

SBCapSeq Protocol: a method for selective cloning of *Sleeping Beauty* transposon insertions using liquid capture hybridization and Ion Torrent semiconductor sequencing

CURRENT STATUS: POSTED

Karen M. Mann

Cancer Research Program, Houston Methodist Research Institute

✉ kmmann@HoustonMethodist.org *Corresponding Author*

Michael B. Mann

Cancer Research Program, Houston Methodist Research Institute

✉ mbmann@HoustonMethodist.org *Corresponding Author*

Liliana Guzman-Rojas

Cancer Research Program, Houston Methodist Research Institute

Felipe Amaya-Manzanares

Cancer Research Program, Houston Methodist Research Institute

Devin J. Jones

Cancer Research Program, Houston Methodist Research Institute

Justin Y. Newberg

Cancer Research Program, Houston Methodist Research Institute

Nancy A. Jenkins

Cancer Research Program, Houston Methodist Research Institute

Neal G. Copeland

Cancer Research Program, Houston Methodist Research Institute

DOI:

10.1038/protex.2016.053

SUBJECT AREAS

Biotechnology *Biological techniques*

KEYWORDS

Sleeping Beauty transposon mutagenesis, insertional mutagenesis, DNA

transposable elements, cancer genomics, genetic cancer models, genomic sequencing

Abstract

SBCapSeq is a transposon-based liquid-phase capture experimental and bioinformatic workflow optimized for Ion Torrent sequencing. SBCaptureSeq permits selective, semi-quantitative, and scalable deep sequencing of *Sleeping Beauty* transposon insertion sites from adaptor-ligated, barcoded Ion Torrent libraries created from populations of cells (bulk specimens) and single cells from both tumor and non-tumor genomes. This protocol includes detailed procedures for genomic DNA isolation, library preparation, capture hybridization, target sequence enrichment, sequencing, and data analysis. The SBCapSeq method, which takes about 7–10 days to perform, permits semi-quantitative sequencing and has been specifically optimized for Sleeping Beauty transposon mutagenesis genetic studies.

Introduction

Sleeping Beauty (SB) transposons are an important insertional mutagenesis tool for cancer gene discovery by performing forward genetic screens in mice¹⁻³. To broaden the capabilities and impact of SB-driven cancer gene discovery models, we set out to develop a new methodology that would allow rapid and reproducible identification of SB transposon insertion sites from bulk tumor tissue down to the single-cell level. We devised an approach that relies on random shearing of genomic DNA to discrete 200 bp fragments that are enriched for transposon sequences using liquid-capture hybridization with probes directed to the inverted repeats of the SB transposon ends. SBCapSeq is both an experimental sequencing approach and bioinformatic workflow optimized for Ion Torrent semiconductor sequencing. SBCapSeq permits the selective, semi-quantitative, and scalable deep sequencing of *Sleeping Beauty* Transposon insertion sites from adaptor-ligated, bar-coded DNA libraries from both bulk tumor specimens and single tumor cell genomes.

Reagents

Agencourt AMPure XP DNA Purification Kit (Beckman Coulter, Cat. No. A63881)

Absolute Ethanol (Fisher Scientific, Cat. No. BP2818-500)

2% Agarose, dye-free, w/ int stds, Pippin Prep 100-600 bp Marker L (Sage Science, Cat. No. CDF2010)

Axygen™ 8-Strip PCR Tubes (Axygen Scientific, Cat. No. PCR0208CPC)

Digital electrophoresis chips (Agilent High Sensitivity DNA Kit, Cat. No. 5067-4626)

DNA LoBind, 1.5 mL tubes (Eppendorf, Cat. No. 22431005)

Dynabeads[®] M-270 Streptavidin (Life Technologies, Cat. No. 65305)

Ethanol (Fisher Scientific, Cat. No. BP2818-500)

Gentra Puregene Tissue Kit (Qiagen, Cat. No. 158667)

Ion library TaqMan Quantitation Kit (Life Technologies, Cat. No. 4468802)

Ion Plus Fragment Library Kit (Life Technologies, Cat. No. 4471252)

Ion PI[™] Hi-Q[™] Chef Kit Thermo Fisher Scientific, Cat. No. A27198)

Ion PI[™] Chip Kit v3 (Thermo Fisher Scientific, Cat. No. A26771)

Ion Xpress[™] Plus Fragment Library Kit (Thermo Fisher Scientific, Cat. No. 4471269)

Ion Xpress[™] Barcode Adapters 1-16 Kit (Thermo Fisher Scientific, Cat. No. 4471250)

KAPA HiFi HotStart Readymix DNA polymerase (Kapa Biosystems, Cat. No. KK2601)

Mouse Cot-1 DNA[®] (Life Technologies, Cat. No. 18440-016)

Qubit[®] dsDNA HS Assay Kit (Life Technologies, Cat. No. Q32851)

Qubit[®] Assay Tubes (Life Technologies, Cat. No. Q32856)

SeqCap[®] EZ Hybridization and Wash Kits (Roche NimbleGen, Cat. No. 05634253001)

TE⁻¹ pH 8.0 Buffer (Tris-HCl 10mM, EDTA 0.1mM)(Teknova, Cat. No. T0227)

xGen Universal Blocking Oligo IT-A (Integrated DNA Technology, Cat. No. 118124656)

xGen Universal Blocking Oligo IT-P1 (Integrated DNA Technology, Cat. No. 118124657)

Custom oligonucleotides supplied by Integrated DNA Technologies, Inc. as 4 nmole Ultramer DNA

Oligos with Standard Desalting:

1. IRDRR Capture probe — 5'Target_r1_1:

/5Biosg/TACAGTTGAAGTCGGAAGTTTACATACACTTAAGTTGGAGTCATTA AAACTCGTTTTTCAACTACTCCAC/

2. IRDRL Capture probe — 3'Target_r1_1:

/5Biosg/TGGTGATCCTAACTGACCTAAGACAGGGAATTTTTACTAGGATTAATGTCAGGAATTGTGAAAAAGTG/

3. IRDRL Blocking probe 1 — UpstreamBlockers_r1_1:

/5Biosg/CCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCC/

4. IRDRL Blocking probe 2 — UpstreamBlockers_r1_2:
/5Biosg/GCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAG
5. IRDRR Blocking probe 1 — DownstreamBlockers_r1_1:
/5Biosg/TAGGGATCCTCTAGCTAGAGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCTTTAGTGAC
6. IRDRR Blocking probe 2 — DownstreamBlockers_r1_2:
/5Biosg/ATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCT
7. IRDRR Blocking probe 3 — DownstreamBlockers_r1_3:
/5Biosg/AGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGG

