2bRAD sample preparation

Galina Aglyamova, Mikhail Matz (matz@utexas.edu)
version June 2019

Overview

This is a modification of the protocol described in Wang et al, Nature Methods 2012 2b-RAD paper (doi:10.1038/nmeth.2023) re-designed for Illumina HiSeq and BcgI enzyme. BcgI is a relatively frequent cutter, with 75-100k sites per genome, and in our experience is one of the most efficient of the IIb-type restriction enzymes. It produces 36-base fragments with two-base overhangs and can be inactivated by heating, both of which facilitate ligation of adaptors. Another enzyme that can be used with this protocol, either on its own or together with BcgI, is AlfI.

The latest modifications include the use of barcoded ligated adaptors so the samples can be pooled by 12 after ligation, and use of degenerate 5'-adaptor that makes it possible to remove PCR duplicates.

The protocol involves the following steps:

1. Restriction digest. Genomic DNA is digested with a type IIB restriction enzyme to produce restriction fragments of uniform length.
2. Ligation. Adaptors, barcoded for 3' end and generic for 5', are ligated to the cohesive ends generated by restriction digest.
3. Heat-inactivate the ligase and pool the ligations with different 3' barcodes into groups of 12.
4. Amplification and barcoding of pooled ligations.
5. Purification of the target 180 bp band by Pippin or gel-electrophoresis (the only purification step in the whole procedure).

NOTE on experimental design: we strongly recommend including three pairs of genotyping replicates in each 2bRAD experiment. Replicates should be independently prepped DNA samples from the same individual. ***They will be extremely useful at the analysis step to assess the overall accuracy of genotyping, set quality filtering criteria, and quantify relatedness between samples.***

Materials

Reagent	Vendor/Cat #
BcgI	NEB R0545S
(optional) AlfI	Thermo Scientific ER1801
NEBuffer #3	NEB B7003S
T4 DNA ligase	NEB M0202L
Titanium Taq, with buffer	Clontech 639208
dNTPs	NEB/N0447S
Eppendorf twin-tec 96-well PCR plate	Eppendorf (VWR:95041-440):
Adhesive PCR plate Foil Seal	Fisher/UT Mkt: AB-0626
Ligated adaptor oligos * #	IDTDNA
Barcode (index) oligos * &	IDTDNA

* see **2bRAD_oligos_order.xlsx** file; it lists oligos sufficient for a 576-sample experiment. If you have more samples per project, order more TruSeq_Un oligos with additional indices (substitute GCCAAT, CAGATC, GATCAG, or CTTGTA for the sequence written in red font). Each additional TruSeq_Un oligo adds capacity for 96 additional samples.

Prepare ligated pseudo-double stranded adaptors by combining corresponding pairs of oligos.

- For **Adaptor 1**, mix 5ILL-NNRW (10 μM) with the same volume of 10 μM Anti5ill-NNRW.
- For **Adaptor2** (versions 1 to 12), combine 10 μM 3illBC(1-12) with the same volume of corresponding 10 μM antiBC.
- Incubate at 42°C for 5 minutes then keep at room temperature until ligation.

& Dilute ILL-BC and Truseq_UN oligos to 1 μM in 10mM Tris-HCl, pH 8.0; store at -20°C for a few months. Each sample would require 0.25 μl of each of these stocks.

Digest

1. Dilute samples to contain the **same amount of DNA (50-100 ng) in 4 µl**. The DNA has to be high purity, with OD 260/230 ratio >1.8. it is very important to equalize your input samples well, based on Qbit or picogreen (not just nanodrop). Concentrating can be accomplished by drying under vacuum. Pipet samples into 96-well plate(s).

2. Prepare a digestion master mix:

Number of samples x 1.1 =

Reagent	per sample	Volume in master mix
NEB Buffer #3	0.6	
SAM (included with BcgI)	0.4	
H_2O	0.5	
BcgI (1 U μl^{-1})	0.5	

3. Pipet master mix into 8-tube strip (26 µl per tube per sample plate). Use multi-channel pipette to **combine 2 µl master mix with each 4 µl DNA sample** (6 µl total volume). Cover the plate with PCR film, incubate at 37°C (in air incubator or in a thermocycler with heated lid) for 1 hr.

4. **Inactivate the enzyme at 65°C for 10 min** then let the thermocycler return to room temp while keeping the heated lid on. Hold samples on ice after this.

