier in the final library, even if sequenced, will not map to a eukaryotic genome. Second, as *E. coli* is a mesophilic species, its housekeeping genes are optimised for the temperature range appropriate for SLIC-CAGE. The carrier sequence is also embedded with homing endonuclease recognition sites to allow specific degradation of DNA derived from the

carrier RNA molecules. The target, sample-derived library remains intact, as the homing endonuclease recognition sites are long (I-CeuI = 27 bp; I-SceI = 18 bp) and statistically unlikely to be found in eukaryotic genomes. After specific degradation of the carrier and removal of fragments by size exclusion, the target library is PCR amplified and ready for next-generation sequencing. Depending on the starting RNA amount (1–100 ng), between 13–18 PCR amplification cycles are expected to be required. The final amount of DNA per each sample ranges between 5–50 ng, yielding enough material for very deep sequencing. When using only 1–2 ng of total RNA, true TSSs can be detected; however, the libraries are expected to be of lower complexity. Lastly, as SLIC-CAGE is based on the nAnT-iCAGE protocol[11], it enables multiplexing of up to eight samples prior to sequencing.

PROTOCOL:

1. Preparation of the carrier

1.1. Preparation of DNA templates for in vitro transcription

1.1.1. Prepare the PCR mix for each PCR template by combining 41 µL of water, 20 µL of 5x HF buffer, 8 µL of 2.5 mM dNTPs, 10 µL of 10 µM unique forward primer (PCR_GN5_f1, **Table 2**; primers are dissolved and diluted in water), 10 µL of 2 ng/µL template plasmid containing the synthetic carrier gene and 1 µL Phusion polymerase. Mix the PCR mix by pipetting. A master mix for all 10 templates can be prepared at once (prepare for 11 reactions).

1.1.2. Add 90 µL of the PCR mix to 10 µL of each 10 µM reverse primer (PCR_N6_r1-r10, **Table 2**). Mix by pipetting.

1.1.3. PCR amplify the templates using the following program: 98 °C for 60 s, (98 °C for 10 s, 50 °C for 30 s, 72 °C for 30 s) 35 cycles, 72 °C for 10 min, hold at 4 °C.

1.1.4. **Gel purification of PCR-amplified DNA templates**

1.1.4.1. Prepare a 1% agarose gel (low-melting agarose is recommended).

1.1.4.2. To decrease the volume, concentrate the PCR reaction mixtures from 100 µL to 20 µL total volume using the vacuum concentrator at a low-medium temperature (30–40 °C).

1.1.4.3. Add 6 µL of the 6x loading dye, mix well, and load on the gel. Run electrophoresis for 30 min in 1x TAE buffer at the voltage appropriate for the used electrophoresis tank (5–10 V/cm). In parallel run a 100 bp or 1000 bp DNA ladder.

1.1.4.4. Using a clean scalpel, excise the gel slices containing the target PCR product. Avoid excess agarose gel. Purify the PCR products using a gel extraction kit (according to the manufacturer's instructions).

NOTE: A260/A230 ratios of DNA isolated from agarose gels are typically low (0.1–0.3). Expected target product and side products are shown in **Figure 2A**. Expected yields from 100 µL PCR reactions are 1.2–3 µg. Reactions can be scaled up to get a higher yield.

1.2. **In vitro transcription of carrier molecules**

1.2.1. Transcribe carrier RNA in vitro using the T7 RNA polymerase according to the manufacturer's instructions. Set up 10–20 µL reactions (the recommended kit is in **Table of Materials**).

1.2.2. Purify the in vitro transcribed RNA using an RNA purification kit. Set up DNA digestion in solution using DNase I following the manufacturer's standard instructions, and elute the RNA in 50 µL of water. To increase the elution yield, leave the water in the column for 5 min before centrifugation.

NOTE: Be careful not to exceed the maximum binding capacity of the columns (in the kit mentioned in the **Table of Materials**, capacity is up to 100 µg). The expected yield from PCR templates 1–10 (1 kbp to 200 bp in length) is 25–50 µg from 10 µL in vitro transcription reactions. Reactions can be scaled up to get a larger stock of carrier molecules.

1.3. **Capping of in vitro transcribed carrier RNA molecules**

1.3.1. Prepare the capping mix by combining 2 µL of 10x capping buffer, 1 µL of 10 mM GTP, 1 µL of 2 mM SAM (freshly diluted) ,and 1 µL of Vaccinia capping enzyme per carrier RNA.

1.3.2. Mix up to 10 µg of each carrier molecule in 15 µL of total volume and denature for 10 min at 65 °C. Place on ice immediately to prevent secondary structure formation.

1.3.3. Mix the denatured carrier RNA with 5 µL of the capping mix and incubate for 1 h at 37 °C.

1.3.4. Purify capped RNA molecules using an RNA purification kit – follow the manufacturer's clean-up protocol. Elute RNA in 30 µL of water. To increase the elution yield, leave the water in the column for 5 min before centrifugation.

NOTE: Measure concentration using the microvolume spectrophotometer. Expected A260/A280 ratio is >2 and A260/A230 is >2. Note that for some RNA samples A260/A230 may be between 1.3–2. Expected yield when using 10 µg of uncapped RNA is 9–10 µg of capped RNA.

1.4. Prepare the mix of the capped and uncapped carrier by combining the amounts described in **Table 3**. Mix well by flicking the tube and measure the concentration using the microvolume spectrophotometer.

NOTE: If a higher concentration of the carrier is required to fit in the reverse transcription

reaction (see below), the carrier mix can be concentrated using the vacuum concentrator at low-medium temperature (30-35 °C) until reaching the desired final concentration. Steps 2–14 are modified from the standard nAnT-iCAGE protocol reported by Murata et al.[11]

2. Reverse transcription

2.1. Combine 1 µL of the RT primer (2.5 mM TCT-N6 dissolved in water, for sequence see **Supplementary Table 1**), 10 ng of total RNA of interest and 4990 ng of carrier mix (**Table 3**) in 10 µL of total volume in a low-binding PCR plate. Mix by flicking the tube.

NOTE: If the sample RNA is too diluted for reverse transcription (see below), combine it with the appropriate amount of the carrier, concentrate using the vacuum concentrator to 9 µL total volume, and add 1 µL of the RT primer. Adding the carrier earl, to reach 5 µg of RNA in total prevents sample loss.

2.2. Heat the mix from step 2.1 at 65 °C for 5 min, and place on ice immediately to prevent renaturation.

2.3. Prepare of the reverse transcription (RT) mix.

2.3.1. For each sample combine 6.1 µL of water (RNase- and DNase-free), 7.6 µL of 5x first-strand buffer, 1.9 µL of 0.1 M DTT, 1 µL of 10 mM dNTPs, 7.6 µL of trehalose/sorbitol mix (see recipe in Murata et al.[11]) and 3.8 µL of the recommended reverse transcriptase (see **Table of Materials**). Mix well by flicking the tube.

2.4. Add 28 µL of the RT mix into the PCR tube with 10 µL of RNA, carrier and the RT primer (total volume 38 µL). Mix well by pipetting.

NOTE: The mix is highly viscous due to trehalose/sorbitol. Mix until visibly homogenous.

2.5. Incubate in a thermal cycler using the following program: 25 °C for 30 s, 50 °C for 60 min, and hold at 4 °C.

2.6. **Purification of cDNA:RNA hybrids using SPRI magnetic beads**

2.6.1. Add 68.4 µL of the recommended RNAse- and DNase-free SPRI beads (see **Table of Materials**) to 38 µL of the RT mix (beads to sample ratio 1.8:1). Mix well by pipetting and incubate for 5 min at room temperature (RT).

2.6.2. Separate the beads on a magnetic stand for 5 min. Discard the supernatant and wash the beads twice with 200 µL of 70% ethanol (freshly prepared).

NOTE: The ethanol is added to the beads without mixing and while the tube is on the magnetic stand. The added ethanol is immediately removed. Care should be taken not to lose any beads

during washes as it may lead to sample loss.

2.6.3. While the tube is still on the magnetic stand, remove all traces of ethanol. Droplets of ethanol can be removed and pushed out of the tube using a P10 pipette. Do not let the beads dry.

2.6.4. Add 42 µL of water preheated at 37 °C to the beads and elute the sample by pipetting up and down 60x.

NOTE: Be careful not to cause foaming by pipetting as it may cause loss of beads (i.e., bound sample) in the foam.

2.6.5. Incubate at 37 °C for 5 min without the lid to allow evaporation of trace amounts of ethanol.

2.6.6. Separate the beads on a magnetic stand for 5 min and transfer the supernatant to a new plate.

NOTE: Try to retrieve all the supernatant to prevent sample loss while avoiding bead carryover. Use the P10 pipette to get the last sample droplets.

3. Oxidation

3.1. Add 2 µL of 1 M NaOAc (pH 4.5) into the purified RT reaction. Mix by pipetting, add 2 µL of 250 mM $NaIO_4$ and mix again.

3.2. Incubate on ice for 45 min. Cover the plate with aluminum foil to avoid light.

3.3. Add 16 µL of Tris-HCl (pH 8.5) into the oxidation mix to neutralize the pH.

3.4. Purify oxidized cDNA:RNA hybrids using SPRI magnetic beads. Add 108 µL of SPRI beads to 60 µL of the oxidation mix (1.8:1 beads to sample ratio). Repeat the purification as described in steps 2.6.1–2.6.6. Elute using 42 µL of water preheated to 37 °C.

NOTE: Freshly prepare 250 mM $NaIO_4$ by adding 18.7 µL of water per 1 mg of $NaIO_4$. $NaIO_4$ is light-sensitive; therefore, keep the solution in a tube covered with aluminum foil or in a light-resistant tube.

4. Biotinylation

4.1. Add 4 µL of 1 M NaOAc (pH 6.0) into the tube containing the purified oxidized sample and mix by pipetting.

4.2. Add 4 µL of 10 mM biotin solution, mix by pipetting and incubate for 2 h at 23 °C in a

thermal cycler to avoid light.

NOTE: Prepare the biotin solution by mixing 50 mg of biotin with 13.5 mL of DMSO. Make single-use aliquots and freeze at -80 °C.

4.3. Purify biotinylated cDNA:RNA hybrids using SPRI magnetic beads. Add 12 µL of 2-propanol and mix by pipetting. Add 108 µL of SPRI beads (1.8:1 beads to sample ratio) and repeat the purification as described in steps 2.6.1–2.6.6. Elute using 42 µL of water preheated at 37 °C.

NOTE: The protocol can be paused here, and samples frozen at -80 °C.

5. RNase I digestion

5.1. Prepare the RNase I mix by mixing 4.5 µL of 10x RNase I buffer with 0.5 µL of RNase I (10 U/µL) per each sample. Mix by pipetting.