Equipment

AB Library Builder™ System (Applied Biosystems by Thermo Fisher Scientific, Cat. No. 4463794)

Agilent 2100 Bioanalyzer (Agilent, Cat. No. G2939AA)

Thermal cycler for PCR (General laboratory supplier)

Computer with 64-bit architecture, 8 GB RAM, 20 GB disk space, Intel Core i5 processor with MacOS 10.9+, Xcode 5+, and bash 3.2+

Covaris S220 Focused-ultrasonicator (Covaris, Inc., Cat. No. S220)

Covaris microTube AFA Fiber Pre-Slit Snap-Cap 6x16mm (Covaris Inc., Cat. No. 520045)

Covaris microTube Holder (Covaris, Inc., Cat. No. 500114)

DynaMag™ -2 magnet (Thermo Fisher Scientific, Cat. No. 12321D)

DynaMag™ -PCR magnet (Thermo Fisher Scientific, Cat. No. 49-2025)

Eppendorf ThermoMixer Temperature Control Device (Thermo Fisher Scientific, Cat. No. 05412503)

Ion Chef Instrument (Thermo Fisher Scientific, Cat No. 4484177)

Ion Proton System and Torrent Suite™ Software (Thermo Fisher Scientific, Cat No. 4476610)

Microcentrifuge (General laboratory supplier)

NanoDrop™ Lite Spectrophotometer (Thermo Fisher Scientific, Cat No. ND-NDL-PR)

Pippin Prep DNA Size Selection System (Sage Science, Cat. No. PIP0001)

Pipettors, 0.1–2 µL, 0.5–10 µL, 1–20 µL, 20–200 µL, 100–1,000 µL (General laboratory supplier)

Qubit® 3.0 Fluorometer (Thermo Fisher Scientific, Cat. No. Q33216)

Vacuum concentrator (General laboratory supplier)

ViiA™ 7 Real-Time PCR System (Applied Biosystems, Cat. No. 4453534)

Vortex mixer (General laboratory supplier)

Procedure

An overview of the SBCapSeq experimental workflow is shown in "Figure

1":http://www.nature.com/protocolexchange/system/uploads/4239/original/Figure_1.jpg?1454608223.

Detailed protocol steps follow:

A. Genomic DNA extraction (Timing overnight, 2 hours hands-on)

1. Choose one source of genomic DNA below.
 - a. Isolate genomic DNA from flash frozen specimens — Perform DNA extraction from frozen tissue according to the Gentra Puregene Handbook protocol for DNA purification from tissue frozen in liquid nitrogen. Using a razor blade or scalpel, shave off fine sections of frozen tissue equaling 10 mg, working on ice at all times, and add directly to Cell Lysis Solution. Follow the protocol exactly, including a 55 °C overnight incubation step in Proteinase K for maximum yields. In the final steps of isolation, re-suspend DNA pellet in 100 µL of DNA Hydration Solution. Quantify once DNA is completely dissolved —this step can be enhanced by overnight incubation at room temperature or 4 °C.
 - b. Isolate genomic DNA from whole genome amplified single cells — Single cells can be isolated by fluorescence-activated cell sorting followed by whole genome amplification according to methods outlined in Leung *et al.*⁴ and Wang, Y. *et al.*⁵. Genomic DNA obtained from these methods provides high-quality DNA in amounts amendable to the library preparation and sequencing methods described below.
2. Quantify genomic DNA extraction by either NanoDrop™ Lite Spectrophotometer or Qubit® 3.0 Fluorometer (using Qubit dsDNA BR Assay Kit per manufacturer's recommended instructions) to ensure sufficient yield to proceed with barcode library preparation. Repeat A.1 as needed to obtain ≥500 ng of

genomic DNA for library preparation.

B. Library preparation (Timing 4.5 hrs total, 75 mins hands-on)

SB capture hybridization sequencing libraries are constructed from 500 ng genomic DNA. Sheared genomic DNA of approximately 200 bp is achieved using one method of DNA fragmentation below.

1. Fragment genomic DNA by enzymatic shearing:
 - a. Follow the Ion Xpress™ Plus and Ion Plus Library Preparation for the AB Library Builder™ System User Guide for barcoded libraries.
 - b. Dilute 5 µL of Ion Xpress™ P1 Adapter and 5 µL of Ion Xpress™ Barcode Adapter in 40 µL E1 Buffer.
 - c. When starting the run, choose Auto SizeSelection and Auto Shearing to 200 bp target shear-size.
 - d. After the run, samples may contain small amounts of Agencourt AMPure XP Reagent. Remove by placing samples in a DynaMag™ -2 magnet for two minutes and transferring the supernatant to a new tube.
2. Fragment genomic DNA by acoustic shearing:
 - a. Reference the Covaris S220 System S220 Series User Manual for appropriate setup and maintenance of the Covaris System.
 - b. Use protocols for fragmentation in 130 µL volume to produce more consistent results.
 - c. Use a buffer of at least 0.1X TE for shearing to produce more consistent results.
 - d. Use screw cap microtubes with double rods to decrease sample splattering during a run.
 - e. Carefully pipette genomic DNA into the microTUBE AFA Fiber Pre-Slit Snap-Cap, ensuring that no air bubbles remain within the DNA and buffer.
 - f. Run the Covaris DNA Shearing protocol for 200 Target BP Peak found in the User Manual using the following settings: Peak Incident Power (W): 140, Duty Factor: 10%, Cycles per Burst: 200, Treatment Time (s): 180, Temperature (°C): 7, Water Level: 13.
 - g. Collect the fragmented genomic DNA by pipetting through the microTUBE cap slit.
 - h. Continue with library preparation according to the Ion Plus Library Preparation for the AB Library Builder™ System User Guide for barcoded libraries. Reference the specific protocol for library

preparation from pre-sheared genomic DNA. Elute DNA libraries in water and carefully transfer all supernatant to a new 1.5 mL tube.

Optional protocol pause point: barcoded DNA libraries can be stored at 4 °C overnight or at -20 °C for up to 1 week.

