Preparing sequencing libraries for 2bRAD genotyping.
Last updated: August 25, 2016

Overview

- Preparation of sequencing libraries for 2b-RAD genotyping consists of treating genomic DNA in a series of enzymatic reactions. All reactions are conducted in small volumes in PCR strip tubes or plates.
- Genomic DNA is first digested with a type IIb restriction enzyme, producing fragments with uniform lengths and degenerate sticky ends.
- Next, adaptors are ligated to the cohesive ends generated by restriction digest.
- Finally, sequencing constructs are amplified for a small number of PCR cycles to incorporate sample-specific barcodes.

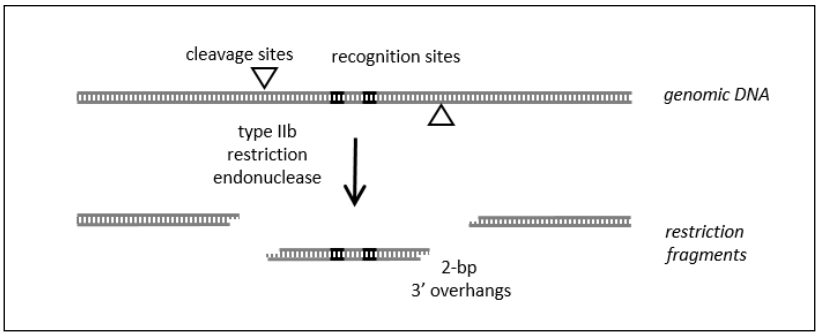
Starting material

1. The protocol begins with 1 µg of purified, high-molecular weight genomic DNA.
2. Samples should be quantified* and visualized on a gel prior to beginning library prep. For high-throughput projects, quantify all samples and check a subset on a gel.
3. Concentrations ≥125 ng µl are required. If necessary, samples can be concentrated prior to library preparation by ethanol precipitation or by drying under vacuum. If the DNA is completely dried during this process (e.g. as in ethanol precipitation), be sure to dissolve the pellet overnight at 4°C before proceeding.

** Note: if A260 is used for quantifying DNA, be sure to include RNAse treatment in your extraction to avoid contaminating RNA. If DNA-specific assays are used, RNAse may be omitted.*

Digest

In this step, DNA is digested with a type IIb restriction endonuclease to produce restriction fragments with degenerate sticky ends.



1. Transfer each sample into a PCR strip tube or plate, 1 µg of genomic DNA at ≥125 ng μl^{-1}.

2. Add nuclease free water (NFW) to each sample for a final volume of 8 µl.

3. Prepare a digestion master mix. The following recipe is for a single reaction, so multiply by the number of samples plus some small amount for pipetting error.

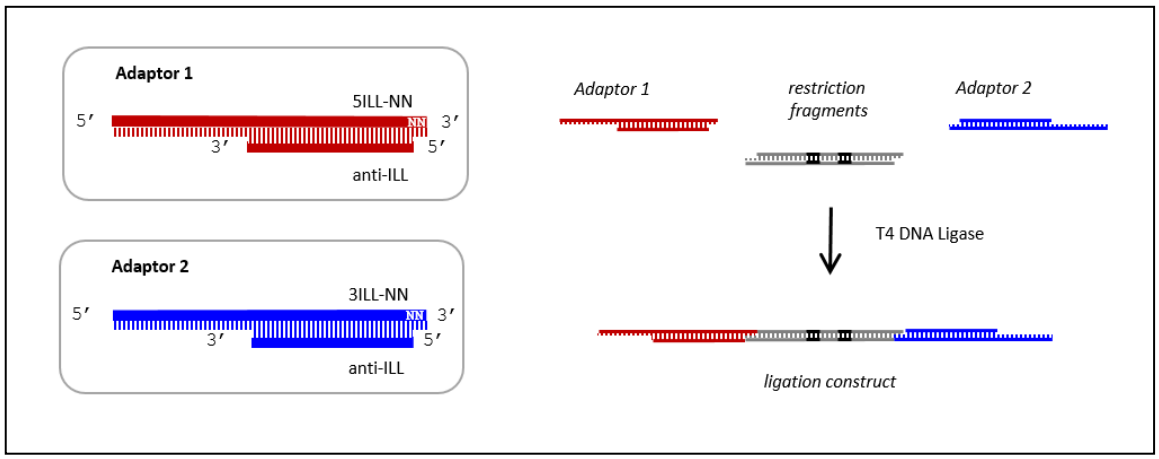
10X buffer R	1.2 µl	
150 µM SAM	0.8 µl	
AlfI (2 U μl^{-1})	0.5 µl	(Thermo Scientific # ER1801)
NFW	1.5 µl	

4. Add 4 µl master mix to each sample.

5. Incubate at 37°C at least 1 and up to 16 hr. If using a heat-inactivatable enzyme (e.g. AlfI), inactivate the enzyme at 65°C for 15 min then transfer samples to ice.

