BAC ends library generation for Illumina sequencing

Karine Labadie (klabadie@genoscope.cns.fr)
Wincker Lab
Adriana Alberti (aalberti@genoscope.cns.fr)
Wincker Lab
Laura Briñas
Wincker Lab
Céline Orvain
Wincker Lab
Caroline Belser
Wincker Lab
Corinne Cruaud
Wincker Lab
Laurie Bertrand
Wincker Lab
Valerie Barbe
Wincker Lab
Jean-Marc Aury
Wincker Lab
Patrick Wincker
Wincker Lab

Method Article

Keywords: BAC ends libraries, Illumina sequencing, Cre-Lox recombination, Next Generation Sequencing

Posted Date: June 2nd, 2015

DOI: https://doi.org/10.1038/protex.2015.048

Abstract

Bacterial artificial chromosome \(BAC) libraries are still a valuable tool for _de novo_ assembly of complex genomes, such as many plants genomes. Shotgun sequencing of BACs, individually or by pools, produces first assemblies which usually need further improvement towards finished quality. We developed a new approach to obtain BAC ends libraries for Illumina sequencing \(BES), overcoming the expensive and time consuming BAC ends Sanger sequencing. This new method could be useful for improving _de novo_ assembly, especially in the case of highly repeated genomes.

Introduction

Here, we describe an easily transposable method for BES library preparation and sequencing on Illumina platform from standard BACs libraries \(Figure 1). DNA is mechanically fragmented and ends are repaired. After ligation of Cre-Lox adaptors, fragments of sizes around 8kb are purified from agarose gel and circularized. BAC ends are selected by a first 20-cycles PCR with vector specific primers. A nested PCR is performed to increase specificity and sensitivity in the selection and to join the first part of the Illumina adaptors sequences. The remaining sequences of Illumina adaptors are added in a third 5-cycles PCR. We validated the protocol by preparing BES libraries from pools of 384 BACs belonging to the Minimum Tiling Path \(MTP) of _Triticum aestivum_ 1B chromosome long arm \(1BL). The MTP includes 8,597 BAC clones \(average insert size of 150 kb) in pIndigo-BAC5 vector \(7.5 kb). They cover 502 Mb \(94% of the chromosome arm)[1]. This protocol can be adapted to any BAC library providing that: i) Primers for BES selective amplification should be designed on the vector in which BACs are cloned. ii) Primers used in the second PCR, for Illumina adaptor sequences addition, should be nested. iii) Size selection should be adapted to the vector size \(add 0.5-1.5 kb to the vector size). iv) The number of BACs to be included in each library should be adapted to the sequencing and assembly strategies used to produce the genome draft sequence.

Reagents

- Covaris g-TUBEs - Covaris, cat. no. 520079 - DNA 12000 reagents – Agilent Technologies, cat. no. 5067-1508 - DNA High sensitivity reagents – Agilent Technologies, cat. no. 5067-4626 - Qubit dsDNA HS assay kit - Life Technologies, cat. no. Q-33120 - Nucleospin Extract II DNA purification kit - Macherey-Nagel, cat. no. 740609.250 - End-It DNA end-repair kit - Epicentre, cat. no. 035ER81050 - T4 DNA Ligase kit - Promega, cat. no. M1804 - _Bst_ DNA Polymerase, Large Fragment \(8 000 U/ml) - NEB, cat. no. M0275S - PCR Nucleotide Mix \(10 mM each) - Roche, cat. no. 11581295001 - _Cre_ recombinase \(1 000 U/ml) - NEB, cat. no. M0298L - Exonuclease I \(20 000 U/ml) - NEB, cat. no. M0293L - ATP, lithium salt, pH 7 \(100 mM) - Roche, cat. no. 11140965001 - 1,4-Dithiothreitol \(ultrapure) - General lab supplier - Plasmid safe ATP-dependent DNase - Epicentre, cat. no. 035E3105K - AMPure XP 5 ml kit - Beckman Coulter, cat. no. A63880 - Megabase Agarose - Bio-Rad, cat. no. 161-3109 - TAE 50X - Bio-Rad, cat. no. 161-0743 - 1 kb DNA step ladder - Promega, cat. no. G6941 - SybrGreen - Sigma, cat. no. S9430 - ZymocleanTM Large Fragment DNA Recovery Kit - ZymoResearch, cat. no. D4046 - EDTA 0.5 M, pH8 - General lab supplier -