C. Quality control check and purification of bar-coded DNA library (Timing 40 min)

1. Use 1 µL of eluate from step B.1.d or B.2.h above to check the bar-coded DNA library integrity with an Agilent High Sensitivity DNA nanochip designed for use in the Bioanalyzer 2100 to measure the average fragment length of the sheared DNA libraries. Follow the Agilent High Sensitivity DNA Kit Quick Start Guide and select DNA, High Sensitivity in the Bioanalyzer software. An example sample trace at this step is provided in "Figure 2a":http://www.nature.com/protocolexchange/system/uploads/4241/original/Figure_2.jpg?1454608225.
2. Purify bar-coded DNA library fragments using Agencourt® AMPure XP beads:
 - a. Add 1.5X volume (45 µL) of Agencourt® AMPure XP beads to the bar-coded DNA library eluate.
 - b. Follow the bind and wash steps in the Agencourt® AMPure protocol, except use 80% ethanol for the washes.
 - c. Elute in 32 µL of TE⁻¹, pH8.0. Save 1 µL for analysis using Bioanalyzer 2100.
 - d. Transfer 30 µL of eluted product to a fresh 1.7 mL low-bind tube, ensuring no beads are carried over.

Optional protocol pause point: bar-coded DNA libraries may be stored at 4°C overnight or at -20°C for up to 1 week.

D. DNA library size selection (Timing 2 hrs)

1. Combine 30 µL of the eluate from step C.2.d with 10 µL of the Pippin Prep loading dye with Marker L. Mix well with micropipette and briefly spin to collect in bottom of tube.
2. Load 40 µL onto one lane of a 2% agarose gel for size selection of DNA libraries using Pippin

Prep according to the manufacturer's protocol.

3. Set the following parameters to ensure the majority of fragments are within the recommended length for Ion Torrent DNA sequencing libraries: Target = 315 bp, Start = 290 bp, End = 340 bp.
4. After run is complete, collect the eluted sample (~40 μ L) containing the size-selected DNA library fragments.
5. Purify size-selected DNA library fragments using Agencourt[®] AMPure XP beads:
 - a. Add 1.5X volume (60 μ L) of Agencourt[®] AMPure XP beads to the bar-coded DNA library eluate.
 - b. Follow the bind and wash steps in the Agencourt[®] AMPure protocol, except use 80% ethanol for the washes.
 - c. Elute in 25 μ L of TE⁻¹, pH8.0 and transfer all of eluted product into a fresh 1.7 mL low-bind tube, ensuring no beads are carried over.
 - d. An example sample trace at this step is provided in "Figure 2b":http://www.nature.com/protocolexchange/system/uploads/4241/original/Figure_2.jpg?1454608225.

Optional protocol pause point: Purified Pippin Prep size-selected DNA library fragments may be stored at -20 °C for up to 1 week.

E. PCR enrichment of DNA libraries after size selection (Timing 2.5 hrs)

Size selected DNA libraries are enriched using Platinum PCR mix, and Library Amplification Primer Mix from the Ion Plus Fragmented Library Kit.

1. Mix the following components in a PCR tube:
 - a. Size selected DNA library (from step D.5.c): 25 μ L
 - b. Platinum PCR High Fidelity mix: 100 μ L
 - c. Library Amplification Primer mix: 5 μ L
 - d. Total volume: 130 μ L
 - e. Split the 130 μ L in two PCR tubes, 65 μ L per tube.

2. Place the tubes in Thermocycler and run the following program: 95 °C for 5 min
95 °C for 15 sec, 58 °C for 15 sec, 70 °C for 60 sec (8-12 cycles)
70 °C for 2 min

Hold at 4 °C (do not leave samples at the hold temperature for more than 10 mins, proceed to clean up step immediately)

3. Clean up the PCR reaction with AMPure XP beads:
- Add 1.8X volume of beads (234 µL) to the reaction and follow the cleanup protocol.
 - At final step, elute DNA in 52 µL TE-1 pH 8.0 buffer and carefully transfer supernatant to a new 1.5 mL tube.
 - Use 1 µL to measure the concentration of DNA with a Qubit dsDNA HS Assay Kit using manufacturer's recommended instructions. Using Qubit readings, determine the concentration of each DNA library before proceeding to Capture Hybridization.
 - An example sample trace at this step is provided in "Figure 2c":http://www.nature.com/protocolexchange/system/uploads/4241/original/Figure_2.jpg?1454608225.

Optional protocol pause point: Purified Pippin Prep size-selected DNA library fragments may be stored at -20°C for up to 1 week.

F. Liquid capture hybridization (Timing 2 days)

Capture Hybridization is prepared according to the Integrated DNA Technologies protocol

“Hybridization capture of DNA libraries using xGen[®] Lockdown[®] Probes and Reagents.” Lockdown probes are designed to recognize the sequences located at the end of the transposon, and blocking oligos are designed to prevent capture of unmobilized transposons remaining at the transposon concatamer donor site. Two biotinylated lockdown capture probes and five biotinylated blocking oligos are designed for the capture hybridization.

- Prepare capture probes and blocking oligos.
 - Prepare transposon Lockdown[®] Probes[®] Blocking Oligos

b. Resuspend the dry capture probes and blocking oligos in 10mM Tris-HCl, 0.1mM EDTA, pH 8.0 (TE⁻¹) to a final pool concentration of 0.75 pmol/μL:

i. 5'Target_r1_1

ii. 3'Target_r1_1

iii. UpstreamBlockers_r1_1

iv. UpstreamBlockers_r1_2

v. DownstreamBlockers_r1_1

vi. DownstreamBlockers_r1_2

vii. DownstreamBlockers_r1_3

c. Mix thoroughly and briefly spin down to collect contents at the bottom of tube.

2. Prepare xGen Universal Blocking Oligos IT-P1 and IT-A by suspending in TE⁻¹ pH 8.0 to a final concentration of 1 μL/library (1X final concentration).

3. Combine universal blocking oligos, mouse Cot-1 DNA and the genomic DNA library and lyophilize. a. Mix the following in a low-bind 1.5 mL PCR tube:

Barcoded library: 60 ng

Mouse Cot-1 DNA: 5 μg

xGen[®] Universal Blocking Oligo (IT-A): 1 μL

xGen[®] Universal Blocking Oligo (IT-P1): 1 μL

b. Dry the contents of the tube using a vacuum concentrator set at 65 °C or lower.

Optional protocol pause point: After lyophilization, tubes can be stored overnight at 4 °C.