5.2. Add 5 µL of the mix to each purified sample (45 µL in total). Mix by pipetting and incubate for 30 min at 37 °C.

6. Preparation of streptavidin beads

6.1. For each sample, mix 30 µL of the streptavidin beads slurry with 0.38 µL of 20 mg/mL tRNA. Incubate on ice for 30 min and mix every 5 min by flicking the tube.

NOTE: Resuspend the streptavidin beads slurry well before pipetting by flicking the bottle. The tRNA solution should be prepared according to Murata et al.[11]

6.2. Separate the beads on the magnetic stand for 2–3 min. Remove the supernatant.

6.3. Wash the beads by resuspending in 15 µL of buffer A. Separate the beads on the magnetic stand for 2–3 min and remove the supernatant. Repeat the wash and remove the supernatant.

6.4. Resuspend the beads in 105 µL of buffer A and add 0.19 µL of 20 mg/mL of tRNA. Mix well by pipetting.

NOTE: The beads should be prepared fresh prior to use. Start preparation of the beads during RNase I digestion. For multiple samples prepare the beads together in a single tube.

7. Cap-trapping

7.1. **Sample binding**

7.1.1. Add 105 µL of prepared streptavidin beads to 45 µL of the RNase I-treated sample. Mix well by pipetting and incubate at 37 °C for 30 min. Mix by pipetting every 10 min.

7.1.2. Separate the beads on the magnetic stand for 2–3 min. Remove the supernatant.

7.2. **Washing beads**

7.2.1. Add 150 µL of wash buffer A and resuspend the beads by pipetting. Separate the beads on the magnetic stand for 2–3 min and remove the supernatant.

7.2.2. Add 150 µL of the wash buffer B and resuspend the beads by pipetting. Separate the beads on the magnetic stand for 2–3 min and remove the supernatant.

7.2.3. Add 150 µL of the wash buffer C and resuspend the beads by pipetting. Separate the beads on the magnetic stand for 2–3 min and remove the supernatant.

NOTE: Buffers B and C should be preheated to 37 °C. Recipes for wash buffers A, B, and C are as described in Murata et al.[11]

7.3. **cDNA release**

7.3.1. Prepare 1x RNase I buffer by mixing 58.5 µL of water with 6.5 µL of 10x RNase I buffer.

7.3.2. Resuspend the beads in 35 µL of 1x RNase I buffer. Incubate at 95 °C for 5 min and transfer directly on ice for 2 min to prevent reassociation of cDNA. Hold the lids during transfer to ice as they may pop-off due to pressure build up.

7.3.3. Separate the beads for 2–3 min on a magnetic stand and transfer the supernatant to a new plate.

7.3.4. Resuspend the beads in 30 µL of 1x RNase I buffer. Separate the beads on the magnetic stand for 2–3 min and transfer the supernatant to the previously collected supernatant (total volume of eluted cDNA should be about 65 µL).

8. RNA removal by RNase H and RNase I digestion

8.1. Per sample, combine 2.4 µL of water, 0.5 µL of 10x RNase I buffer, 0.1 µL of RNase H, and 2 µL of RNase I.

8.2. Add 5 µL of the mix to the 65 µL of the released cDNA sample and mix by pipetting. Incubate at 37 °C for 15 min and hold at 4 °C.

8.3. Purify cDNA from the RNase digestion mix using SPRI magnetic beads. Add 126 µL of SPRI beads to 70 µL of degradation reaction and mix by pipetting. Follow purification steps as described for SPRI beads purification in 2.6.1–2.6.6. Elute using 42 µL of water preheated at 37 °C as described.

8.4. Prepare RNase I mix by combining 4.5 µL of 10x RNase I buffer and 0.5 µL of RNase I.

8.5. Add 5 µL of the RNase mix to the 40 µL of the purified cDNA sample. Mix by pipetting and incubate at 37 °C for 30 min. Hold at 4 °C.

8.6. Purify the sample using SPRI magnetic beads. Add 81 µL of SPRI beads to 45 µL of degradation reaction and mix by pipetting. Follow purification steps as described for SPRI beads purification in 2.6.1–2.6.6. Elute using 42 µL of water as described.

9. Ligation of 5' linker

9.1. Concentrate the purified cDNA sample to 4 µL using the vacuum concentrator. Keep the temperature at 30–35 °C. Test the volume using a pipette. If the sample has dried to completeness, dissolve by adding 4 µL of water.

NOTE: It is better to avoid drying to completeness to prevent sample loss.

9.2. Incubate the concentrated sample at 95 °C for 5 min and immediately place on ice for 2 min to prevent renaturation. Hold the lids while transferring the tubes as the lids may pop-off due to pressure build-up.

9.3. Incubate 4 µL of the 2.5 µM 5' linker at 55 °C for 5 min and immediately place on ice for 2 min to prevent renaturation.

9.4. Mix 4 µL of the 2.5 µM 5' linker with 4 µL of the sample.

NOTE: The 5' linker should be prepared according to **Supplementary Table 2**, **Supplementary Table 3**, **Supplementary Table 4**, and **Supplementary Table 5**. Dilute the 10 µM 5' linker to a 2.5 µM concentration using 100 mM NaCl prior to use.

9.5. Add 16 µL of the ligation premix (see **Table of Materials**) to the mixed 5' linker and the sample and mix well by pipetting. Incubate at 16 °C for 16 h.

9.6. Purify the ligation mix using SPRI magnetic beads. Add 43.2 µL of SPRI beads and follow steps 2.6.1–2.6.6. Elute as described using 42 µL of water preheated at 37 °C.

9.7. Repeat the purification done in step 9.6 by adding 72 µL of SPRI beads to the transferred supernatant (1.8:1 beads to sample ratio).

NOTE: 5' linkers contain barcodes which allows pooling of up to eight samples prior to sequencing (eight trinucleotide barcodes are available, as described in Murata et al.[11] and **Supplementary Table 1**).

10. Ligation of 3’ linker

10.1. Concentrate the purified sample to 4 µL using the vacuum concentrator as described in step 9.1**.**

10.2. Incubate the concentrated sample at 95 °C for 5 min and immediately place on ice for 2 min to prevent renaturation. Hold the lids while transferring the tubes as the lids may pop off due to pressure build-up.

10.3. Incubate 4 µL of the 2.5 µM 3’ linker at 65 °C for 5 min and immediately place on ice for 2 min to prevent renaturation.

10.4. Add 4 µL of the 2.5 µM 3’ linker to the 4 µL of the concentrated sample.

10.5. Add 16 µL of the ligation premix and mix well by pipetting. Incubate at 16 °C for 16 h.

10.6. Purify the ligation mix using SPRI magnetic beads. Add 43.2 µL of SPRI beads and follow steps 2.6.1–2.6.6. Elute as described using 42 µL of water preheated to 37 °C.

NOTE: The 3’ linker should be prepared according to **Supplementary Tables 6** and **Supplemental Table 7**. Dilute the 10 µM 3’ linker to a 2.5 µM concentration using 100 mM NaCl.

11. Dephosphorylation

11.1. Prepare the SAP mix by combining 4 µL of water, 5 µL of 10x SAP buffer, and 1 µL of SAP enzyme.

11.2. Add 10 µL of SAP mix to the purified ligated sample (total volume 50 µL) and incubate in the thermocycler using the following program: 37 °C for 30 min, 65 °C for 15 min, and hold at 4 °C.

12. Degradation of 3’ linker upper strand using uracil specific excision enzyme

12.1. Add 2 µL of uracil specific excision enzyme (see **Table of Materials**) to the dephosphorylated sample, mix by pipetting and incubate in the thermocycler using the following program: 37 °C for 30 min, 95 °C for 5 min, and immediately place on ice for 2 min to prevent reannealing of the fragmented upper strand.

12.2. Purify the reaction mixture by adding 93.6 µL of SPRI magnetic beads to the 52 µl mixture and mix well by pipetting. Repeat purification steps 2.6.1–2.6.6. Elute with 42 µL of water preheated at 37 °C as described.

13. Second strand synthesis

13.1. Prepare the second strand synthesis mix (volumes are expressed per sample) by combining 5 µL of 10x DNA polymerase reaction buffer, 2 µL of water, 1 µL of 10 mM dNTPs, 1 µL of 50 µM nAnT-iCAGE second strand primer (sequence is in **Supplementary Table 1**) and 1 µL of DNA exonuclease-deficient polymerase (see recommended polymerase in **Table of Materials**).

13.2. Add 10 µL of the mix to the purified sample and mix well by pipetting (total volume is 50 µL). Incubate in the thermal cycler using the following program: 95 °C for 5 min, 55 °C for 5 min, 72 °C for 30 min, and hold at 4 °C.

14. Degradation of second strand synthesis primer using Exonuclease I

14.1. Add 1 µL of Exonuclease I to the second strand synthesis mixture. Mix well by pipetting and incubate at 37 °C for 30 min followed by holding at 4 °C.

14.2. Purify double stranded DNA by adding 91.8 µL of SPRI magnetic beads to 51 µL of the Exonuclease I-treated sample. Repeat purification steps described in 2.6.1–2.6.6. and elute with 42 µL of water preheated to 37 °C as described.

14.3. Concentrate the sample using the vacuum concentrator to 15 µL as described in step 9.1.

15. Quality and quantity control

15.1. Use 1 µL of the concentrated samples and run a high sensitivity DNA chip on a DNA quality analyzer. Expected profile/quantity is presented in **Figure 3**.

16. First round of carrier degradation

16.1. Prepare the degradation mix by combining 2 µL of water, 2 µL of 10x restriction enzyme buffer, 1 µL of I-SceI, and 1 µL of I-CeuI.

16.2. Add 6 µL of the degradation mix to 14 µL of the concentrated sample and mix by pipetting. Incubate at 37 °C for 3 h followed by 20 min deactivation at 65 °C and hold at 4 °C.

16.3. Purify the degradation mix using SPRI magnetic beads. Add 5 µL of water to increase the volume of the degradation mix and add 45 µL of SPRI beads (1.8:1 beads to sample ratio). Repeat purification as described in steps 2.6.1–2.6.6. and elute with 42 µL of water preheated to 37 °C.

16.4. Concentrate the eluted sample from 42 µL to 20 µL of the total volume as described in step 9.1.

17. Control of degradation level and determining the number of PCR amplification cycles.

17.1. Prepare qPCR mix for amplifying whole libraries (adaptor mix). Combine 3.8 µL of water, 5 µL of qPCR premix (2x), 0.1 µL of 10 µM adaptor_f1 primer (5'-AATGATACGGCGACCACCGA-3'), and 0.1 µL of 10 µM adaptor_r1 primer (5'-CAAGCAGAAGACGGCATACGA-3') for each sample (see **Table of Materials** for recommended qPCR premix).