Tris 1M, pH8 - General lab supplier - NaCl 5M - General lab supplier - Kapa Hifi Hotstart NGS Library Amplification Kit - KapaBiosystems, cat. no. KK2612 - DNA gel loading dye \(6x) - Life Technologies, cat. no. R0611 - Ultrapure nuclease-free water - General lab supplier - Elution buffer: 10mM Tris-Cl, pH 8.5 - Qiagen, cat. no. 19086

Equipment

- 0.2 ml and 1.5 ml microtubes - General lab supplier - 1.5 ml DNA LoBind tubes - Eppendorf, cat. no.0030 108.051 - MiniSpin Plus microcentrifuge - Eppendorf, cat. no. 5453000011 - Heat blocks \(20°C - 70°C) - General lab supplier - 2100 Bioanalyzer - Agilent Technologies, cat. no. G2938C - Mini-centrifuge for quick ~2000 g spins - General lab supplier - Micro-wave oven - General lab supplier - Horizontal electrophoresis system, with 20-well combs, 15 x 15 cm UV- transparent tray, casting gates - Bio-Rad, cat. no. 1704402 - Standard power supply - Fisher Scientific, cat. no. S65533Q - Transilluminator- General lab supplier - Magnetic separation device - Life Technologies, cat. no. 12321D - Qubit fluorometer 2.0 - Life Technologies, cat. no. Q32866 - Thermal cycler - General lab supplier - General lab equipment \(pipettors, barrier tips, vortexer, lab vessels, scalpels, weighing scale, filters) - General lab supplier

Procedure

DNA input preparation and requirements The quality and quantity of starting material are important for successful BES library generation. The procedure described here has been optimized starting from 30 µg of pooled BACs DNA. BACs should be grown and extracted individually on 96-well plates using modified alkaline lysis extraction[2], which minimizes _E. coli_ genomic DNA contamination. After extraction, we recommend to perform a 0.5x AMPure beads cleanup: this step allows eliminating shorter useless DNA fragments \(<1kb) and concentrating DNA. Indeed, for shearing constraints, the final volume containing 30 µg should not exceed 900 µl. STEP 1: DNA fragmentation 1.1 Dispatch BAC pool DNA in 6 aliquots of 5 µg in 150 µL elution buffer. 1.2 Perform a first fragmentation on one aliquot in a Covaris g-TUBE device, following the manufacturer’s instructions. Spin in an Eppendorf MiniSpin Plus at 9,000 rpm for 60 sec. Invert the g-TUBE, and spin it again using the same speed and duration. 1.3 After the first fragmentation, check the approximate size range by loading 1 µl of a 1:2 DNA dilution on the Bioanalyzer 12000 chip or by running a 0.5% agarose gel. The expected DNA fragment size range is 7 to 10kb with a peak around 8kb. 1.4 If the size is correct, fragment the next five aliquots and pool them all together. 1.5 Nucleospin Extract II DNA Purification kit \(Macherey Nagel) is used to concentrate sheared DNA. Split shared DNA on two columns and follow the manufacturer’s protocol instructions for purification. The final product of each column is eluted with 65 µl NE buffer. STEP 2: sheared DNA repair 2.1 For each tube, combine and mix the following components in a LoBind tube using the End-It DNA end-repair kit: Sheared DNA 65 µl End Repair Buffer \(10x) 10 µl dNTP mix 10 µl ATP 10 µl End repair enzyme mix 5 µl Total reaction volume 100 µl 2.2 Mix by gently inverting the tube 5 times and incubate the reaction at 20 °C for 1 hour. 2.3 Purify each reaction on a Nucleospin Extract II DNA Purification column and elute in 100 µl NE buffer. At this point, DNA can be stored at -20 °C. STEP 3: Circularization adaptors ligation Adaptors sequences