4. Hybridize DNA capture probes with the library: a. Add the following to the tube from Step F.3.b, and incubate at room temperature for 5-10 min:

8.5 μL SeqCap 2X Hybridization Buffer

3.4 μL SeqCap Hybridization Component A

1.1 μL Nuclease-Free Water

b. Pipette up and down to mix, and transfer to a low-bind 0.2 mL PCR tube.

- c. Incubate in a thermal cycler at 95°C for 10 min.
- d. Remove samples from thermal cycler and immediately add 4 µL of the pool of 2 lockdown probes and 5 blocking oligos.

Note: Final volume will be 17 µL.

- e. Vortex and briefly spin down.
- f. Incubate samples in a thermal cycler at 65°C (with the heated lid at 75°C) for 4 hr.

5. Prepare wash buffers: a. For each capture, dilute the following SeqCap[®] EZ

Hybridization and Wash Kit buffers to create 1X working solutions as follows:

- i. Prepare 2.5X Bead Wash Buffer by adding 200 µL of concentrated buffer plus 300 µL of Nuclease-free water
 - ii. Prepare 1X Wash Buffer I by adding 30 µL of concentrated buffer plus 270 µL of Nuclease-free water
 - iii. Prepare 1X Wash Buffer II by adding 20 µL of concentrated buffer plus 180 µL of Nuclease-free water
 - iv. Prepare 1X Wash Buffer III by adding 20 µL of concentrated buffer plus 180 µL of Nuclease-free water
 - v. Prepare 1X Stringent Wash Buffer by adding 40 µL of concentrated buffer plus 360 µL of Nuclease-free water
 - vi. Note: 1X working solutions can be stored at room temperature for up to 2 weeks.
- b. For 1X Wash Buffer I and 1X Stringent Wash Buffer, create 2 aliquots per reaction and store at the temperature specified as following:
- i. 1X Wash Buffer I: 100 µL at 65°C and 200 µL at Room temp (15–25°C)
 - ii. 1X Stringent Wash Buffer: Two aliquots of 200 µL at 65°C
 - iii. Important: Preheat buffers in a 65°C water bath for at least 2 hours before use in Step 8.
- c. Keep the remaining 1X buffers at room temperature.

6. Prepare the streptavidin beads: a. Important: Beads should be prepared immediately

before use. Do not allow beads to dry out.

b. Equilibrate Dynabeads[®] M-270 Streptavidin beads at room temperature for approximately 30 min before use.

Note: Use only Dynabeads[®] M-270 Streptavidin magnetic beads per IDFT protocol.

c. Mix the beads thoroughly by vortexing for 15 sec.

d. Aliquot 100 μ L of beads per capture into a single 1.7 mL low-bind tube (e.g. for 1 capture reaction, use 100 μ L of beads and for 2 capture reactions, use 200 μ L of beads. For more than 6 capture reactions, more than one tube will be required).

e. Place the tube in a magnetic separation rack, allowing beads to fully separate from the supernatant.

f. Remove and discard the clear supernatant, ensuring that the beads remain in the tube.

g. Perform the following wash:

i. Add 200 μ L of 1X Bead Wash Buffer per capture, and vortex for 10 sec.

ii. Place the tube in the magnetic rack, allowing beads to fully separate from the supernatant.

iii. Carefully remove and discard the clear supernatant.

h. Perform a second wash by repeating Step F.6.g.

i. Add 100 μ L of 1X Bead Wash Buffer per capture (refer to Step F.6.d) and vortex.

j. Transfer 100 μ L of the resuspended beads into a new 0.2 mL low-bind tube for each capture reaction.

k. Place the tube in a magnetic rack, allowing beads to fully separate from the supernatant.

l. Carefully remove and discard the clear supernatant.

m. Note: Small amounts of residual Bead Wash Buffer will not interfere with downstream binding of the DNA to the beads.

7. Bind hybridized target to the streptavidin beads: a. Transfer the hybridization

samples (from Step F.4.f) to the tube containing prepared beads (from Step F.6.l).

b. Mix thoroughly by pipetting up and down 10 times.

- c. Bind the DNA to the beads by placing the tube into a thermal cycler set to 65°C (with the heated lid at 75°C) for 45 min.
 - d. Every 15 min during the 65 °C incubation, carefully open the PCR tube lid and pipet up and down 8 times to ensure that the beads remain in suspension.
8. Wash streptavidin beads to remove unbound DNA:
- a. Note: Use the 1X wash buffers from Step F.5.
 - b. 65°C washes. Important: Work quickly to ensure the temperature remains close to 65 °C.
 - i. Add 100 µL preheated 1X Wash Buffer I to the tube from Step F.7.d.
 - ii. Important: Vortex for 10 sec.
 - iii. Transfer the mixture to a new low-bind 1.5 mL tube.
 - iv. Important: Vortex briefly.
 - v. Place the tube in the magnetic rack, allowing beads to fully separate from the supernatant.
 - vi. Pipet and discard the supernatant, which contains unbound DNA.
 - vii. To wash, add 200 µL of preheated 1X Stringent Wash Buffer, and slowly pipet up and down 10 times.
 - viii. Important: Do not create bubbles during pipetting.
 - ix. Incubate in a water bath at 65°C for 5 min.
 - x. Place the tube in the magnetic rack, allowing beads to fully separate from the supernatant.
 - xi. Pipet and discard the supernatant, which contains unbound DNA.
 - xii. Repeat wash with Stringent Wash Buffer (Step F.8.b.vii-xi).
 - c. Room temperature washes
 - i. Add 200 µL of room temperature 1X Wash Buffer I and vortex for 2 min.
 - ii. Place the tube in the magnetic rack, allowing beads to fully separate from the supernatant.
 - iii. Pipet and discard the supernatant.
 - iv. Add 200 µL of room temperature 1X Wash Buffer II and vortex for 1 min.
 - v. Place the tube in the magnetic rack, allowing beads to fully separate from the supernatant.
 - vi. Pipet and discard the supernatant.

- vii. Add 200 μL of room temperature 1X Wash Buffer III and vortex for 30 sec.
 - viii. Place the tube in the magnetic rack, allowing beads to fully separate from the supernatant.
 - ix. Pipet and discard the supernatant.
- d. Bead resuspension (not elution)
- i. Remove the tube from the magnetic rack and add 23 μL of Nuclease-Free Water to the beads.
 - ii. Pipet up and down 10 times and ensure any beads stuck to the side of the tube have been resuspended.
 - iii. Important: Use the entire 23 μL of resuspended beads with captured DNA in Step G. Do not place the tube on the magnetic rack again.
- e. An example sample trace at this step is provided in "Figure 2d":http://www.nature.com/protocolexchange/system/uploads/4241/original/Figure_2.jpg?1454608225.