Ligation

1. Prepare 12 master mixes for ligations (one for each of the twelve Adaptor 2), in strip-tubes.

Number of rows in all 96-well sample plates x 1.1 =

Reagent	per sample	Volume in master mix
10mM ATP	0.5	
10x T4 ligase buffer)	2	
Adaptor 1	1	
Adaptor 2 (1 through 12)	1	
H_2O	14.5	
T4 DNA ligase	1	

NOTE: this is the stage at which reduced tag representation (RTR) can be achieved by using modified 5ILL-NNRW and 3illBC oligos, not with NN but with NG bases at their 3'-ends.

2. Use multichannel pipette to combine **20 µl** master mix with digested DNA, **each master mix going into a specific column** on the plate. Mix by pipetting. Keep the plate on ice while mixing.

3. Incubate at **16°C for at least 2 hours** and up to overnight.

4. Heat at **65°C for 20 min** to inactivate the ligase, in a thermocycler with heated lid.

5. Pool ligations by row in strip-tubes, using **3 µl from each well.** The 96-well plate is now reduced to 8 pooled samples, each corresponding to the original row.

Amplification and barcoding

1. Pipet 4 µl of pooled ligations into wells of a new 96-well plate. **Each column in this "pooled ligations plate" now corresponds to one whole original 96-well sample plate.**

2. Pipet the eight **1 µM ILL-BC oligos** into a strip tube. Use multi-channel pipette to add 3 µl of them to the plate with pooled ligations (**each oligo to its own row**).

3. Pipet as many **1 µM TruSeq_UN oligos** as there are columns in the pooled plate into a strip tube. Use multi-channel pipette to add 3 µl of them to the pooled ligations plate (**each oligo to its own column**).

4. Prepare the following master mix:

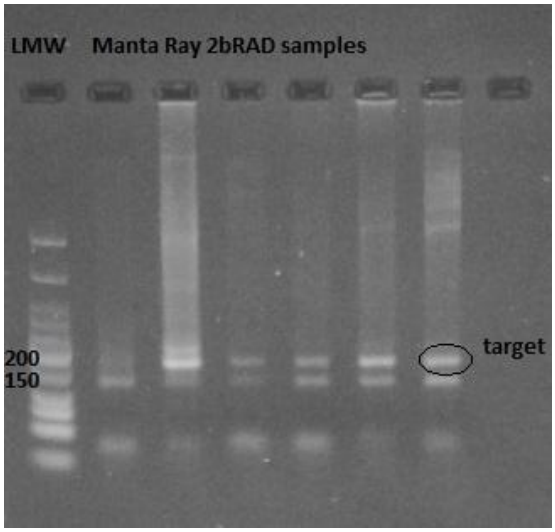
Number of pooled samples x 1.1 =

Reagent	per sample	Volume in master mix
dNTPs 2.5 mM	2.5	
H_2O	3.7	
10 µM IC1-P5	0.4	
10 µM IC1-P7	0.4	
10x Titanium buffer	2	
Titanium Taq	1	

5. Aliquot the master mix into strip-tubes, then use multi-channel to add 10 µl to each pooled ligaton.

6. Amplify as follows:
   70°C 30 sec then (95°C 20 sec, 65°C 3 min, 72°C 30 sec) X 15 cycles

7. For a random subset of 4-6 samples. load 5 µl on a 2% agarose gel alongside LMW ladder (NEB N3233S) or other marker that has 150 and 200 bp bands.



NB: There should be <u>no substantial variation in product amount</u> among samples at this stage (unlike this gel photo); if there is, something went wrong with DNA quantity/quality of some original samples!