LoxP_Ad1_top 5’-CGATAACTTCGTATAATGTATGCTATACGAAGTTATTACG-3’ LoxP_Ad1_bottom 5’-CGTAATAACTTCGTATAGCATACATTATACGAAGTTATCGAC*C-3’ LoxP_Ad2_top 5’-TCGTATAACTTCGTATAATGTATGCTATACGAAGTTATGCAC*C-3’ LoxP_Ad2_bottom 5’-GCATAACTTCGTATAGCATACATTATACGAAGTTATACGA-3’ * phosphorothioate modified base 3.1 Adaptors preparation Prepare 10X Annealing buffer: 100 mM Tris 10 mM EDTA, pH 8.0 500 mM NaCl For 10ml: Tris 1M 1 ml EDTA 0,5M 0.2 ml NaCl 5M 1 ml H_20 7.8 ml Total final volume 10 ml Resuspend lyophilized adaptor oligonucleotides at a concentration of 250 µM in H_2O. In a 0.2 ml tube, mix the following components: Annealing buffer \(10X) 5 µl LoxP_Ad1_top \(250 µm) 10 µl LoxP_Ad1_bottom \(250 µm) 10 µl H_20 25 µl Total final volume 50 µl Prepare the same mixture using LoxP_Ad2_top and LoxP_Ad2_bottom oligonucleotides. Anneal oligonucleotides in a thermal cycler 5 min at 95 °C, then decrease temp 0.1 °C/sec to reach 4 °C. Hold at 4 °C. Freeze annealed adaptors in 10 µl aliquots at -20 °C. 3.2 To each tube containing the repaired DNA, add: Promega T4 DNA ligase buffer \(10X) 15 µl LoxP_Ad1 3 µl LoxP_Ad2 3 µl H_20 19 µl Mix by gently inverting the tube 5 times. 3.3 Add 10 µl Promega T4 DNA ligase \(final reaction volume = 150 µl). Mix again by gently inverting the tube 10 times, then spin the tube briefly to collect the contents at the bottom. 3.4 Incubate for 16 hours at 16 °C. STEP 4: gel size selection 4.1 Prepare 0.5% Megabase agarose gel \(150 ml) using 1X TAE buffer. Use a large gel tray \(15x15 cm) and create large wells on the 20-wells comb by joining them together with tape. Each sample well should be approx. 2.4 cm wide. 4.2 Add 20 µl of 6X loading dye to each 150 µl ligation and load the two reactions on two separate large combs. Prepare 30 µl of 1 kb DNA step ladder \(900 ng ladder + loading dye) and load 10 µl into the lanes on either side of the sample lanes, leaving one empty lane between them. 4.3 Run the gel for 4 hours at 85 V \(constant voltage). 4.4 After migration, stain the gel by incubation in a diluted SybrGreen bath according to the manufacturer instructions. Examine the gel on a transilluminator to visualize DNA. 4.5 Using a clean scalpel blade, and the 1 kb DNA step ladder as a size guide, cut the gel slices containing DNA fragments ranging from 8 to 9 kb. 4.6 Purify the two collected DNA slices separately using Zymoclean Large Fragment DNA Recovery Kit, following the manufacturer instructions. Elute each purified DNA in 40 µl elution buffer. Combine elutions from the two matching columns, to give a total 80 µl of sample. 4.7 Quality control: run 1 µl of the elution on a Bioanalyzer 12000 LabChip to confirm the size of the extracted fraction and quantify how much DNA was recovered. STEP 5: fill-in reaction 5.1 Prepare two reactions in 1.5 ml LoBind tubes, each containing the following components: Size selected DNA 38 µl ThermoPol buffer \(10X) 5 µl PCR nucleotide mix \(10 mM each) 4 µl _Bst_ DNA polymerase, large fragment \(8 U/µl) 3 µl Total reaction volume 50 µl Mix by gently inverting the tubes 10 times, then spin briefly to collect the contents at the bottom. 5.2 Incubate the fill-in reactions for 15 minutes at 50 °C. 5.3 At the end of the incubation, pool the two reactions together and purify on a Nucleospin Extract II DNA Purification column. Elute in 50 µl NE buffer. 5.4 Quality control: remove 1 µl, dilute 1/10 and use 1 µl of the diluted sample for Qubit High Sensitivity quantitation. At this stage, sample can be stored at -20 °C overnight. STEP 6: DNA circularization 6.1 Set up 7 circularization reactions, each containing 300 ng DNA. Prepare each 300 ng aliquot of the filled-in DNA in a total volume of 80 µl \(adjust volume with ultrapure water). For each reaction, add the following components in a 0.2 ml tube, in the order shown: _Cre_ buffer 10X 10 µl Size selected DNA \(300 ng) 80 µl _Cre_ recombinase \(1 U/µl) 10 µl Total reaction volume 100 µl Mix by gently inverting the tubes 10 times, then spin briefly to