Ligation

In this step adaptors are ligated to the restriction fragments produced above. If reduced representation libraries are intended (targeting a fraction of the restriction fragments), selective adaptors must be chosen at this stage.



1. Prepare a pair of double-stranded adaptors fresh for each set of preps by combining complementary oligonucleotides from stock solutions.
    a. To prepare adaptor 1, choose a 5ILL oligo with the appropriate end. e.g. 5ILL-NN for full libraries targeting all sites, or 5ILL-NG for reduced libraries targeting 1/16th of sites.
    b. Prepare the double-stranded adaptor 1 by combining your chosen 5ILL oligo with anti-ILL, each at 2 µM final concentration, by diluting in NFW.
    c. Repeat steps a-b to prepare an equal volume of the double stranded adaptor 2 (also at 2 µM) using 3ILL oligo and anti-ILL.
    d. Incubate each double stranded adaptor mixture at room temperature for at least 10 minutes.

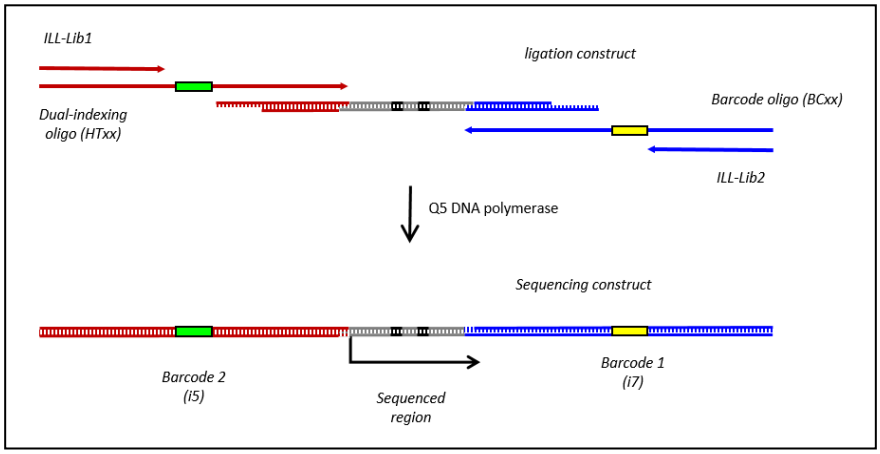
2. Prepare the following master mix for ligations. This recipe is for a single reaction, so scale up as needed.

10 mM ATP	1.0 µl	
10X T4 ligase buffer	4.0 µl	
Adaptor 1 (2 µM)	5.0 µl	
Adaptor 2 (2 µM)	5.0 µl	
T4 DNA ligase	1.0 µl	(New England Biolabs, M0202L)
BSA (20 mg ml^{-1})	0.5 µl	(New England Biolabs, B9000S)
Nuclease-free water (NFW)	23.5 µl	

a. Add 40 µl ligation master mix to each ligation product. Incubate at 16°C for 1 to 16 hr, then hold on ice.

Amplification

In this step, ligation constructs are amplified using a set of four primers that introduce sample-specific barcodes and the Illumina specific sequencing and cluster amplification primer binding sites.



1. Prepare a test-scale PCR to determine optimum cycle number and evaluate relative yield across samples.

a. For each reaction prepare the following master mix. The following recipe is for a single reaction, so multiply by the number of samples plus some small amount for pipetting error.

NFW	5.5 µl	
10 mM (each) dNTP	0.5 µl	
10 µM ILL-Lib1	0.5 µl	
10 µM ILL-Lib2	0.5 µl	
1 µM ILL-HT	1.25 µl	
1 µM ILL-BC	1.25 µl	
BSA (20 mg ml^{-1})	0.25 µl	(New England Biolabs, B9000S)
5X Q5 buffer	5.0 µl	
Q5 DNA polymerase	0.25 µl	(New England Biolabs, M0491L)

b. Combine 15 µl master mix with 10 µl ligation. Amplify on the following profile:

(98°C 5 sec, 60°C 20 sec, 72°C 10 sec) X *N* cycles

c. Sample 5 µl from each reaction at N=10, 15, 20, & 25 cycles. Analyze these products on a 2% agarose gel* with a low molecular weight marker to confirm molecular weight of PCR product.