17.2. Combine 9 µL of qPCR adaptor mix with 1 µL of sample from step 16.4 and mix well by pipetting.

17.3. Prepare qPCR mix for amplifying DNA derived from the carrier (carrier mix). Combine 3.8 µL of water, 5 µL of qPCR premix (2x), 0.1 µl of 10 µM carrier_f1 primer (5'-GCGGCAGCGTTCGCTATAAC-3'), and 0.1 µL of 10 µM adaptor_r1 primer for each sample
17.4. Combine 9 µL of qPCR carrier mix with 1 µL of the sample from step 16.4 and mix well by pipetting.

17.5. Set qPCR program: 95 °C for 3 min (95 °C for 20 s, 60 °C for 20 s, 72 °C for 2 min) repeated 40x, followed by instrument-specific denaturation curve (65–95 °C), and hold at 4 °C.

NOTE: Prepare a negative control by replacing the sample with water.

18. PCR amplification of the target library

18.1. Prepare the PCR amplification mix by combining 6 µL of water, 0.5 µL of 10 µM adaptor_f1 primer, 0.5 µL of 10 µM adaptor_r1 primer and 25 µL of PCR premix (2x). Mix by pipetting (see **Table of Materials** for the recommended PCR premix).

18.2. Add 32 µL of the PCR mix to 18 µL of the sample from step 16.4. Mix thoroughly by pipetting.

18.3. Set the PCR amplification: 95 °C for 3 min, (98 °C for 20 s, 60 °C for 15 s, 72 °C for 2 min) 12-18 cycles, 72 °C for 2 min and hold at 4 °C.

NOTE: The exact number of PCR cycles is determined by qPCR results and corresponds to the Ct value obtained with the adaptor primer mix (number of PCR cycles is equal to the Ct value).

18.4. Purify the amplified sample by adding 90 µL of SPRI magnetic beads to 50 µL of the amplified sample and mix thoroughly by pipetting. Repeat the purification steps described in steps 2.6.1–2.6.6. and elute the sample using 42 µL of water as described.

19. Second round of carrier degradation

19.1. Repeat steps 16.1–16.3.

19.2. Purify the degradation mix using SPRI magnetic beads. Add 10 µL of water to the sample

to increase the volume and mix with 30 µL of SPRI beads (1:1 beads to sample ratio). Repeat purification as described in steps 2.6.1–2.6.6. and elute with 42 µL of water preheated at 37 °C as described.

19.3. Concentrate the eluted sample from 42 µL to 30 µL of total volume.

20. Library size selection

20.1. Mix 24 µL of SPRI magnetic beads with 30 µL of the sample from step 19.3. (0.8:1 beads to sample ratio). Repeat the purification steps as described in steps 2.6.1–2.6.6. and elute the sample in 42 µL of water as described.

20.2. Concentrate the sample to approximately 14 µL as described in step 9.1.

21. Quality control

21.1. **Assessment of size-distribution**

21.1.1. Run 1 µL of the sample on the high sensitivity DNA chip. Expected results are presented in **Figure 4**.

NOTE: If fragments shorter than 200 bp are visible (see example in **Figure 4A,C**), size selection (steps 20.1–20.2) should be repeated until the short fragments are removed (**Figure 4B,D**). Usually one additional round of size selection is enough. If the amount of short fragments is severe (as in **Figure 4C**), the beads-to-sample ratio should be decreased to 0.6:1.

21.2. **Carrier degradation quality control**

21.2.1. Repeat steps 17.1–17.5.

NOTE: Depending on the concentration of the libraries estimated in the HS DNA chip run (region analysis), the samples need to be diluted prior to qPCR. Use 0.5 µL of the sample to avoid sample loss and dilute 100–500x in water (dilute to 1–20 pg/µL final concentration). Expected difference between Ct values obtained with adaptor and carrier mix is 5–10.

21.3. **Library quantification**

21.3.1. Prepare the working dilution of the lambda DNA standard by mixing 20 µL of 100 mg/mL lambda DNA standard with 980 µL of 1x TE (prepare by diluting 20x TE provided in the DNA quantification kit). Dilution of the lambda DNA can be stored at -20 °C.

21.3.2. Prepare the lambda DNA standard serial dilutions by mixing the diluted lambda standard and 1x TE according to **Supplementary Table 8**.

NOTE: For higher accuracy, it is recommended to add 100 µL of 1x TE buffer to all tubes and remove 1x TE volume as required per volume of the diluted lambda to be added. Do not use more than 1 µL of the library; use of 384 well plates for this measurement is recommended.

REPRESENTATIVE RESULTS:

This report describes the full SLIC-CAGE protocol for obtaining sequencing-ready libraries from nanograms of starting total RNA material (**Figure 1**). To obtain the synthetic RNA carrier mix, first, PCR carrier templates need to be prepared and gel-purified to eliminate PCR side products (**Figure 2A**). Each PCR template (ten in total) is produced by using a common forward, but a different reverse primer (**Table 2**), leading to different lengths of the PCR template to enable size variability of synthetic RNA carriers. Once purified, PCR templates are used for in vitro transcription of the carrier molecules. A single RNA carrier product is expected if the templates are gel-purified (see representative gel-analysis in **Figure 2B**). Preparation of the carrier can be upscaled depending on the need, and when prepared, mixed and frozen at -80 °C for future use.

Using the recommended minimal amount of sample total RNA (10 ng) combined with 16–18 cycles of PCR amplification, high complexity SLIC-CAGE libraries can be achieved. Number of PCR cycles required to amplify the final library highly depends on the amount of total input RNA used (the expected number of cycles is presented in **Table 4**).

After the first round of degradation, in qPCR results (step 17), the expected difference between Ct values obtained using adaptor_f1 or carrier_f1 primer is 1–2, with Ct values obtained with adaptor_f1 lower than with carrier_f1.

The distribution of the fragment lengths in the final library is between 200–2000 bp with the average fragment size of 700–900 bp (based on the region analysis using Bioanalyzer software, **Figure 4B,D**). Shorter fragments, as presented in **Figure 4A,C**, have to be removed by additional rounds of size-exclusion (steps 20–21). These short fragments are PCR amplification artefacts and not the target library. Note that shorter fragments cluster better on the sequencing flow cells and may cause sequencing problems.

The expected amount of library material obtained per sample is between 5–50 ng. Significantly lower amounts are indicative of sample loss during the protocol. If the obtained low quantity is enough for sequencing (2–3 ng of the pooled libraries is needed), the libraries may be of lower complexity (see below).

Depending on the sequencing machine, quantity of the library loaded onto the flow cell may need to be optimised. Using an Illumina HiSeq 2500, loading 8–12 pM SLIC-CAGE libraries gives on average 150–200 million reads, with >80% of reads passing quality score Q30 as threshold. The obtained reads are then mapped to the reference genome [for 50 bp reads, Bowtie2[12] can be used with default parameters that allow zero mismatches per seed sequence (22 bp)]. Expected mapping efficiencies depend on the total RNA input amount and are presented in

Table 5. The uniquely mapped reads can then be loaded into R graphical and statistical computing environment[13] and processed using CAGEr (Bioconductor package[14]). The package vignette is easy to follow and explains the workflow and processing of the mapped data in detail. An easy visual control of the library complexity is the distribution of promoter width, as low-complexity libraries will have artificially narrow promoters (**Figure 5A**, SLIC-CAGE library derived from 1 ng of total RNA, for details see previous publication[10]). However, even the low-complexity SLIC-CAGE libraries allow identification of true CTSSs, with greater precision than alternative methods for low/medium-input TSS mapping (**Figure 5B,C**).

FIGURE AND TABLE LEGENDS:

Figure 1: Steps in the SLIC-CAGE protocol. Sample RNA is mixed with the RNA carrier mix to achieve 5 µg of total RNA material. cDNA is synthesised through reverse transcription and the cap is oxidized using sodium periodate. Oxidation allows attachment of biotin to the cap using biotin hydrazide. Biotin gets attached to the mRNA's 3' end, as it is also oxidized using sodium periodate. To eliminate biotin from mRNA:cDNA hybrids with incompletely synthesized cDNA and from the 3' ends of mRNA, the samples are treated with RNase I. cDNA that reached the 5' end of mRNA is then selected by affinity purification on streptavidin magnetic beads (cap-trapping). After release of cDNA, 5'- and 3'-linkers are ligated. The library molecules that originate from the carrier are degraded using I-SceI and I-CeuI homing endonucleases and the fragments are removed using SPRI magnetic beads. The library is then PCR amplified.

Figure 2: Representative gel-analysis of carrier PCR templates and carrier *in vitro* transcripts. (A) Carrier PCR templates prior to gel purification: the first well contains the 1 kbp marker, followed by carrier PCR templates 1, 1–10. (**B)** Carrier in vitro transcripts: the first well contains the 1 kbp marker, followed by carrier transcripts 1–10. Carrier transcripts were denatured by heating for 5 min at 95 °C prior to loading.

Figure 3: Representative DNA quality (high sensitivity DNA chip) trace of SLIC-CAGE prior to first round of carrier degradation.

Figure 4: Representative DNA quality (high sensitivity DNA chip) traces of SLIC-CAGE libraries after PCR amplification. (A) SLIC-CAGE library that requires additional size-selection for removal of short fragments. (**B)** SLIC-CAGE library after size-selection using 0.6x SPRI beads to sample ratio. (**C)** SLIC-CAGE library of lower output amount that requires size-selection for removal of short fragment. (**D)** SLIC-CAGE library of lower output amount after size-selection using 0.6:1 SPRI beads to sample ratio.

Figure 5: Validation of SLIC-CAGE libraries. (A) Distribution of tag cluster interquantile widths in SLIC-CAGE libraries prepared from 1, 5, or 10 ng of *S. cerevisiae* total RNA, and in the nAnT-iCAGE library prepared from 5 µg of *S. cerevisiae* total RNA. A high amount of narrow tag clusters in the 1 ng SLIC-CAGE library indicates its low complexity. (**B**) ROC curves for CTSS identification in *S. cerevisiae* SLIC-CAGE libraries. All *S. cerevisiae* nAnT-iCAGE CTSSs were used as a true set. (**C**) ROC curves for CTSS identification in *S. cerevisiae* nanoCAGE libraries. All *S.*

cerevisiae nAnT-iCAGE CTSSs were used as a true set. Comparison of ROC curves shows that SLIC-CAGE strongly outperforms nanoCAGE in CTSS identification. Data from ArrayExpress E-MTAB-6519 was used.

Table 1: Sequence of the carrier synthetic gene. I-SceI sites are bold and italicized in purple, and I-CeuI recognitions sites are green.