collect the contents at the bottom. 6.2 Run the following incubation program in a thermal cycler, with the heated lid on: 45 min at 37 °C, 10 min at 70 °C. STEP 7: linear DNA digestion 7.1 Prepare a stock of 1 M 1,4-Dithiothreitol \(DTT) solution in ultrapure water. Filter through a 0.45 µm filter and store in single use aliquots at -20°C. 7.2 Prepare a fresh DTT solution by mixing 2 µl 1M DTT with 18 µl ultrapure water. 7.3 At the end of the 70 °C incubation, add 1.1 µl 0.1M DTT to each tube. Gently mix by inverting the tubes 6 times and then spin briefly to collect the contents at the bottom. 7.4 Then, add the following reagents: ATP, lithium salt \(100 mM) 1.1 µl Plasmid-safe ATP-dependent DNase \(10 U/µl) 5 µl Exonuclease I \(20 U/µl) 3 µl Gently mix by inverting the tubes 10 times and then spin briefly to collect the contents at the bottom. 7.5 Run the following incubation program in a thermal cycler, with the heated lid on: 30 min at 37 °C, 10 min at 70 °C. 7.6 Immediately at the end of the incubation, add 5 µl EDTA 0.5 M and mix by inverting the tubes. 7.7 Split all the reactions in two 1.5 ml LoBind tubes. Purify the reactions on a unique Nucleospin Extract II DNA Purification column and elute in 24 µl NE buffer. Proceed immediately to amplification or store at 4 °C. Step 8: 1st PCR for BAC ends selective amplification Oligonucleotides sequences used for PCR amplifications are listed in Table 1. 8.1 Combine and mix the following components in a 0.2 ml tube: KAPA HiFi HotStart Ready Mix \(2X) 25 µl pIndigo_F \(10 µM) 1 µl pIndigo_R \(10 µM) 1 µl Circularized DNA 23 µl Total reaction volume 50 µl 8.2 Amplify using the following PCR cycling conditions: 45 sec at 98 °C, \[15 sec at 98 °C, 30 sec at 63 °C, 30 sec at 72 °C] 20 cycles total, 7 min at 72 °C, hold at 4 °C. 8.3 Clean up the reaction by adding 50 µL \(1 volume) AMPure XP beads and mix by short vortexing. Incubate for 5 minutes, then bind the beads and remove the supernatant. Add 200 µL 70% ethanol \(made fresh each time), incubate 30 seconds and remove. Repeat once. Let the pellet dry completely \(5 to 10 minutes), then elute in 25 µL elution buffer. STEP 9: 2nd PCR for Illumina sequences addition 9.1 Combine and mix the following components in a 0.2 ml tube: KAPA HiFi HotStart Ready Mix \(2X) 25 µl PCR_2_Illum_F \(10 µM) 1 µl PCR_2_Illum_R \(10 µM) 1 µl First PCR product 23 µl Total reaction volume 50 µl 9.2 Amplify using the following PCR cycling conditions: 45 sec at 98 °C, \[15 sec at 98 °C, 30 sec at 64 °C, 30 sec at 72 °C] 10 cycles total, 7 min at 72 °C, hold at 4 °C. 9.3 Clean up the reaction using AMPure XP beads as described at step 8.3. Elute in 25 µL elution buffer. 9.4 Quality control: remove 1 µl, dilute 1/20 with ultrapure water and use 1 µl of the diluted sample for Qubit High Sensitivity quantitation. Load 1 µl of diluted sample on a High Sensitivity DNA LabChip and run on a Bioanalyzer instrument. STEP 10: 3rd PCR for Illumina sequences addition 10.1 Perform four PCR reactions, each using 10 ng DNA from previous PCR as template. Combine and mix the following components in 0.2 ml tubes: KAPA HiFi HotStart Ready Mix \(2X) 25 µl PCR_3_Illum_F \(10 µM) 1 µl PCR_3_Illum_IndN_R * 1 µl PCR product \(10 ng) x µl H_2O 23-x µl Total reaction volume 50 µl *Use one of 12 different primers containing 6-bases index. 10.2 Amplify using the following PCR cycling conditions: 45 sec at 98 °C, \[15 sec at 98 °C, 30 sec at 60 °C, 30 sec at 72 °C] 5 cycles total, 7 min at 72 °C, hold at 4 °C. 10.3 Pool the four PCR reactions together in a 1.5 ml tube and clean up by adding 200 µl \(1 volume) AMPure XP beads. Follow the same procedure described at step 8.3. Elute in 25 µL elution buffer. 10.4 Final quality control: measure the concentration of the library by using 1 µl on a Qubit fluorometer with the dsDNA HS Assay kit. Prepare a 1ng/µl dilution, load 1 µl on a High Sensitivity DNA LabChip and run on a Bioanalyzer instrument. The expected library size range is 700 bp to 2000 bp \(Figure 2a). STEP 11: paired end sequencing on Illumina Miseq Sequence the final QC validated library on a Miseq instrument

