Sample preparation for 2b-RAD genotyping

Shi Wang (swang@ouc.edu.cn)
College of Marine Life Sciences, Ocean University of China
Eli Meyer (eli.meyer@science.oregonstate.edu)
Department of Zoology, Oregon State
John McKay
Department of Bioagricultural Sciences and Graduate Degree Program in Ecology, Colorado State University
Mikhail Matz
School of Biological Sciences, University of Texas - Austin

Method Article

Keywords: Genotyping, SNP discovery, Restriction-associated DNA, type IIB restriction endonuclease,

Posted Date: June 6th, 2012

DOI: https://doi.org/10.1038/protex.2012.010

Abstract

High-throughput sequencing of restriction site-associated genomic DNA \(RAD) is now a widely used approach for simultaneously discovering and profiling genetic polymorphisms. We describe an alternative approach for RAD genotyping with several useful features, called 2b-RAD for the use of type IIB restriction endonucleases \(AlfI, BsaXI). The fragments excised by type IIB enzymes are uniform in length, streamlining library preparation and producing even sequencing coverage across the genome. The cohesive ends present on both 5' and 3' ends of each fragment make it possible to target a defined subset of restriction sites using modified adaptors. Marker density can be adjusted as needed in this way, greatly reducing sequencing coverage requirements and costs. 2b-RAD offers a cost-effective strategy for high-throughput genotyping that is widely suitable for linkage mapping and profiling genetic variation in natural populations.

Introduction

Sample preparation for 2b-RAD genotyping is accomplished in a single well of a multi-well plate within about 4 hours. 2b-RAD constructs are produced in a series of three reactions \(restriction digest, adaptor ligation, and barcode incorporation), then gel-purified and sequenced \(Fig 1). First, genomic DNA is digested using type IIB restriction endonucleases to excise the target fragments used for genotyping \(Fig 1). To maintain buffer concentrations in subsequent reactions, the initial digest is conducted in a small volume, which necessitates high DNA concentrations \(at least 250 ng μl^{-1}). Samples with lower concentrations can be easily concentrated using standard methods including ethanol precipitation or drying under vacuum. Before processing a new set of samples \(e.g. a new species, etc.), it's advisable to test the effectiveness of the chosen restriction enzyme by comparing digested genomic DNA \(~100-200 ng genomic DNA) against the uncut DNA and a no-enzyme control. While partially degraded DNA can be used effectively for 2b-RAD genotyping, for the purposes of this test a perfectly intact sample should be chosen. Digested samples show a subtle smear extending downward from the original high-molecular weight genomic DNA band. No-enzyme controls should look identical to the original DNA after 1 hr at 37°C. Next, partially double-stranded adaptors are ligated to the restriction fragments using T4 DNA ligase. The sequences of these adaptors, and all other oligonucleotides used in this protocol, are shown in Table 1. Note that the choice of adaptors at this stage determines the set of restriction sites to be included in the library, and the sequencing platform to be used. The fully degenerate overhangs \(5'-NNN-3') featured in the standard protocol target all sites. Adaptors with reduced degeneracy can be used instead to target a subset of the sites \(e.g. adaptors bearing overhangs 5'-NNG-3' target 1/16th of the sites). The fraction targeted can be chosen based on genome size, polymorphism frequency, and the number of polymorphisms required for a particular experiment. In the final reaction, ligation products are amplified using oligonucleotide primers complementary to the adaptors, and incorporating sample-specific barcodes for multiplex sequencing of pooled samples. This is accomplished using a step-out PCR procedure in which two long oligos present at low concentrations incorporate barcode sequences and binding sites for two shorter primers present at standard concentrations. To minimize any distortions

of allele frequencies, samples are collected across a range of cycle numbers to identify the minimum number of cycles required to produce a visible product in gel electrophoresis. Only a small number of cycles should be required \(4-10) for standard libraries, while for reduced schemes additional cycles proportional to the targets' lower abundance are required \(e.g. 1/16th reductions require an additional ~4 cycles). Finally, the barcoded PCR products are gel purified using standard methods to remove primer dimers and residual products of restriction digest. The purified products are then pooled for multiplex sequencing on either SOLiD or Illumina platforms. Note that the fragments produced by 2b-RAD are shorter than the maximum read length of these technologies, so short read lengths should be specified when submitting samples \(36 bp or less) to minimize sequencing costs. The details given in Procedures correspond to a standard AlfI library to be sequenced on the SOLiD System. A wide variety of modifications are possible, depending on the choice of enzyme, the reduction scheme, and the sequencing platform, as described in the associated manuscript.

Reagents

Major reagents: 1) AlfI \(Fermentas, cat no. ER1801) 2) T4 ligase \(NEB, cat no. M0202) 3) ATP \(NEB, cat no. P0756) 4) Phusion High-Fidelity DNA polymerase \(NEB, cat no. M0530) 5) dNTP \(NEB, cat no. N0447) 6) Agarose 7) All adaptor and primer oligos are ordered from IDT.