Table 2: Primers for carrier template amplification. Forward primer is the same for all carrier templates. Underlined is the T7 promoter sequence. PCR_GN5_f1: TAATACGACTCACTATAGNNNNNCAGCGTTCGCTA. Using differing reverse primers, PCR templates and hence carrier RNAs of differing length are produced.

Table 3: RNA carrier mix. In total 49 µg of the carrier mix 0.3–1 kbp: uncapped = 44 µg, capped = 5 µg.

Table 4: Expected number of PCR cycles in dependence of sample total RNA input. Approximate number of cycles is based on experiments performed using *Saccharomyces cerevisiae*, *Drosophila melanogaster,* and *Mus musculus* total RNA.

Table 5: Expected mapping efficiency and in dependence of total RNA input amount. Approximate numbers are presented and based on experiments performed using *Saccharomyces cerevisiae* and *Mus musculus* total RNA.

DISCUSSION:

For successful SLIC-CAGE library preparations, it is critical to use low-binding tips and tubes to prevent sample loss due to sample adsorption. In all steps involving retrieval of the supernatant, it is recommended to recover the entiresample volume. As the protocol has multiple steps, continuous sample loss will lead to unsuccessful libraries.

If CAGE (nAnT-iCAGE) has not been performed routinely, it is best to test SLIC-CAGE with different input amounts (10 ng, 20 ng, 50 ng, 100 ng, 200 ng) of the same total RNA sample and compare to nAnT-iCAGE libraries that are prepared using 5 µg of total RNA. If the nAnT-iCAGE library is unsuccessful (less than 0.5–1 ng of the DNA library obtained per sample), SLIC-CAGE is unlikely to work, and sample loss needs to be minimized.

A critical step to ensure high quality libraries devoid of uncapped degraded RNA or rRNA is the cap-trapping described in section 7. It is highly important that the streptavidin beads are thoroughly resuspended in wash buffers and that the wash buffers are removed prior to continuing to the next wash step or elution of cDNA.

If results from the qPCR after the first round of carrier degradation show no difference between the use of adaptor_f1 and carrier_f1 primers, continuing the protocol is still recommended. If after the second round of carrier degradation, the difference in Ct values is less than five, a

third round of carrier degradation is recommended. We have never found a third round of degradation necessary, and if it occurs, it is recommended to replace the homing endonuclease stocks.

Additional rounds of PCR amplification may be added to the protocol if the final amount of the library obtained is not enough for sequencing. PCR amplification can then be set with minimal number of amplification cycles needed to yield enough material for sequencing, taking into account sample loss that cannot be avoided in size selection. Purification or size selection using SPRI magnetic beads should then be performed until all small (<200 bp) fragments are removed (if needed, use 0.6:1 beads to sample ratio), and the library should be quantified using Picogreen.

Libraries can be sequenced in single-end or paired-end mode. Using paired-end sequencing, information about transcript isoforms can be obtained. In addition, as reverse transcription is performed using a random primer (TCT-N_6, N_6 being a random hexamer), information from the sequenced 3'-end can be used as unique molecular identifiers (UMI) to collapse PCR duplicates. As a moderate number of PCR amplification cycles is used (up to 18), the use of UMIs unnecessary has been previously found to be unnecessary.

As the core of the protocol relies on nAnT-iCAGE[11], SLIC-CAGE uses eight barcodes. Therefore, multiplexing more than eight samples is currently not supported. In addition, both SLIC-CAGE and nAnT-iCAGE are not suitable for capturing RNAs shorter than 200 bp, as the protocols are designed to remove linkers and PCR artefacts through size-exclusion with AMPure XP beads. SLIC-CAGE is the only unbiased low-input single-nucleotide resolution method for mapping transcription initiation start sites using nanograms of total RNA material. Alternative methods rely on the template switching activity of the reverse transcriptase to barcode capped RNA instead of cap-trapping (e.g., NanoCAGE[15] and NanoPARE[16]). Due to template switching, these methods exhibit sequence-specific biases in TSSs detection, leading to increased numbers of false positive TSSs and decreased numbers of true TSSs[9,10].

ACKNOWLEDGEMENTS:

This work was supported by The Wellcome Trust grant (106954) awarded to B. L. and Medical Research Council (MRC) Core Funding (MC-A652-5QA10). N. C. was supported by EMBO Long-Term Fellowship (EMBO ALTF 1279-2016); E. P. was supported by the Medical Research Council UK; B. L. was supported by the Medical Research Council UK (MC UP 1102/1).

DISCLOSURES:

A patent for degradable carrier RNA/DNA has been filled.

REFERENCES:

1 Shiraki, T. *et al.* Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proceedings of the*

National Academy of Sciences of the United States of America. **100** (26), 15776-15781, doi:10.1073/pnas.2136655100 (2003).

2 Haberle, V., Lenhard, B. Promoter architectures and developmental gene regulation. *Seminars in Cell and Developmental Biology.* **57** 11-23, doi:10.1016/j.semcdb.2016.01.014 (2016).

3 Haberle, V., Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nature Reviews Molecular Cell Biology.* **19** (10), 621-637, doi:10.1038/s41580-018-0028-8 (2018).

4 Andersson, R. *et al.* An atlas of active enhancers across human cell types and tissues. *Nature.* **507** (7493), 455-461, doi:10.1038/nature12787 (2014).

5 Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature.* **489** (7414), 57-74, doi:10.1038/nature11247 (2012).

6 Celniker, S. E. *et al.* Unlocking the secrets of the genome. *Nature.* **459** (7249), 927-930, doi:10.1038/459927a (2009).

7 Consortium, F. *et al.* A promoter-level mammalian expression atlas. *Nature.* **507** (7493), 462-470, doi:10.1038/nature13182 (2014).

8 Boyd, M. *et al.* Characterization of the enhancer and promoter landscape of inflammatory bowel disease from human colon biopsies. *Nature Communications.* **9** (1), 1661, doi:10.1038/s41467-018-03766-z (2018).

9 Adiconis, X. *et al.* Comprehensive comparative analysis of 5'-end RNA-sequencing methods. *Nature Methods.* doi:10.1038/s41592-018-0014-2 (2018).

10 Cvetesic, N. *et al.* SLIC-CAGE: high-resolution transcription start site mapping using nanogram-levels of total RNA. *Genome Research.* **28** (12), 1943-1956, doi:10.1101/gr.235937.118 (2018).

11 Murata, M. *et al.* Detecting expressed genes using CAGE. *Methods in Molecular Biology.* **1164** 67-85, doi:10.1007/978-1-4939-0805-9_7 (2014).

12 Langmead, B., Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods.* **9** 357, doi:10.1038/nmeth.1923 (2012).

13 R Core Team, A language and environment for statistical computing. <https://www.R-project.org/> (2017).

14 Haberle, V., Forrest, A. R., Hayashizaki, Y., Carninci, P., Lenhard, B. CAGEr: precise TSS data retrieval and high-resolution promoterome mining for integrative analyses. *Nucleic Acids Research.* **43** (8), e51, doi:10.1093/nar/gkv054 (2015).

15 Poulain, S. *et al.* NanoCAGE: A Method for the Analysis of Coding and Noncoding 5'-Capped Transcriptomes. *Methods in Molecular Biology.* **1543** 57-109, doi:10.1007/978-1-4939-6716-2_4 (2017).

16 Schon, M. A., Kellner, M. J., Plotnikova, A., Hofmann, F., Nodine, M. D. NanoPARE: parallel analysis of RNA 5' ends from low-input RNA. *Genome Research.* **28** (12), 1931-1942, doi:10.1101/gr.239202.118 (2018).