G. Post-capture PCR enrichment (Timing 2.5 hrs)

1. Prepare the PCR mix in 0.2 mL low-bind PCR tubes as follows:
 - a. 2x KAPA HiFi Hot Start Ready mix DNA Polymerase: 25 μL
 - b. Library Amplification Primer mix (from the Ion Plus Fragmented Library Kit): 2 μL
 - c. Beads with captured DNA (from Step F.8.d.iii): 23 μL
 - d. Total volume: 50 μL
2. Briefly vortex and spin the PCR mix, but ensure that the beads remain in solution.
3. Place the PCR tube in the thermal cycler, and run the following program with the heated lid set at 105°C:
 - 98 °C for 45 sec
 - 98 °C for 15 sec, 65 °C for 30 sec, 72 °C for 30 sec (11-14 cycles)
 - 72 °C for 1 min
 - Hold at 4 °C

Optional protocol pause point: After post-capture PCR enrichment, tubes may be stored at 4°C overnight.

H. Post-capture PCR fragment clean-up (Timing 0.5 hrs)

1. Add 1.5X volume (75 μ L) of Agencourt[®] AMPure XP beads to the PCR-enriched capture.
2. Follow the bind and wash steps in the Agencourt[®] AMPure protocol, except use 80% ethanol for the washes.
3. Elute in 35 μ L of TE⁻¹ pH 8.0
4. Transfer 35 μ L of eluted product to a fresh 1.7 mL low-bind tube, ensuring no beads are carried over.

Optional protocol pause point: Purified PCR fragments may be stored at 4 °C overnight or at -20 °C for up to 1 week.

I. Validate and quantify library (Timing 4.5 hrs total, 3.5 hrs hands-on)

1. Using 1 μ L of eluate from step H.4, measure the average fragment length of the captured library on an Agilent 2100 Bioanalyzer using a high sensitivity DNA nanochip.
2. Using 2 μ L of eluate from step H.4, quantify libraries using ViiA[™] 7 Real-Time PCR System and the Ion library TaqMan Quantitation Kit, as directed by the manufacturer.
3. Determine the concentration of the ~30 μ L of eluate from step H.4 for sequencing.

J. Sequencing of SBCapSeq libraries (2-3 days)

1. Dilute sequencing libraries for Ion Torrent sequencing according to experimental design:
 - a. For sequencing a single SB Capture Hybridization library per PI[™] chip, prepare a diluted stock of eluted library from step H.4 to 50 pmol.
 - b. For multi-plex sequencing of 2–16 individual SB Capture Hybridization libraries per PI[™] chip, prepare a diluted stock of consisting of a pool of 2–16 eluted libraries from step H.4 to 50 pmol.
2. Set up Ion Chef for Ion PI[™] Hi-Q[™] Chef Kit according to the manufacturer's instructions using the sequencing library/pooled-libraries mixture from step J.1. Ion Chef chip loading runs overnight.
3. The next morning, prepare Ion Torrent Proton for sequencing and load PI[™] chip using the

manufacturer's instructions. A 500-flow sequencing run will complete in approximately 4 hours.

4. After a sequencing run is complete, Torrent Suite™ Software will align reads to a Reference Genome and generate a chip run report. Typically this analysis runs overnight.
5. It is important to note that the portion of the PI™ chip run report pertaining to the metrics for sequence data after aligning sequenced reads to a reference genome (for mouse the default is mm10) is expected to look similar to "Figure 3":http://www.nature.com/protocolexchange/system/uploads/4243/original/Figure_3.jpg?1454608227. Specifically, the range of expected Reference Coverage is between 1.1X to 2.9X and Unaligned bases is between 40% to 70% when successful sequencing of SBCapSeq libraries is achieved (see Figure 3). The variation in values differs depending upon the number of pooled libraries per chip and the complexity (e.g. transposon copy number, etc.) of SB insertions in a given tumor, or pooled set of tumors. This occurs because of SB Capture Hybridization enrichment of non-Reference Genome reads belonging to the SB transposon are not able to map well to the Reference genome.
6. Once the sequencing run and genome alignment is complete, download and archive the Ion Torrent Suite generated sequencing data files.

K. SBCapSeq data analysis (~1 day)

Note that elements below contained in square brackets ([text]) are variables that the user must set; before running the code replace [text] with your non-bracketed text. Refer to the README.md file at "<https://github.com/jynewberg/sbcapseq>":<https://github.com/jynewberg/sbcapseq> for detailed instructions and a demonstration of this workflow on a toy data set.

1. Prepare computer workstation for data analysis by downloading the required installer and python scripts from "<https://github.com/jynewberg/sbcapseq>":<https://github.com/jynewberg/sbcapseq>. If you download the master archive ("<https://github.com/jynewberg/sbcapseq/archive/master.zip>":<https://github.com/jynewberg/sbca>

pseq/archive/master.zip), make sure to unzip the contents. Once you do this, you should see an installer script (`'install.sh'`), a python directory, and some documentation and license files.

2. Install the toolbox and reference indexes by executing the installer script (`'install.sh'`). Using Terminal.app (or similar terminal emulator), navigate to the directory containing the installer, then run the following command:

```
<code>bash install.sh</code>
```

3. Add the sbcapseq toolbox scripts to your system paths by running the following command:

```
export PATH=$HOME/nnlab/sbcapseq/bin:$PATH
```

4. Convert the Ion Torrent Suite sequencing data to FASTQ format, if it is not already in this format. This can be done using samtools⁶ and the `sam2fastq.py` script:

- a. Converting from SAM format to FASTQ format:

```
<code>samtools view -S \[reads.sam] | sam2fastq.py - > \[reads.fastq]</code>
```

- b. Converting from BAM format to FASTQ format:

```
<code>samtools view \[reads.sam] | sam2fastq.py - > \[reads.fastq]</code>
```

5. Map the FASTQ reads to mouse TA sites using the `fastq2tas.py` script:

```
<code>fastq2tas.py $HOME/nnlab/sbcapseq/srv \[reads.fastq]</code>
```

This will generate a folder containing a file detailing the TA sites (`insertions.txt`) and a log file giving a breakdown of reads that made mapped to TA sites (`log.json`).