Equipment

Major equipments: 1) NanoDrop 1000 Spectrophotometer \(Thermo Scientific) 2) Horizontal Electrophoresis Apparatus Set 3) Gel Documentation System 4) DNA Engine Tetrad 2 thermal cycler \(Bio-Rad)

Procedure

Restriction digest **1 |** Prepare intact, high-quality genomic DNA samples each containing 1 µg at a high concentration \(at least 250 ng μl^{-1}). Dilute all samples to the same volume \(4 µl) with nuclease free water. **2 |** Prepare a master mix for restriction digest by combining the following components in a single tube and mixing thoroughly by back-pipetting. The volumes listed here are intended for a single reaction, so multiply by the number of samples plus some small amount for pipetting error. 0.6 µl 10X Buffer R 0.4 µl 150 µM SAM 1.0 U AlfI Nuclease-free water \(NFW) to total volume 2.0 µl **3 |** Combine 2 µl master mix with each 4 µl DNA sample, and incubate 1 hr at 37°C, then inactivate the enzyme at 65°C for 20 min and hold samples on ice. **Adaptor ligation** **1 |** Prepare two partially double stranded adaptors by combining each pair of oligonucleotides at a final concentration of 4 µM each. Adaptor 1 is prepared by combining sld-ada1-AlfI and anti-sld-ada1, and adaptor 2 by combining sld-ada2-AlfI and anti-sld-ada2 \(Table 1). **2 |** Prepare a master mix for ligation by combining the following components in a single tube. The volumes listed here are intended for a single reaction, so scale up as needed. 0.5 µl 10 mM ATP 2.0 µl 10X T4 ligase buffer 2.5 µl 5 µM Adaptor 1 2.5 µl 5 µM Adaptor 2 11.5 µl NFW **3 |** Combine 20 µl master mix with 5 µl of digested DNA. Incubate for 1 hour \(16°C for heat-inactivatable

enzymes e.g. AlfI, or at 4°C for enzymes that cannot be inactivated, e.g. BsaXI), then hold on ice. **Barcode incorporation** **1 |** Conduct a test-scale PCR to determine the minimum number of cycles required to produce a visible product, and evaluate relative yield across samples. Prepare a master mix for PCR by combining the following components in a single tube, and mix thoroughly by backpipetting. The volumes listed here are intended for a single reaction, so scale up as needed. 6.5 µl NFW 2.5 µl 2.5 mM dNTP 0.4 µl 10 µM sld-p5 0.4 µl 10 µM sld-p6 1.0 µl 1 µM sld-p3 1.0 µl 1 µM sld-p4 \(barcode) 4.0 µl 5X HF buffer 0.2 µl Phusion polymerase **2 |** Combine 16 µl master mix with 4 µl ligation from each sample, and amplify on the following profile: \(98°C 5 sec, 60°C 20 sec, 72°C 10 sec) X 12 cycles. **3 |** Sample 5 µl from each reaction at two-cycle intervals \(n = 6, 8, 10, & 12 cycles). The thermocycler can be paused during each sampling interval. **4 |** Visualize PCR products using standard gel electrophoresis methods \(2% agarose TBE gel), including a low-molecular weight marker. Select the minimum number of cycles required to produce a visible product at ~130 bp. **5 |** Prepare a master mix for preparatory-scale PCR by combining the following components in a single tube, and mix thoroughly by backpipetting. The volumes listed here are intended for a single reaction, so scale up as needed. 32.5 µl NFW 12.5 µl 2.5 mM dNTP 2.0 µl 10 µM sld-p5 2.0 µl 10 µM sld-p6 5.0 µl 1 µM sld-p3 5.0 µl 1 µM sld-p4 \(barcode) 20.0 µl 5X HF buffer 1.0 µl Phusion polymerase **6 |** Combine 80 µl master mix with 20 µl ligation and amplify using the optimal cycle number determined in step 4. **Gel purification** **1 |** Resolve samples on 2% agarose gels using standard electrophoresis methods and reagents. Large wells are required to accommodate each 100 µl sample in a single well. **2 |** Using a UV transilluminator set to low intensity, excise the target band \(~130 bp) from each lane, being careful to limit the UV exposure of each sample to no more than 30 seconds. **3 |** Extract preparations from excised gel slices either: A. using a commercial gel-extraction kit of your choice according to the manufacturers instructions; or B. by eluting directly from the gel slice into water \(40 µl NFW in a 1.5 ml microcentrifuge tube). Gel slices are held overnight at 4°C, and the eluent collected the following morning. **4 |** Collect a small aliquot \(~2-10 µl) from each preparation and combine in a multiplex pool for sequencing. Remaining preparations can be held for 6 months at -20°C and sampled again for sequencing if additional coverage is required for any samples.