Figure 1

Click here to access/download;Figure;Figure1.eps

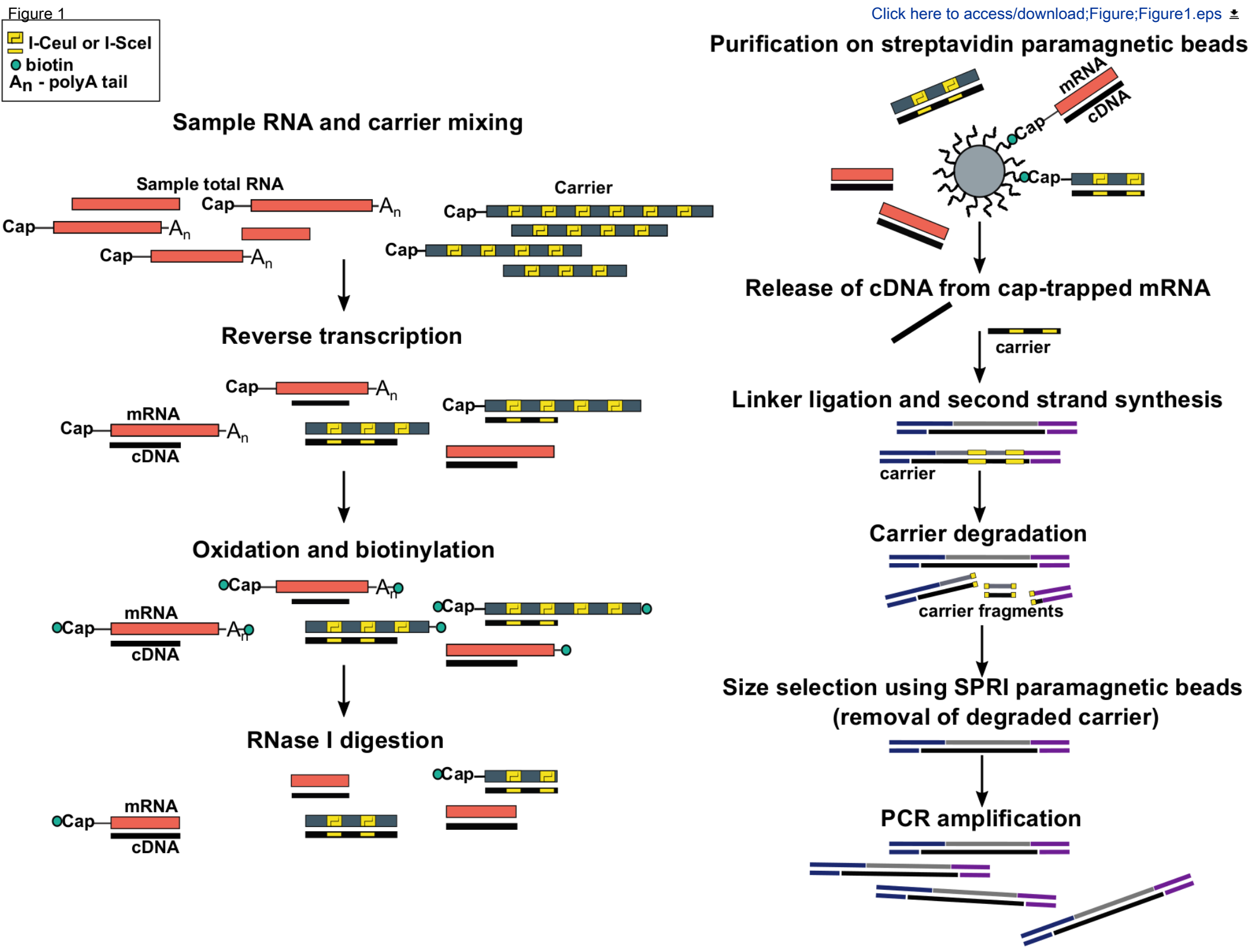

Figure 2

Click here to access/download;Figure;Figure2.eps

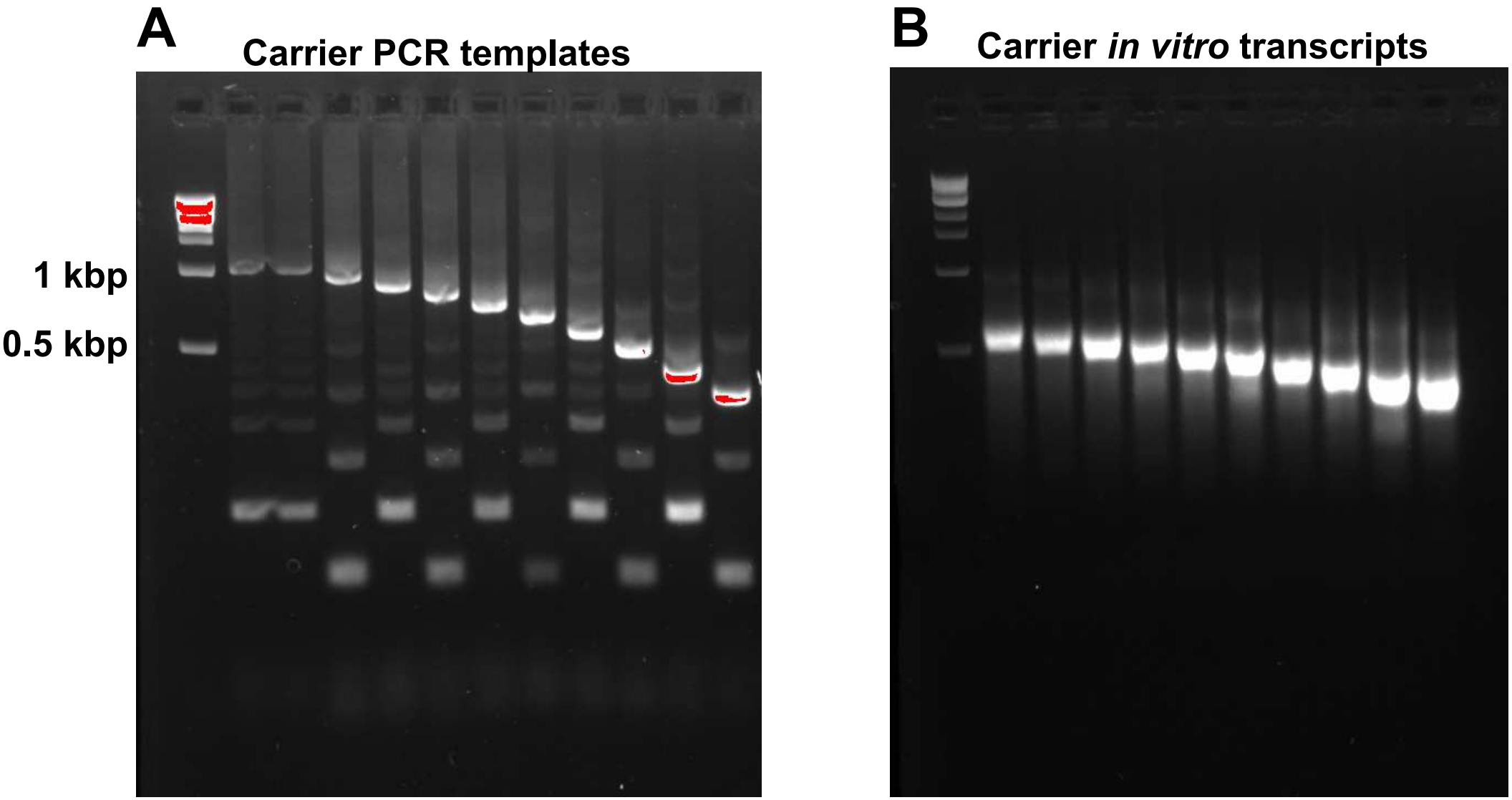

Figure 3

Click here to access/download;Figure;Figure3.eps

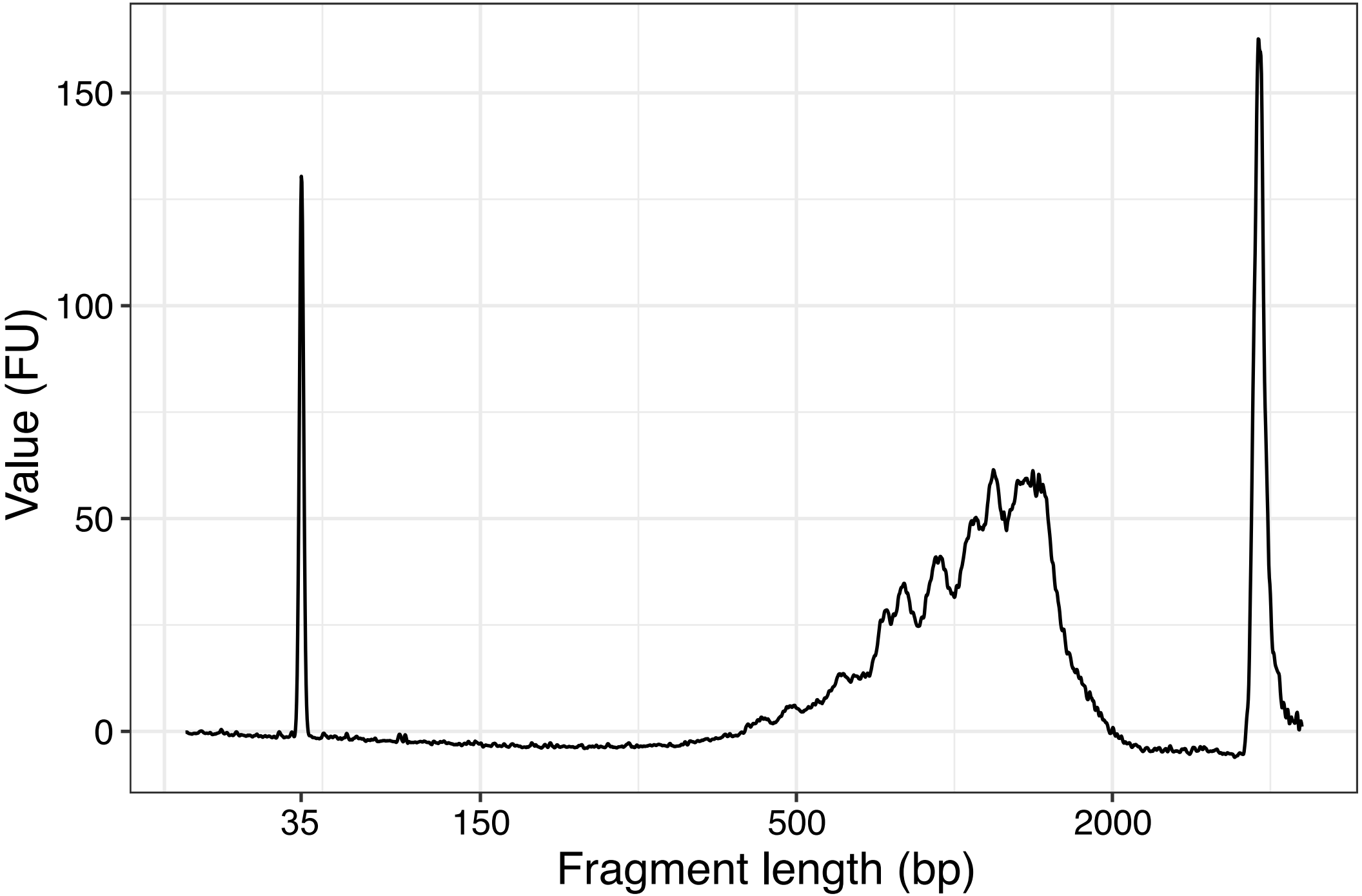

Figure 4

Click here to access/download;Figure;Figure4.eps

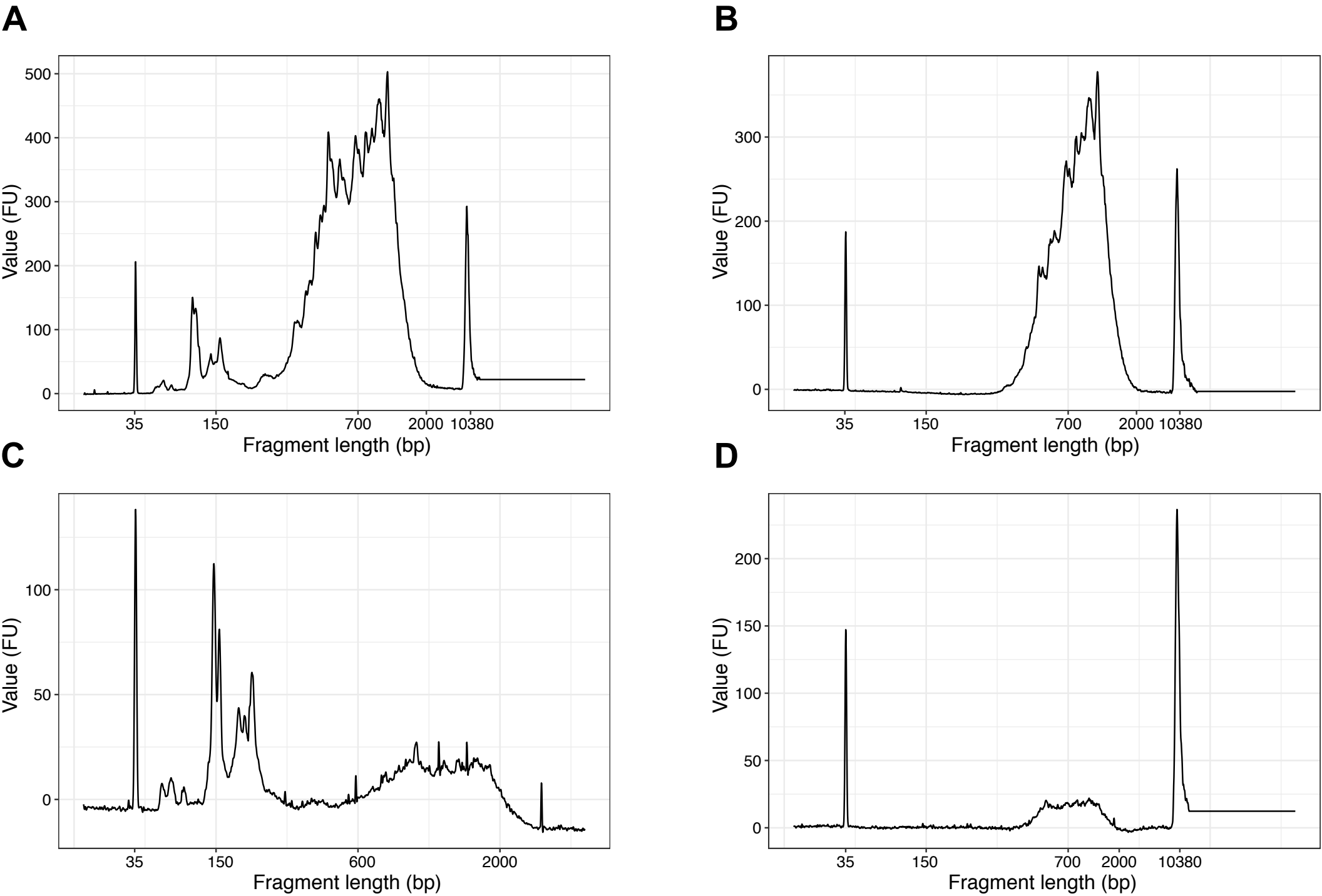

Figure 5

Click here to access/download;Figure;Figure5.eps

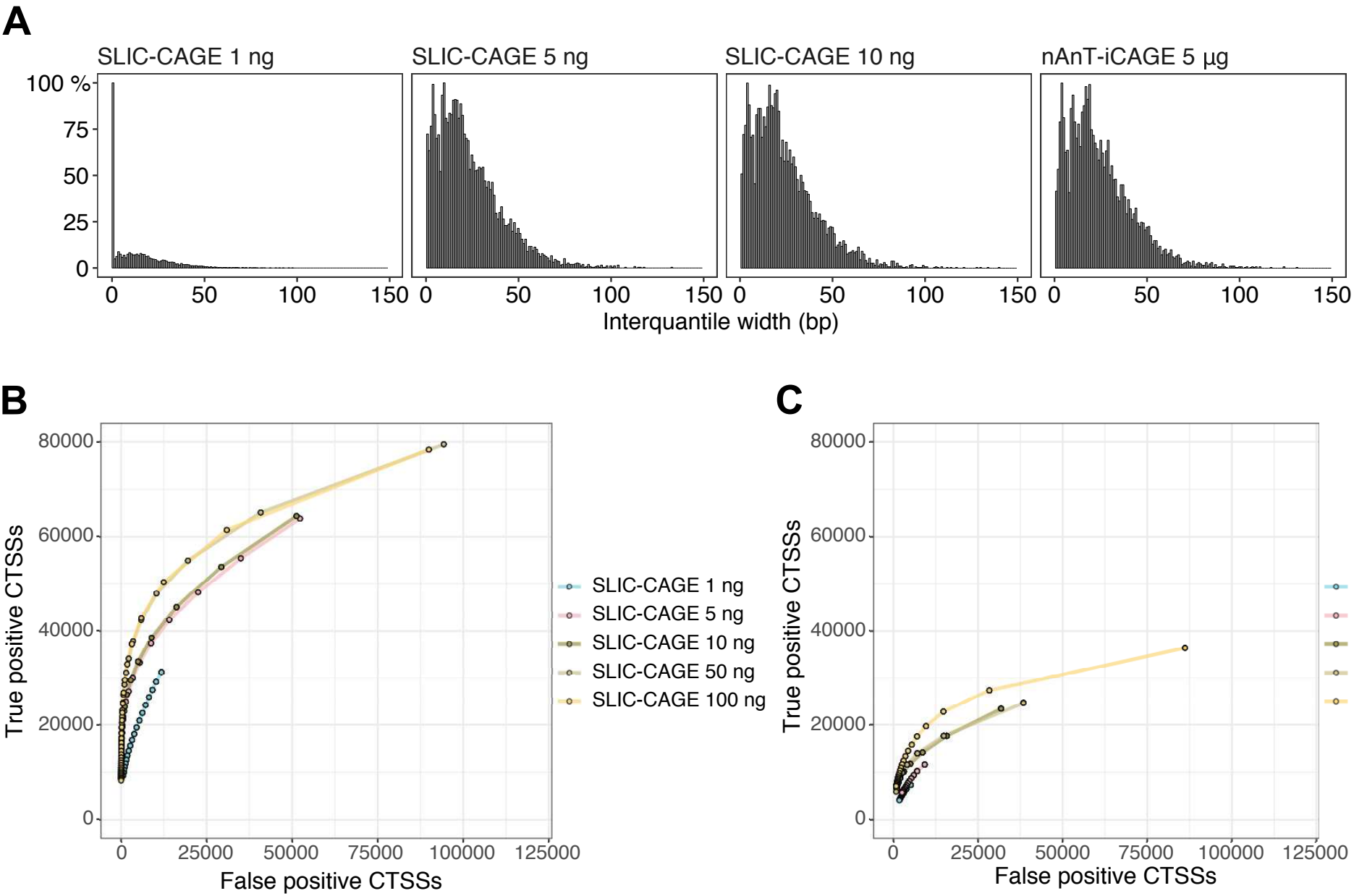

Table 1

5'CAGCGTTCGCTA***TAACTATAACGGTCCTAAGGTAGCGAA*** ATGCAAGAGCAATACCGCCCGGAAGAGATAGAA
TCCAAAGTACAGCTTCA***TAGGGATAACAGGGTAAT*** TTGGGATGAGAAGCGCACATTTGAAGTAACCGAAGACG
AGAGCAAAGAGA***TAACTATAACGGTCCTAAGGTAGCGAA*** AGTATTACTGCCTGTCTATGCTTCCCTATCCTTCTG
GTCGACTACACATG*TAGGGATAACAGGGTAAT* GGCCACGTACGTAACTACACCATCGGTGACGTGATCGCCCGC
TACCAGCG***TAACTATAACGGTCCTAAGGTAGCGAA*** TATGCTGGGCAAAAACGTCCTGCAGCCGATCGGCTGGG
I-SceI: ***TAGGGATAACAGGGTAAT***
I-CeuI: ***TAACTATAACGGTCCTAAGGTAGCGAA***

Table 2

carrier	reverse primer 5’-3’	PCR product length / bp
1	PCR_N6_r1: NNNNNNCTACGTGTCGCAGACGAATT	1034
2	PCR_N6_r2: NNNNNNTATCCAGATCGTTGAGCTGC	966
3	PCR_N6_r3: NNNNNNCACTGCGGGATCTCTTTACG	889
4	PCR_N6_r4: NNNNNNGCCGTCGATAACTTGTTCGT	821
5	PCR_N6_r5: NNNNNNAGTTGACCGCAGAAGTCTTC	744
6	PCR_N6_r6: NNNNNNGTGAAGAATTTCTGTTCCCA	676
7	PCR_N6_r7: NNNNNNCTCGCGGCTCCAGTCATAAC	599
8	PCR_N6_r8: NNNNNNTATACGCGATGTTGTCGTAC	531
9	PCR_N6_r9: NNNNNNACCGCCGCGCCTTCCGCAGG	454
10	PCR_N6_r10: NNNNNNCAGGACGTTTTTGCCCAGCA	386

* Forward primer is the same for all carrier templates.
Underlined is the T7 promoter sequence. PCR_GN5_f1:
TAATACGACTCACTATAGNNNNNCAGCGTTCGCTA

Table 3

carrier	length	uncapped / µg	capped / µg
1	1034	3.96	0.45
2	966	8.36	0.95
3	889	4.4	0.5
4	821	6.6	0.75
5	744	4.4	0.5
6	676	3.08	0.35
7	599	4.4	0.5
8	531	3.96	0.45
9	454	2.64	0.3
10	386	2.2	0.25

Table 4

Total RNA input /ng	PCR cycles
1 ng	18
2 ng	17
5 ng	16
10 ng	15-16
25 ng	14-15
50 ng	13-15
100 ng	12-14

Table 5

Total RNA input /ng	% overall mapped	% uniquely mapped	% carrier
1 ng	30	20-30	30
2 ng	60	20-50	10
5 ng	60-70	40-60	5-10
10 ng	60-70	40-60	5-10
25 ng	65-80	40-70	0-5
50 ng	65-80	40-70	0-3
100 ng	70-85	40-70	0-2

Material

2-propanol, Bioultra, for molecular biology, ≥99.5%
3' linkers
5' linkers
Agencourt AMPure XP, 60 mL
Agencourt RNAClean XP Kit
Axygen 0.2 mL Polypropylene PCR Tube Strips and Domed Cap Strips
Axygen 1 x 8 strip domed PCR caps
Axygen 1.5 mL Maxymum Recovery Snaplock Microcentrifuge Tube
Axygen 96 well no skirt PCR microplate
Bioanalyzer (or Tapestation): RNA nano and HS DNA kits
Biotin (Long Arm) Hydrazide
Cutsmart buffer
Deep Vent (exo-) DNA Polymerase
DNA Ligation Kit, Mighty Mix
dNTP mix (10 mM each)
Dynabeads M-270 Streptavidin
DynaMag-2 Magnet
DynaMag-96 Side Skirted Magnet
Ethanol, BioUltra, for molecular biology, ≥99.8%
Exonuclease I (E. coli)
Gel Loading Dye, Purple (6x), no SDS
HiScribe T7 High Yield RNA Synthesis Kit
Horizontal electrophoresis apparatus
I-CeuI
I-SceI
KAPA HiFi HS ReadyMix (2x)
KAPA SYBR FAST qPCR kit (Universal) 2x
Micropipettes and multichannel micropipettes (0.1-10 µl, 1-20 µl, 20-200 µ)
Microplate reader
nuclease free water
PCR thermal cycler