using 150-bp length paired end mode \(V3 chemistry), in order to obtain BES 10x coverage. For optimal clusterization \(800-900 K cluster/mm²), after denaturation step, dilute library to 6pM in HT1. Combine the library with PhiX control spike-in \(≥ 20%).

Timing

The whole procedure takes ~ 5 days.

Troubleshooting

- STEP 1: if most of DNA fragments size range is not between 7 and 10kb, modify g-TUBE fragmentation conditions, starting from centrifugation speed. - STEP 4: low yield after gel size selection. Check the quality of input DNA. A high quality DNA is essential for successful library construction. - STEP 6: the number of circularization reactions has been optimized in order to maximize BAC ends diversity in the circularized DNA template. This should be adjusted by taking in consideration the number of BACs in the pool and a minimal circularization yield needed for BES amplification. We suggest performing at least 5 circularizations. - STEP 9: presence of a discrete peak on Bioanalyzer traces \(Figure 2b). This could be due to amplification of a specific sequence on the genome. In step 10, instead of four PCR, perform at least 10 PCR and purify final PCR product on a gel in order to exclude the discrete band. - STEP 11: low quality of sequencing run metrics \(Q30 and quality mean under Illumina specifications). This could be due to poor diversity of the library in the first bases. Increase PhiX spike-in % or add in the same lane high diversity libraries.

Anticipated Results

The method has been validated on a pool of 384 BACs belonging to the _Triticum aestivum_ 1BL chromosome BAC library. We produced 2,296,478 paired end sequences on an Illumina Miseq instrument \(V3 chemistry). After sequencing, an in-house quality control process was applied to reads that passed the Illumina quality filters, as described by Madoui et al[3]. Furthermore, an additional trimming process removed the sequences of the cloning vector and the circularization adaptor used during BES library construction \(Table 2). After this trimming process, we obtained 1,438,662 useful paired end sequences. They were mapped with bwa aligner \(version 0.6.1)[4] on the reference assembly obtained by standard Illumina sequences. Seventy-five percent of the BACs extremities were covered by the expected paired end sequences.

References

1. Philippe, R. et al. A high density physical map of chromosome 1BL supports evolutionary studies, map-based cloning and sequencing in wheat. Genome Biol 14, R64, doi:10.1186/gb-2013-14-6-r64 \(2013). 2. Choulet, F. et al. Structural and functional partitioning of bread wheat chromosome 3B. Science 345, 1249721, doi:10.1126/science.1249721 \(2014). 3. Madoui, M. A. et al. Genome assembly using

Nanopore-guided long and error-free DNA reads. BMC Genomics 16, 327, doi:10.1186/s12864-015-1519-z \(2015). 4. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754-1760, doi:10.1093/bioinformatics/btp324 \(2009).

Acknowledgements

We thank Robert Olaso and Laurene Giraut for helpfull discussions. Funding was provided by France Génomique \(ANR-10-INBS-0009-08) as part of Investissements d’Avenir Program.