6. Convert the TA sites to `BED` format, and mask potential hotspots, using the following command:

```
<code>tas2bed.py -f \[insertions.txt] | maskbed.py $HOME/nnlab/sbcapseq/srv/protonMaskSites.bed - > \[inse
```

- 
7. The **BED** file format is suitable for several types of analysis (e.g. common insertion site statistical analysis) and may be used to visualize SB insertions in a genomics browser (e.g. UCSC Genome Browser, Integrative Genomics Viewer, or Integrated Genome Browser).

Timing

The entire procedure can be completed in as few as seven days.

Troubleshooting

Manufacturer troubleshooting guidelines are published in product documentation supplied with reagents and equipment.

Anticipated Results

SBCapSeq in our lab has been successfully used on the following Ion Torrent sequencing system configurations (chip loading method/sequencing platform/chip version): Chef/Proton/PI™ v3, Chef™/PGM™/318™ v2 BC, OneTouch™ 2/Proton™/PI™ v2, and OneTouch™ 2/PGM™/318™. For Step J — Sequencing of SBCapSeq libraries, select an appropriate Ion Torrent sequencing system configuration and follow the manufacturer's instructions.

The authors have successfully used the SBCapSeq method and bioinformatic workflow to generate SB insertion site data from over 200 DNA libraries from a diverse set of SB induced tumor genomes, including Myeloid Leukemia (ML) and several solid tumors, for sequencing on the Ion Torrent semiconductor sequencing platforms. DNA libraries have been successfully created from various tumor and non-tumor specimens and from populations of bulk tumor cells and matched single tumor cells. Using the semi-quantitative feature of this method, we were able to unambiguously identify genes mutated by SB that drive tumorigenesis and characterize tumor heterogeneity at an unprecedented level. We applied this methodology to analyze a completely penetrant SB-driven model of myeloid leukemia that recapitulates many of the molecular attributes of human ML (see Associated Publications reference for Mann *et al.*, 2016). Our results show that the unbiased, semi-quantitative, and genome-wide capability for sequencing disease causing SB insertions in our mouse

model of AML using the SBCapSeq method offers the most comprehensive view of single cell tumor heterogeneity to date.

Representative Bioanalyzer traces of a representative library from an SB-driven mouse model of acute myeloid leukemia carried through the various step in the protocol is shown in "Figure

2":http://www.nature.com/protocolexchange/system/uploads/4241/original/Figure_2.jpg?1454608225.

A representative PI™ chip run report pertaining to the metrics for sequence data after aligning

sequenced reads to a reference genome (for mouse the default is mm10) is presented in "Figure

3":http://www.nature.com/protocolexchange/system/uploads/4243/original/Figure_3.jpg?1454608227.

References

1. Copeland, NG. & Jenkins, NA. Harnessing transposons for cancer gene discovery. *Nat. Rev. Cancer.* 10, 696-706 (2010).
2. Mann, KM. *et al.* *Sleeping Beauty* mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma. *Proc. Natl. Acad. Sci. USA* 109, 5934-5941 (2012).
3. Mann, MB. *et al.* *Sleeping Beauty* mutagenesis: exploiting forward genetic screens for cancer gene discovery. *Curr. Opin. Genet. Dev.* 24, 16-22 (2014).
4. Leung ML. *et al.* SNES: single nucleus exome sequencing. *Genome Biol.* 25, 16:55 (2015).
5. Wang Y. *et al.* Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature.* 512, 155-60 (2014).
6. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 25, 2078-2079 (2009).

Acknowledgements

The authors wish to thank Rami Zahr (Integrated DNA Technologies, Inc.) for assistance with capture and blocking probe design and Michael Black (Otago) and Cristin Print (Auckland) for valuable discussions. This work was supported by the Cancer Prevention Research Institute of Texas (N.G.C. and N.A.J.). N.G.C. and N.A.J. are CPRIT Scholars in Cancer Research.