Phusion High-Fidelity DNA Polymerase
QIAquick Gel Extraction Kit (50)
qPCR machine
Quant-iT PicoGreen dsDNA Reagent
Quick-Load Purple 100 bp DNA Ladder
Quick-Load Purple 1 kb Plus DNA Ladder
Ribonuclease H
RNase ONE Ribonuclease
RNase-Free DNase Set
RNeasy Mini Kit
Sodium acetate, 1 M, aq.soln, pH 4.5 RNAse free
Sodium acetate, 1 M, aq.soln, pH 6.0 RNAse free
Sodium periodate
Sorenson low binding aerosol barrier tips, MicroReach Guard, volume range 10 µL, Graduatec
Sorenson low binding aerosol barrier tips, MultiGuard, volume range 1000 µL , Graduated
Sorenson low binding aerosol barrier tips, MultiGuard, volume range 20 µL , Graduated
Sorenson low binding aerosol barrier tips, MultiGuard, volume range 200 µL , Graduated
SpeedVac Vacuum Concentrator
SuperScript III Reverse Transcriptase
Trehalose/sorbitol solution
Tris-HCl, 1M aq.soln, pH 8.5
tRNA (20 mg/mL)
UltraPure Low Melting Point Agarose
USB Shrimp Alkaline Phosphatase (SAP)
USER Enzyme
Vaccinia Capping System
Wash buffer A
Wash buffer B
Wash buffer C

Company	Catalog Number
Sigma-Aldrich	59304-100ML-F
Beckman Coulter	A63881
Beckman Coulter	A63987
Axygen (available through Corning)	PCR-0208-CP-C
Axygen (available through Corning)	PCR-02CP-C
Axygen (available through Corning)	MCT-150-L-C
Axygen (available through Corning)	PCR-96-C
Agilent	
Vector laboratories	SP-1100
NEB	
NEB	M0259S
Takara	6023
ThermoFisher Scientific	18427013
Invitrogen	65305
ThermoFisher Scientific	12321D
ThermoFisher Scientific	12027
Sigma-Aldrich	51976-500ML-F
NEB	M0293S
NEB	B7025S
New England Biolabs	E2040S
NEB	R0699S
NEB	R0694S
Kapa Biosystems (Supplied by Roche)	KK2601
Kapa Biosystems (Supplied by Roche)	KK4600
Gilson	
ThermoFisher Scientific	AM9937

ThermoFisher Scientific	F530S
Qiagen	28704
ThermoFisher Scientific	P11495
NEB	N0551S
NEB	N0550S
Takara	2150A
Promega	M4261
Qiagen	79254
Qiagen	74104
VWR	AAJ63669-AK
Sigma-Aldrich	311448-100G
Sorenson (available through SIGMA-ALDRICH)	Z719390-960EA
Sorenson (available through SIGMA-ALDRICH)	Z719463-1000EA
Sorenson (available through SIGMA-ALDRICH)	Z719412-960EA
Sorenson (available through SIGMA-ALDRICH)	Z719447-960EA
ThermoFisher Scientific	18080044
ThermoFisher Scientific	16520050
Applied Biosystems (Provided by ThermoFisher Scientific)	78390500UN
NEB	M5505S
NEB	M2080S

Comments/Description
Used in RNAclean XP purification.
Sequences are described in Murata et al 2014 and Supplementary Table 1 of this manuscript. Annealing of strands to produce 3'linkers is des
Sequences are described in Murata et al 2014 and Supplementary Table 1 of this manuscript. Annealing of strands to produce 5'linkers is des
Purification of DNA
Purification of RNA and RNA:cDNA hybrids in CAGE steps.
Or any 8-tube PCR strips (used only for water and mixes).
Caps for PCR plates.
Low-binding 1.5 ml tubes, used for enzyme mixes or sample concentration.
Low-binding PCR plates - have to be used for all steps in the protocol. Note that plates should be cut to contain 2 x 8 wells for easier visibility
To determine quality of RNA, efficient size selection and final quality of the library (Tapestation can also be used)
Biotinylation/tagging
Restriction enzyme buffer
Second strand synthesis
Used for 5' and 3'-linker ligation
dNTP mix for production of carrier templates (or any dNTPs suitable for PCR)
Cap-trapping. Do not use other beads as these are optimised with the buffers used.
Magnetic stand for 1.5 ml tubes - used to prepare Streptavidin beads.
Magnetic stand for PCR plates (96 well-plates) - used with cut plates to contain 2 x 8 wells.
Used in AMPure washes. Any molecular biology suitable ethanol can be used.
Leftover primer degradation
agarose gel loading dye
Kit for carrier in vitro transcription
purification of carrier DNA templates from agarose gels
Homing endonuclease used for carrier degradation.
Homing endonuclease used for carrier degradation.
PCR mix for target library amplification
qPCR mix to assess degradation efficiency and requiered number of PCR amplification cycles
Use of Gilson with the low-binding Sorenson tips is recommended. Other micropippetes might not be compatible.. Different brand low-bindi
For Picogreen concentration measurement of the final library. Microplates are used to allow small volume measurement and reduce sample
Or any nuclease (DNase and RNase) free water
incubation steps and PCR amplficication

DNA polymerase for amplification of carrier templates (or any high fidelity polymerase)
Purification of carrier PCR templates from agarose gels.
determining PCR amplification cyle number and degree of carrier degradation
Used to measure final library concentration - recommended as, in our hands, it is more accurate and reproducible than Qubit.
DNA ladder
DNA ladder
Digestion of RNA after cap-trapping.
Degradation of single stranded RNA not protected by cDNA.
Removal of carrier DNA templates after in vitro transcription.
For cleanup of carrier RNA from in vitro transcription or capping
Or any nuclease (DNase and RNase) free solution
Or any nuclease (DNase and RNase) free solution
Oxidation of vicinal diols
Low-binding tips - recommended use throughout the protocol to minimise sample loss.
Low-binding tips - recommended use throughout the protocol to minimise sample loss.
Low-binding tips - recommended use throughout the protocol to minimise sample loss.
Low-binding tips - recommended use throughout the protocol to minimise sample loss.
concentrating samples in various steps to lower volume
Used for reverse transcription (1st CAGE step)
Preparation is described in Murata et al 2014.
1 M solution, DNase and RNase free
tRNA solution. Preparation is described in Murata et al 2014.
Or any suitable pure low-melt agarose.

Degradation of 3'linker's upper strand, Uracil Specific Excision Reagent/Enzyme
Enzymatic kit for in vitro capping of carrier molecules
Cap trapping washes. Preparation is described in Murata et al 2014.
Cap trapping washes. Preparation is described in Murata et al 2014.
Cap trapping washes. Preparation is described in Murata et al 2014.

cribed in the supplementary of this protocol.
cribed in the supplementary of this protocol.

of the samples

ıg tips may not be of equal quality and may increase sample loss.
waste.

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT - UK

Title of Article:	Transcription start site mapping using Super-Low Input Carrier-CAGE
Author(s):	Nevena Cvetesic, Elena Pahita, Boris Lenhard

Item 1: The Author elects to have the Materials be made available (as described at http://www.jove.com/publish) via:

☐ Standard Access ☑ Open Access

Item 2: Please select one of the following items:

☑ The Author is **NOT** a United States government employee.

☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution 3.0 Agreement (also known as CC-BY), the terms and conditions of which can be found at: http://creativecommons.org/licenses/by/3.0/us/legalcode ; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4 and 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video - Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video - Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties**. The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole

jove JOURNAL OF VISUALIZED EXPERIMENTS
1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT - UK

discretion andwithout giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contaminationdue to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13. **Fees.** To cover the cost incurred for publication, Jo VE must receive payment before production and publication the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name:	Nevena Cvetesic
Department:	Institute of Clinical Sciences
Institution:	Imperial College London, Faculty of Medicine
Title:	PhD
Signature:	Nevena Cvetesic
Date:	07/02/2019

Please submit a **signed** and **dated** copy of this license by one of the following three methods:

1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

Dear dr. Steindel,

Thank you for reviewing our manuscript entitled "Transcription start site mapping using Super-Low Input Carrier-CAGE" by Nevena Cvetesic, Elena Pahita and Boris Lenhard.

In this letter we address point-by-point all of the comments raised by the reviewers and all editorial comments. In the revised version of the manuscript with implemented responses to reviewer's comments, all newly added and modified parts are marked in red. We marked the parts for video production in yellow.

Editorial comments:
Changes to be made by the Author(s):
1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues. The JoVE editor will not copy-edit your manuscript and any errors in the submitted revision may be present in the published version.
The manuscript has been thoroughly proofread.

2. Please remove trademark (™) and registered (®) symbols from the Table of Equipment and Materials.
Trademark and registered symbols have now been removed from the Table of Equipment and Materials.

3. Please sort the Materials Table alphabetically by the name of the material.
We have now removed subsections in the Table of Equipment and Materials and sorted everything alphabetically.

4. JoVE cannot publish manuscripts containing commercial language. This includes trademark symbols (™), registered symbols (®), and company names before an instrument or reagent. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials and Reagents.
For example: RNAClean, Superscript, RNase ONE, AMPure XP, etc.
Please remove "AMPure" from Figure 1 as well.
We have now removed commercial language from the manuscript and Figure 1, and used the most generic terms where possible. We have kept the exact reagent/product and company name, along with the Catalogue numbers in the Table of Equipment and Materials.

5. There is a 10 page limit for the Protocol, but there is a 2.75 page limit for filmable content. Please highlight 2.75 pages or less of the Protocol (including headings and spacing) that identifies the essential steps of the protocol for the video, i.e., the steps that should be visualized to tell the most cohesive story of the Protocol. Remember that non-highlighted Protocol steps will remain in the manuscript, and therefore will still be available to the reader.
We have highlighted in yellow the essential steps of the protocol for the video (up to 2.75 pages in length).

6. Please ensure that the highlighted steps form a cohesive narrative with a logical flow

from one highlighted step to the next. Please highlight complete sentences (not parts of sentences). Please ensure that the highlighted part of the step includes at least one action that is written in imperative tense.
We have ensured the aforementioned narrative for the highlighted steps.

7. Please revise the text to avoid the use of any personal pronouns (e.g., "we", "you", "our" etc.).
We have revised the text to avoid the use of any personal pronouns.

8. Please do not abbreviate journal titles.
Full journal titles are now written in references.

Reviewers' comments:
Reviewer #1:

Cvetesic and collaborators propose an improved method for the Cap analysis of gene expression (CAGE) technique. CAGE is a highly sensitive method to detect transcription start sites (TSS) and can also be used for accurate quantification of gene expression. In the recent years, it has been extensively used to fine mapping active promoter and enhancer regions, thus providing critical information for the understanding of the regulation of gene expression. In the present study, the authors describe the SLIC-CAGE, a method allowing to perform CAGE experiments from very low amount of cells. The use and validation of this improved method has been published elsewhere (Genome Res. 28 (12), 1943-1956).