Figures

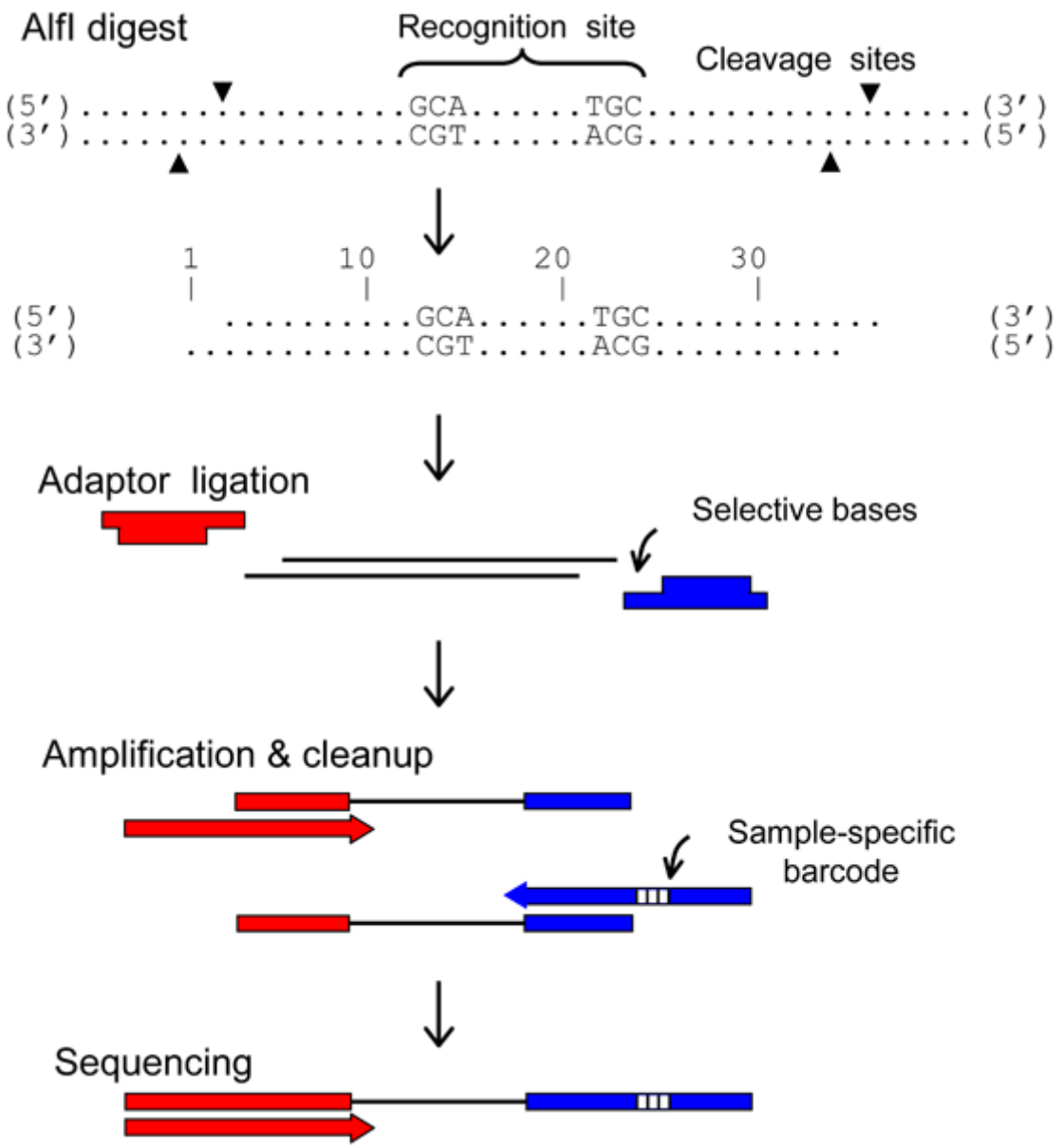

Figure 1

Preparation of samples for 2b-RAD genotyping. Sample preparation for 2b-RAD genotyping is accomplished by restriction digest (AlfI) of genomic DNA, cohesive end ligation of partially double-stranded adaptors with compatible (NN) overhangs, and incorporation of barcodes for multiplex sequencing by PCR.

Name	Sequence (5’-3’)
anti-sld-ada1	ATCACCGACTGCCCAT(invT)
anti-sld-ada2	TTCGCCTTGGCCGT(invT)
sld-ada1-AlfI	CTCTCTATGGGCAGTCGGTGATNN
sld-ada2-AlfI	CTGCTGTACGGCCAAGGCGAANN
sld-3	CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGGTGAT
sld-4	CTGCCCCGGGTTCCTCATTCTCT[bc]CTGCTGTACGGCCAAGGCG
sld-5	CCACTACGCCTCCGCTTTCCTCTCTATG
sld-6	CTGCCCCGGGTTCCTCATTCT

Figure 2

Table 1 Oligonucleotide sequences for 2b-RAD sample preparation bc: barcode, any 6-10 bp DNA sequence easily distinguishable from other barcodes on that same sequencing run.