Figures

A. Genomic DNA extraction

Purify DNA ↓ Quantitation by NanoDrop/Qubit

B. Library preparation

Purify DNA ↓

C. Quality control check and purification of bar-coded DNA library

↓ Bioanalyzer 2100 analysis

D. DNA library size selection

Purify DNA ↓

E. PCR enrichment of DNA libraries after size selection

Purify DNA ↓ Quantitation by Qubit

F. Liquid capture hybridization

↓

G. Post-capture PCR enrichment

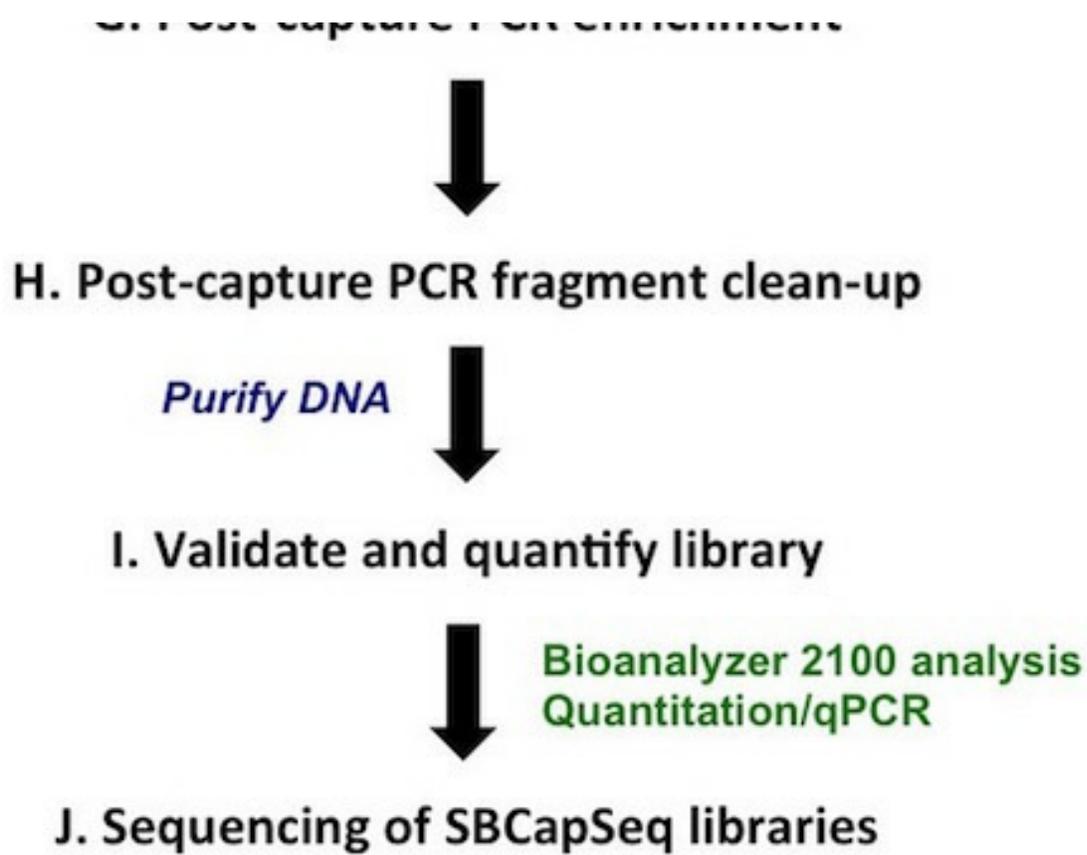
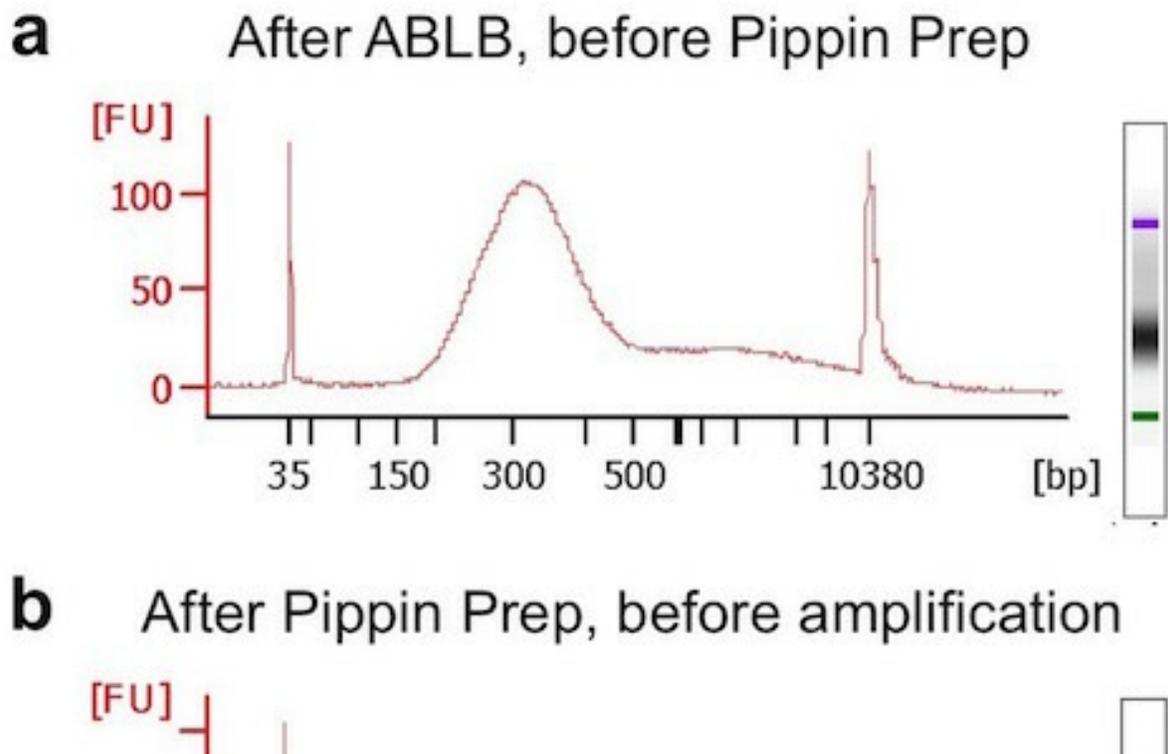
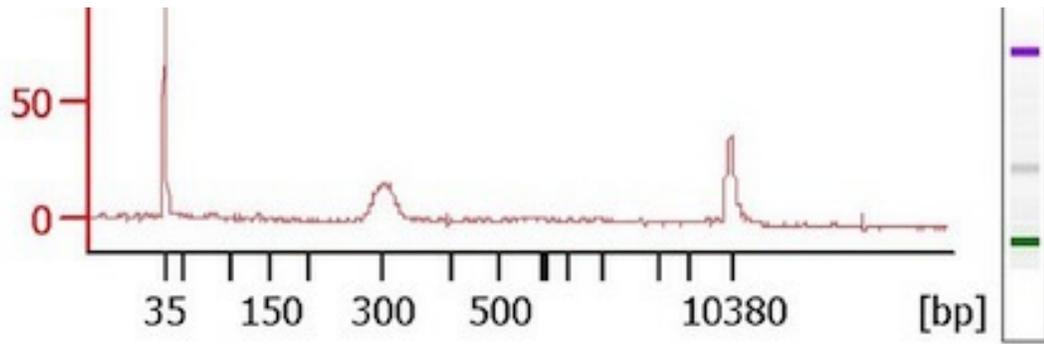


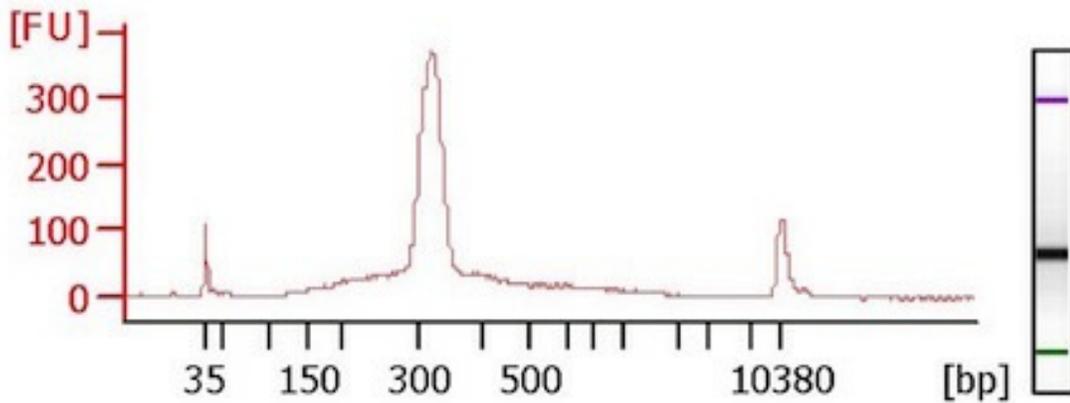
Figure 1

Overview of SBCapSeq experimental workflow. See main text for detailed protocol.