Note – additional cycles may be required depending on the reduction scheme and quality of the input DNA. As a rule of thumb consider adding 1 cycle per 2-fold reduction (e.g. +2 cycles for a ¼ reduction library).

**See note at the end of document about gel electrophoresis.*

d. Select the minimum number of cycles required to produce a visible product at 166 bp. Note that primer dimers are likely to be visible at ~70-90 and ~130 bp, so be sure to run the gel until these sizes can be resolved.

Note – Let's stop for a moment and acknowledge that gel electrophoresis is not terribly precise. Don't try to estimate precise molecular weights from this gel. Instead, compare PCR products with the ladder to identify where the target construct and any artifacts appear in your own gel conditions.
The correct product will consistently run well above 150 bp, and two artifacts may appear below this (one just below 100 and one about halfway between 100 and 150).

2. Run a preparatory-scale PCR based on the cycle numbers determined above.
   a. Prepare the following master mix for each sample:

NFW	22.0 µl
10 mM dNTP	2.0 µl
10 µM ILL-Lib1	2.0 µl
10 µM ILL-Lib2	2.0 µl
5X Q5 buffer	20.0 µl
BSA (20 mg ml^{-1})	1.0 µl (New England Biolabs, B9000S)
Q5 DNA polymerase	1.0 µl (New England Biolabs, M0491L)

   b. Plan and record barcode assignments. Each sample has to be assigned a unique combination of "BC" and "HT" oligos. BC oligos introduce index 1 ("i7") and HT oligos introduce index 2 ("i5").
   c. Using a single tip, aliquot 50 µl master mix into each well of a PCR plate or strip.

d. Using one tip per oligo, add 5 µl of the appropriate "BC" oligonucleotide (1 µM) to each well.
e. Using a new tip for each sample, add 40 µl of each ligation product to the appropriate well or tube.
f. Using a new tip for each reaction, add 5 µl of the appropriate HT oligo to each well.
g. Amplify using the minimum cycle number determined above.

Gel purification

In this step the sequencing constructs are gel-purified to eliminate residual genomic DNA and primer dimers.

1. Prepare a diluted loading buffer to minimize dye carryover during gel purification. Dilute any standard DNA loading buffer (e.g. NEB B7021S) 1:10 using a 50% glycerol solution prepared with NFW.

2. Prepare a 2% agarose gel*. Use a thick comb that can accommodate 120 µl volumes, or tape together two wells if required.

   **See note at the end of document about gel electrophoresis.*

3. Combine each 100 µl PCR product with 20 µl of diluted loading buffer, load entire mixture into the gel, and run the gel long enough to separate target from non target bands (i.e. ~170 vs ~130 bp).

4. View the gel briefly (<30 seconds) on a UV transilluminator set at low intensity to verify the presence of target bands and adequate separation of molecular weight standards to resolve bands in the 50 to 200 bp range. Typically ~5 cm run distance (well to dye front) is sufficient.

5. Cut out the target band (~170 bp) in a narrow gel slice, being careful to avoid any primer dimers (70-90 and ~130 bp) that may be present.

Note: at this stage a commercial gel-extraction kit can be used instead of the following steps.

6. Transfer each gel slice into a 2 ml microcentrifuge tube and add 40 µl NFW.

7. Centrifuge tubes 1 min at high speed to bring gel into contact with the water. Optional: incubate at 4°C overnight to maximize yield. Otherwise, proceed directly to freezer.

8. Freeze at -80°C for at least 30 minutes.

9. Centrifuge at maximum speed for 10 minutes.

10. Press gel slice against side of tube using pipette tip, and withdraw supernatant (at least 60 µl should be recovered). If less than 60 µl is accessible, repeat centrifugation. Transfer the supernatant to a new PCR tube or plate.

Combining libraries for sequencing

Finally, the libraries are pooled in equal ratios in an effort to sequence all samples at equal coverage.

1. Prepare a 1:100 dilution of each library by combining 2 µl of the eluted library (above) with 198 µl NFW. Typically this is done using a multichannel pipette and micro wellplates.

2. To quantify each library using qPCR, prepare a master mix. The following recipe is for one sample, so adjust accordingly. (In experienced hands 1-2 reactions per sample are sufficient, but users new to qPCR may benefit from additional replication).