Figures

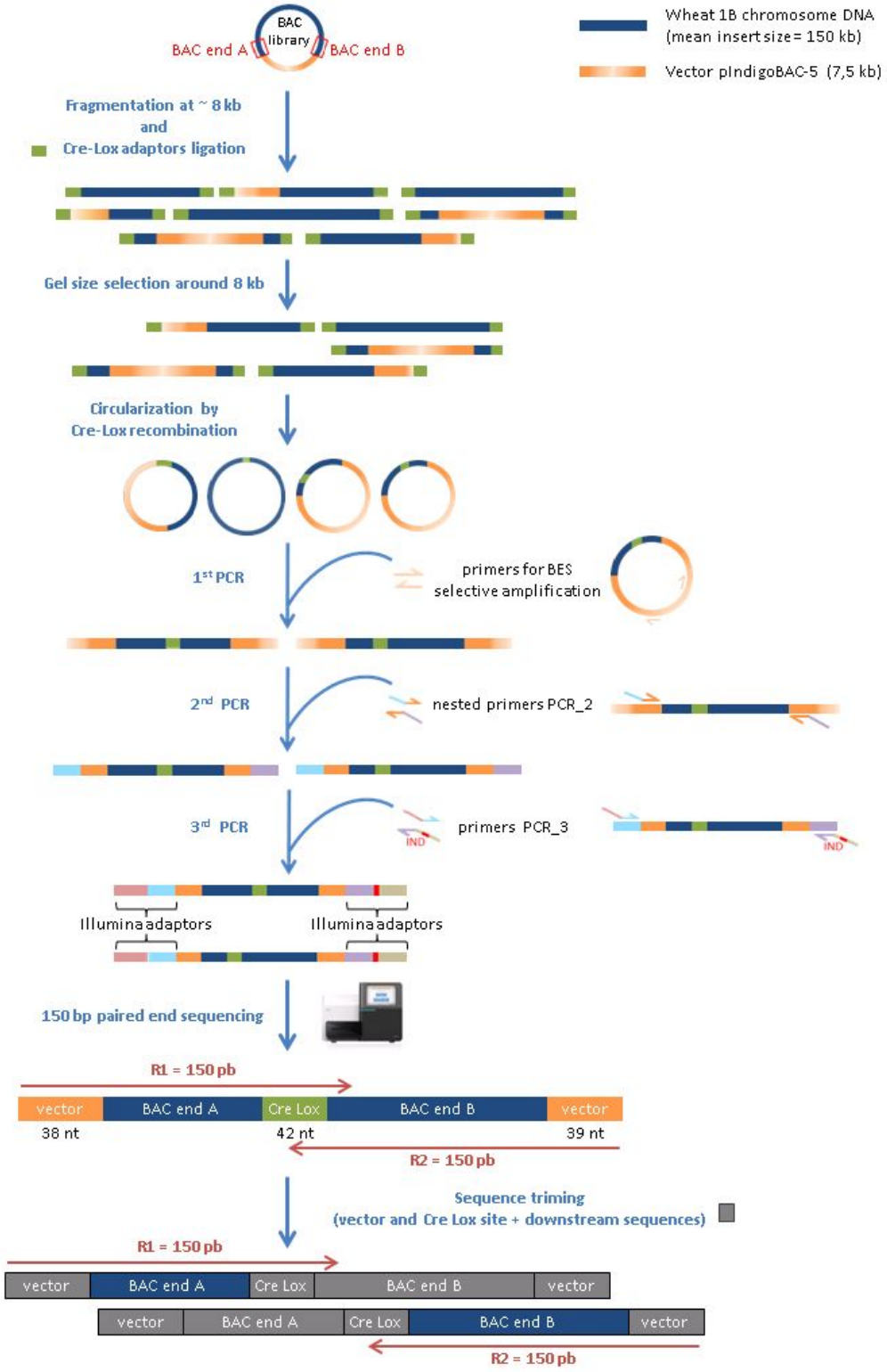
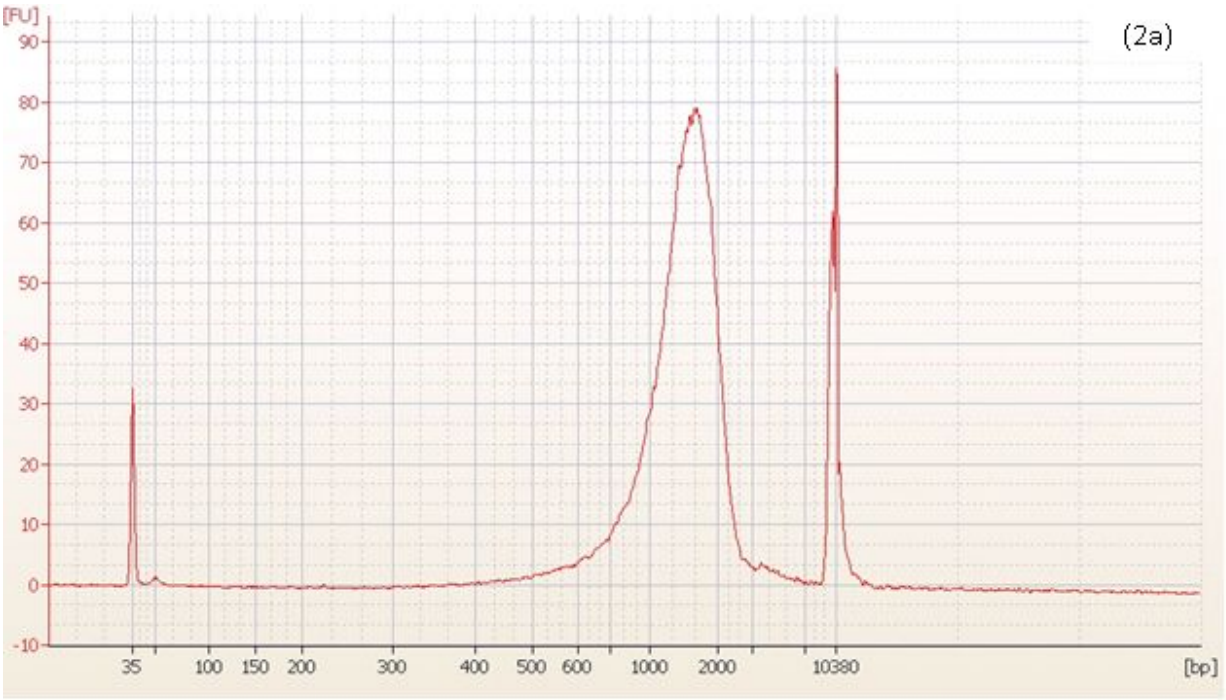