Comments:

1. To illustrate theirs results the authors should show some comparative analyses between data obtained with their method and published experiments performed with standard protocols as well as representative screen shots. If possible, sequencing of low complexity libraries could be shown for comparison.
We thank the reviewer for raising this point. We have now added a figure (Figure 5) that includes the mentioned visual assessment of library complexity – distribution of interquantile widths and include a low-complexity library for comparison. We also now include a comparison of SLIC-CAGE and nanoCAGE (alternative low/medium-input method based on template switching) performance based on CTSS identification (ROC curves) and use a nAnT-iCAGE as dataset to define true CTSSs (Figure 5B, C). We have added a comment at the end of the representative results paragraph pointing to the added Figure 5.

2. How the different size of the synthetic carrier are generated?
Each PCR template (ten in total) is produced by using a common forward, but a different reverse primer (Table 2), leading to a different length of the PCR template to enable size variability of synthetic RNA carriers. This has now been added to the paragraph representative results for clarity.

3. Figure 2: add the legend to the top of each panel.
We have now added the legend to the top of each panel.

4. Bioanalyser profiles shown in figures 3 and 4 should show the size distribution (bp) instead of Time (S).
We have changed the x-axis as suggested to show size distribution in bp.

5. The sequences of mentioned primers (5'linker, TCT-N6, etc) should be shown as supplementary table.
The sequences are now added in the Supplementary Table 1.

6. What is the "USER" enzyme?
USER enzyme is the commercial name for a mix of Uracil DNA glycosylase and DNA glycosylase-lyase Endonuclease VIII (produced by NEB). We have removed the commercial

names from the manuscript (commercial names are now used only in the Materials List). We have substituted USER name with uracil specific excision enzyme.

7. Step 8: how the number of cycles is determined should be specified.
This is specified in the Note under 18.3 step. For clarity we have now added that the number of PCR cycles is equal to the Ct value obtained with the adaptor primer mix.

8. Typos:
a. Line 51: "is added to the sample of interest"
b. Line 77: incomplete sentence: "TSS information is also proving important for profiling healthy and disease tissue »
c. Line 185: incomplete sentences?
We have corrected the mentioned typos and thoroughly checked the manuscript.

Reviewer #2:

In this paper Cvetesic et al., describe a rather long technic: SLIC-CAGE which maps the TSS used in cells/tissues. SLIC-CAGE allows to perform CAGE (CAP Analysis of Gene Expression) using very low amount of starting material (ngs of RNA) and its performance is demonstrated in a paper recently published in Genome Research.

The present paper is well written, easy to follow and well described; a sufficient amount of details are given for the readers to reproduce the protocols.

I only have minor comments:

1. The authors could try to limit the use of the reference of Murata et.al (11) to allow the readers to only refer to their new paper in the future.
We thank the reviewer for the comment and have limited the use of the reference in the revised manuscript. However, as the protocol is already very long, we have to reference some of the steps from Murata et al 2014, as the manuscript would exceed the limit of the journal.

2. It would be very valuable to have paragraph describing he critical points in the protocols to which further experimentalists should pay particular attention.
Critical points in the protocol were already discussed in the very beginning of the Discussion paragraph. We have now expanded this section and added a step critical for the quality of the library.

3. Could the authors explain the rationale of diluting primers in 0.1XTE why not water or TE (just a curiosity of the reviewer) which may also hit readers.
We only dilute the linker components in 0.1xTE (Supplementary Material, and as described in Takahashi et al 2012), all other primers are dissolved in water. We have now made this clear in the manuscript. In general, 1xTE is avoided due to high EDTA concentration that can inhibit enzymatic reactions by chelating Mg^{2+} ions, therefore we rather use 0.1xTE. As it is crucial that the 5' and 3' linkers stay annealed, NaCl addition is more relevant for

stabilization of the annealed linkers. We follow the protocol that was originally described for the tagging CAGE version in 2012 Takahashi manuscript. We have now made it clearer that this protocol for annealing and dilution is taken entirely from that publication.

4. It would be interesting to have a paragraph dedicated to processing of sequencing data.
We agree with the reviewer that it would be good to have a paragraph dedicated to processing of the sequencing data. However, due to length restrictions we cannot include a full paragraph as the manuscript is already at the page limit. We have now added a sentence on mapping and direct the reader that processing of the data and downstream analyses are well described in the CAGEr package vignette and publication.

Supplementary material for

Transcription start site mapping using Super-Low Input Carrier-CAGE

Nevena Cvetesic[1, 2*], Elena Pahita[1,2] and Boris Lenhard[1,2,3*]

[1]Institute of Clinical Sciences, Faculty of Medicine, Imperial College London, London W12 0NN, UK
[2]MRC London Institute of Medical Sciences, London W12 0NN, UK
[3]Sars International Centre for Marine Molecular Biology, University of Bergen, N-5008 Bergen, Norway

*Corresponding authors
Nevena Cvetesic
Institute of Clinical Sciences,
Faculty of Medicine,
Imperial College London and
MRC London Institute of Medical Sciences,
London, W12 0NN, UK
Email Address: ncvetesi@ic.ac.uk

Boris Lenhard
Institute of Clinical Sciences,
Faculty of Medicine,
Imperial College London and
MRC London Institute of Medical Sciences,
London, W12 0NN, UK
Email Address: b.lenhard@imperial.ac.uk

Email Addresses of Co-authors:
Elena Pahita (elena.pahita16@lms.mrc.ac.uk)

Supplementary Table 1. Primer sequences.

Primer name	Sequence	Step
RT primer	Phos TCTNNNNNN	2.1.
5'nAnT-iCAGE_01 N6*	CGACGCTCTTCCGATCTACCNNNNNN Phos	9.
5'nAnT-iCAGE_01 GN5*	CGACGCTCTTCCGATCTACCGN NNNN Phos	9.
5'nAnT-iCAGE_01 Dwn*	Phos GGTAGATCGGAAGAGCGTCG Phos	9.
3'nAnT-iCAGE N3 + AGA Up	NNNAGAUCGGAAGAGCGGUUCAGCAGGAA UGC CGAGACCGAUCUCGUAUGCCG UCUUCUGCUUG	10.
3'nAnT-iCAGE Dwn	CAAGCAGAAGACGGCATACGAGATCGGTCTC GGCATTCCTGCTGAACCGCTCTTCCGA	10.
nAnT-iCAGE 2nd primer	AATGATACGGCGACCACCGAGATCTACACTC TTTCCCTACACGACGCTCTTCCGATCT	13.1.

*barcoding is within the 5'linker. One barcode example is shown (ACC, underlined sequence). Other barcoding options are: CAC, AGT, GCG, ATG, TAC, ACG, GCT.

Supplementary Table 2. Annealing of 5’ and 3’ linkers

[a]Temperature / °C	Time/min
95	5
83	5
71	5
59	5
47	5
35	5
11	hold

[a]Speed of temperature change should be approximately 0.1 °C/s.

Note: linker annealing conditions are from[1] and linker sequences are from [2]

Supplementary Table 3. 5'linker annealing

component	concentration	volume	final concentration
	Annealing of 5' N6 linker		
5'nAnT-iCAGE_01 N6	1 mM	1 µl	100 µM
5'nAnT-iCAGE_01 Dwn	1 mM	1 µl	100 µM
NaCl	1 M	1 µl	0.1 M
0.1 x TE	1 mM Tris-Cl 7.5	7 µl	0.7 mM
	0.1 mM EDTA 8.0		0.07 mM
	Annealing of 5' GN6 linker		
5'nAnT-iCAGE_01 GN5	1 mM	1 µl	100 µM
5'nAnT-iCAGE_01 Dwn	1 mM	1 µl	100 µM
NaCl	1 M	1 µl	0.1 M
0.1 x TE	1 mM Tris-Cl 7.5	7 µl	0.7 mM
	0.1 mM EDTA 8.0		0.07 mM

Note: All single stranded linkers should be dissolved to be 1 mM in 0.1 x TE. Sequences of linkers are as in Supplementary Table 1 (and as in Murata et al 2014). Example of 5'linker 01 is shown (barcode ACC). Other barcoding options are shown below Supplementary Table 1. When annealing, use up (N6 and GN5) and down (Dwn) linkers with the same barcode. After mixing linkers according to Supplementary Table 3, run annealing program described in Supplementary Table 2. When annealing is done, mix the annealed linkers in 1:4 ratio (see Supplementary Table 4).

Supplementary Table 4. 5’linker mixing

component	concentration	volume	final concentration
5’nAnT-iCAGE_01 N6 annealed (N6 + Dwn)	100 µM	2.5 µl	20 µM
5’nAnT-iCAGE_01 GN5 annealed (GN5 + Dwn)	100 µM	10 µl	80 µM

Dilute the mixed 5’linkers to working concentration according to Supplementary Table 5. Store the rest of the linkers at -20 °C.

Supplementary Table 5. 5’linker dilution.

component	concentration	volume	final concentration
5’nAnT-iCAGE_01 (N6 and GN5)	100 µM	1 µl	10 µM
NaCl	0.1 M	9 µl	0.09 M

Store the diluted linkers at -20 °C.

Supplementary Table 6. 3’linker annealing

component	concentration	volume	final concentration
3’nAnT-iCAGE N3 + AGA Up	1 mM	1 µl	100 µM
3’nAnT-iCAGE Dwn	1 mM	1 µl	100 µM
NaCl	1 M	1 µl	0.1 M
0.1 x TE	1 mM Tris-Cl 7.5	7 µl	0.7 mM
	0.1 mM EDTA 8.0		0.07 mM

Note: All single stranded linkers should be dissolved to be 1 mM in 0.1 x TE. Sequences of linkers are as in Murata et al 2014. After mixing the single stranded linkers according to Supplementary Table 6, run the annealing program described in Supplementary Table 2. When annealing is done, dilute the 3’linker according to Supplementary Table 7.

Supplementary Table 7. 3’linker dilution.

component	concentration	volume	final concentration
3’nAnT-iCAGE	100 µM	2 µl	10 µM
NaCl	0.1 M	18 µl	0.09 M

Store the diluted linkers at -20 °C.

Supplementary Table 8. Preparation of standard serial dilutions.

Concentration /ng ml^{-1}	2 µg $µl^{-1}$ lambda DNA / µl	1 x TE / µl
0	0	100
10	1	99
20	2	98
40	4	96
60	6	94
80	8	92
100	10	90
200	20	80

Supplementary references

1 Takahashi, H., Lassmann, T., Murata, M. & Carninci, P. 5' end-centered expression profiling using cap-analysis gene expression and next-generation sequencing. *Nature Protocols.* **7** (3), 542-561, doi:10.1038/nprot.2012.005, (2012).

2 Murata, M. *et al.* Detecting expressed genes using CAGE. *Methods in Molecular Biology.* **1164** 67-85, doi:10.1007/978-1-4939-0805-9_7, (2014).