NFW	4.3 µl
SYBR qPCR master mix, 2X	7.5 µl (e.g. Bioline #BIO-92005)
ILL-Lib1, 10 µM	0.6 µl
ILL-Lib2, 10 µM	0.6 µl

Note: Other PCR master mixes can obviously be substituted here. The important thing is to use a master mix compatible with your PCR instrument; this reaction is shown only as an example.

3. Pipette 13 µl qPCR master mix into each well of a PCR plate, then add 2 µl of each diluted library to the appropriate well.

4. Conduct qPCR and calculate C_T for each sample.

5. To determine volumes of each library for the combined pool:

    a. Rank samples from lowest to highest C_T and identify reference sample (sample with highest C_T)
    b. Calculate proportion of each library to sequence as:
        $P_L = 2 \text{ ^ } [C_T \text{ (sample)} - C_T \text{ (reference)}]$
    c. Calculate the volume of each library to use as:
        $V = P_L * 60$ µl

Note: If you've chosen a reference sample with a very high CT (suggesting a failed library prep) relative to the others, very low volumes (<2 µl) may be calculated at this step. If so, choose the next sample (i.e. next lowest CT) as reference instead, and continue adjusting choice of reference until reasonably high volumes are calculated for the majority of samples.

Note: As a rough rule of thumb, the pool of combined samples used for a single lane of sequencing should be <u>at least</u> 200-500 µl at this stage (it may be substantially higher).

6. Combine libraries using the volumes calculated from qPCR to produce a pool for sequencing.

7. Illumina sequencing typically requires templates in ≤20 µl volume at ≥2 nM concentration. The pooled libraries produced above are typically too dilute for sequencing. To purify and concentrate these libraries, they may be precipitated with isopropanol (below) or using a commercial PCR cleanup kit.

8. To precipitate:
    a. add 0.1 volumes of 3M sodium acetate and 3 volumes 100% isopropanol. Incubate 30 minutes at -20°C.
    b. Centrifuge 20 minutes at maximum speed, 4°C.
    c. Dry pellet 15' at room temperature and resuspend in 25 µl NFW.

9. (Optional) Quantify your pooled library using a method sufficiently precise and sensitive to quantify concentrations in the 1-10 ng μl^{-1} range. You need a concentration of **at least 0.65 ng μl^{-1}** at this stage to have enough for sequencing.

Oligonucleotide sequences

Illumina HiSeq

Name	Sequence (5' – 3')
anti-ILL	AGATCGGAAGAGC(InvdT)
5ILL-NN	CTACACGACGCTCTTCCGATCTNN
3ILL-NN	CAGACGTGTGCTCTTCCGATCTNN
5ILL-NG	CTACACGACGCTCTTCCGATCTNG
3ILL-NG	CAGACGTGTGCTCTTCCGATCTNG
ILL-HT	AATGATACGGCGACCACCGAGATCTACAC**[BC2]**ACACTCTTTCCCTACACGACGCTCTTCCGATCT
ILL-BC	CAAGCAGAAGACGGCATACGAGAT**[BC1]**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC
ILL-Lib1	AATGATACGGCGACCACCGA
ILL-Lib2	CAAGCAGAAGACGGCATACGA

InvdT: inverted T to prevent extension by DNA polymerase that may lead to tandem ligations. BC1: barcode 1 (called "index 1" or "i7" by Illumina); may be any one of the standard 6-bp TruSeq barcodes or other custom 5- to 8-bp barcodes. BC2: barcode 2 (called "index 2" or "i5" by Illumina); custom 5- to 8-bp barcodes.

Example schemes for reduced tag representation.

Fraction of sites targeted	5' oligo	3' oligo
1	-NN	-NN
1/4	-NR	-NR
1/8	-NC	-NG
1/16	-NC	-NC
1/32	-RG	-YG
1/64	-RG	-RG

Gel electrophoresis in 2bRAD library preparation

We use a sodium borate buffer with added EDTA (SBE). This buffer is cheap and can be run at higher voltages with better resolution than the usual Tris based buffers. You can make the 20X buffer as shown at http://openwetware.org/wiki/SB. We add EDTA to 20 mM (in the 20X solution).

We also use ethidium bromide at 0.1 µg ml^{-1}, but other dyes with comparable sensitivity can be used instead. EtBr is very cheap and sensitive, but somewhat toxic and must be illuminated with UV light which damages DNA, so some users prefer to use more expensive dyes that can be viewed with blue light.

Generally, there is a lot of flexibility in the details of gel electrophoresis but users should strive to maximize resolution and minimize damage to the DNA.