Figure 1

Library preparation workflow



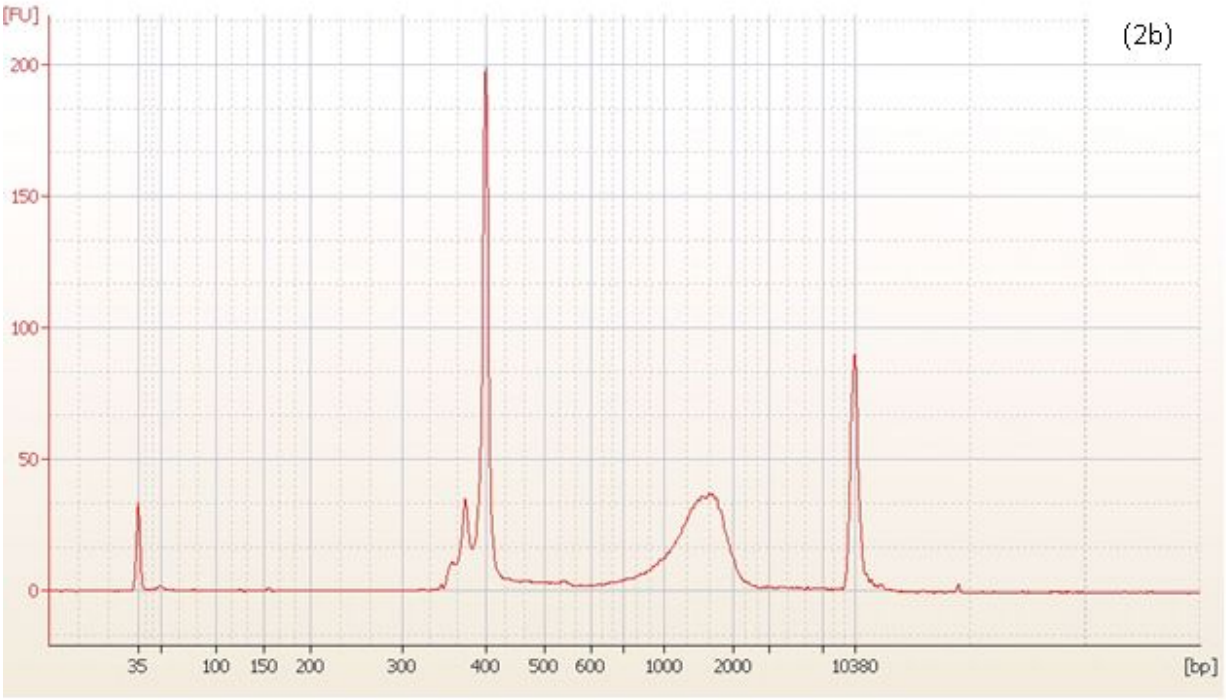


Figure 2: (a) expected final library profile; (b) library with a discrete peak.

Figure 2

Bioanalyzer traces after PCR 3

Supplementary Files

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

- supplement0.docx
- supplement0.docx