c After amplification, before SB Capture



d After SB Capture, before sequencing

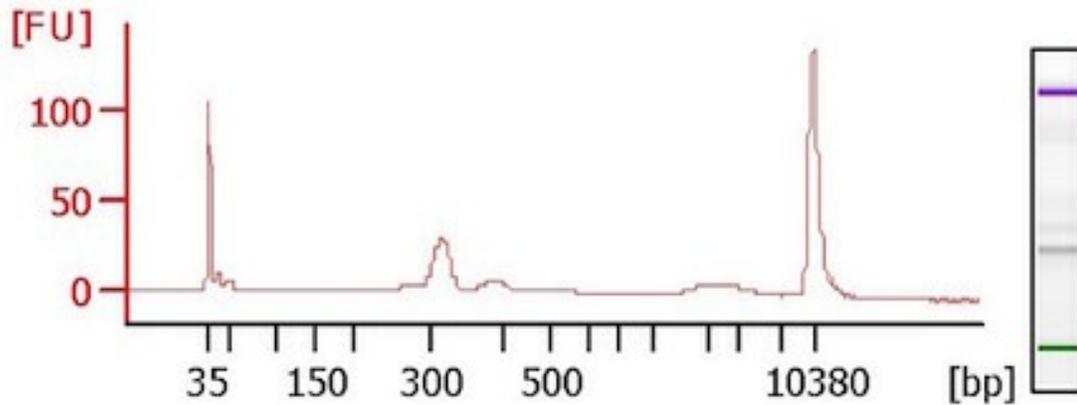


Figure 2

Analysis of DNA library fragment sizes at various steps in SBCapSeq Protocol. Agilent 2100 Bioanalyzer traces of DNA libraries at (a) Step C.1, (b) Step D.5.d, (c) Step E.2.e, and (d) Step F.8.e.

Aligned to Mus musculus 10

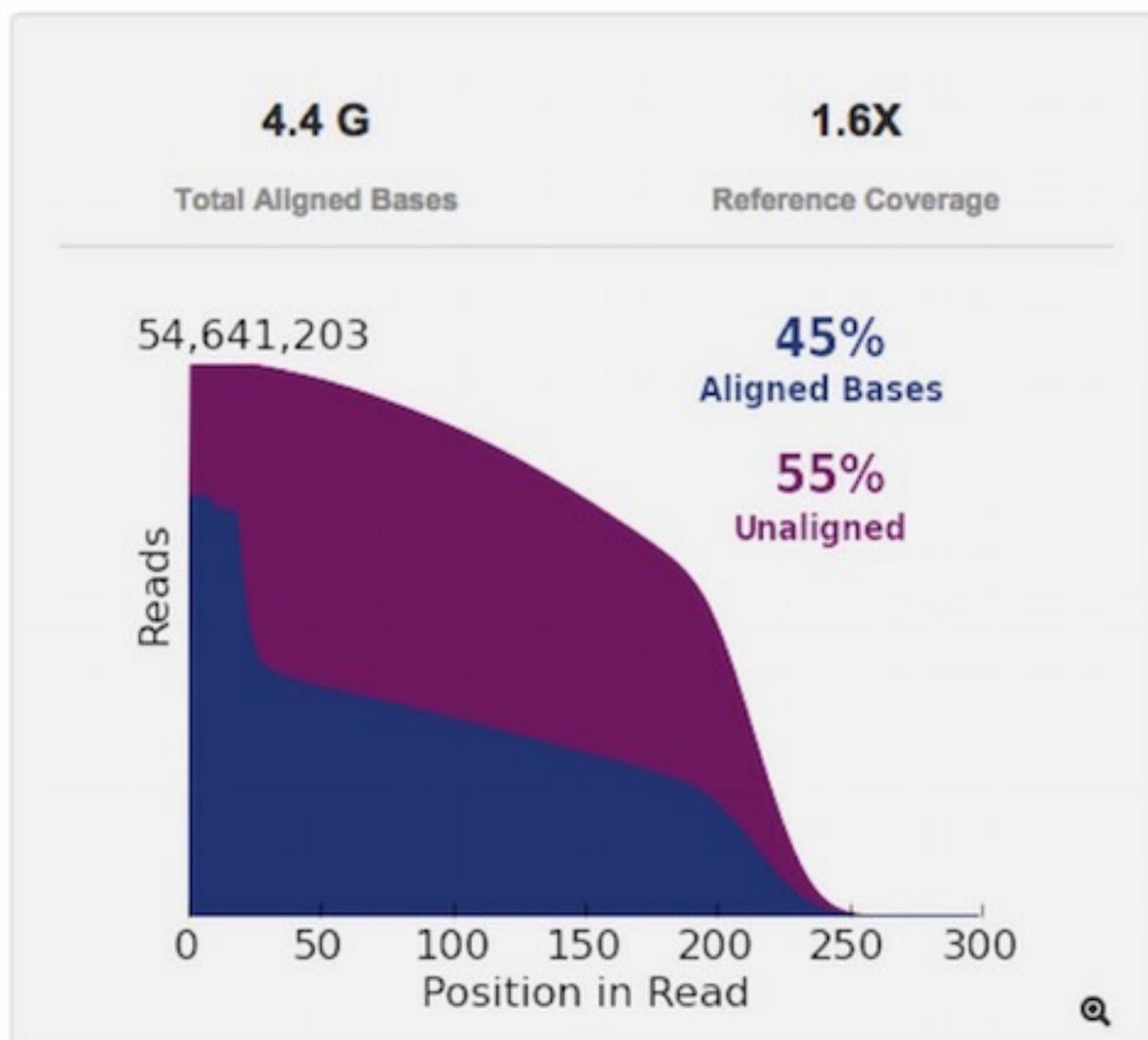


Figure 3

Accumulation of non-reference genome sequencing reads from a representative SBCapSeq experiment. Torrent Suite™ Software run report for PI™ chip pertaining to the metrics for sequence data after aligning sequenced reads to mouse mm10 reference genome at Step

J.5.

Analyzing tumor heterogeneity and driver genes in single myeloid leukemia cells with SBCapSeq

by Karen M Mann, Justin Y Newberg, Michael A Black, +12
Nature Biotechnology (03 August, 2016)