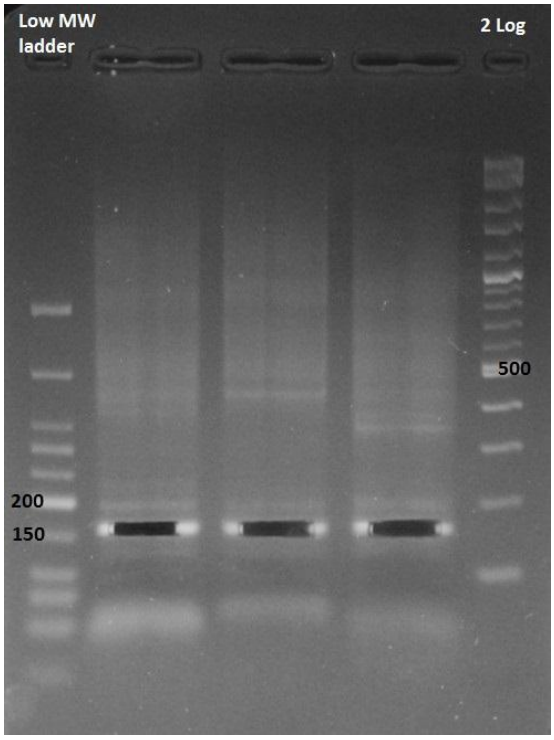
Confirm that all samples have a **visible band at ~180 bp**. You might also see a band below 150 bp, which is an artifact from the carried-over ligase (make sure to heat-inactivate the ligase before mixing the PCR reaction!)

If the 180 b product is visible but barely, add two more cycles to the same reactions.

If the 180 b product is not visible at all after 15 cycles, do not add more cycles. Even though weak samples might be possible to amplify with more PCR cycles, such samples would have poor representation of allelic diversity (will show lower counts after removal of PCR duplicates and will often miss a second allele in a heterozygote). To troubleshoot, optimize previous steps. Three common problems are poor DNA quality (make sure OD 260/230 is 2), incorrect DNA quantification in presence of RNA, and expired enzymes (ligase and restriction enzyme).

If all is well (all samples have visible 180 bp band of similar intensity), **pool all reactions together** (combine 3 µl from each reaction, vacuum-concentrate to 60 µl) and **extract the target 180 bp band** with PippinPrep or BluePippin (170-190 bp range). Most sequencing core facilities would do it for you. If not, gel-purify the band as described below.

Gel-purification



1. Prepare a 2% agarose gel using TBE or TB. Use a wide comb that can accommodate 30-50 µl, or simply tape together two wells.

2. Load 30-50 µl of sample (40 µl sample + 10 µl loading dye) alongside LMW ladder. Run gel at low voltage for a long time until bands at 150bp and 200bp will clearly resolved. In our experience 100V for 70 minutes produce the good separation.

3. View the gel briefly (<20 seconds) on a UV or (possibly longer) on appropriate for your DNA dye black-light transilluminator set at low intensity to verify the presence of target band and adequate separation of molecular weight standards to resolve bands at ~180 bp and (possibly) below 150 bp. Typically ~5 cm run distance is sufficient. Photograph.

4. Cut out target ~180 bp band in a narrow gel slice, avoiding the edges of the lane (i.e., cut out the middle 70-75% of the band). Cut just inside the bottom boundary of the target band to avoid getting anything smaller.

***Note: at this stage a commercial gel-extraction kit can be substituted for the following three steps, if you feel more confident this way. In our practice, simply soaking the gel slice in water overnight, as described below, works just fine.**

5. Transfer each gel slice into a 1.5 ml microcentrifuge tube and add 20 µl NFW.

6. Make sure gel slice is in contact with water (cut or break it into a few, say 4-5, smaller pieces to make sure they sit comfortably at the bottom of the tube). Hold overnight at 4°C.

7. The following day transfer the eluate (~15 µl) into a new tube. This material is now ready for sequencing. Optimal sequencing depth is 2-3M 50xSE reads per sample (10 times less for the RTR version). Sequence on HiSeq 2500 if you can; on HS 4000 add 20% phiX.

Oligonucleotide sequences for Illumina HiSeq

The repository includes the bulk-order excel file called **2bRAD_oligo_order.xlsx**, which can be directly uploaded to IDTDNA (the total cost as of January 2019 was $1845.36). These oligos can barcode 6 x 96 samples and should be sufficient for hundreds of 2bRAD projects in term of quantity. If you have more samples per project, order more TruSeq_Un oligos with additional indices (substitute TGACCA, ACAGTG, GCCAAT, CAGATC, GATCAG, or CTTGTA for the sequence written in red font). Each additional TruSeq_Un oligo adds capacity for 96 more samples. Note that each of TruSeq_Un oligos costs $97, so, for smaller projects, consider ordering less than six of them listed in the excel file.

For reduced-representation 2bRAD, all the NN in the excel file should be replaced with NG.