Conservation Genet Resour (2014) 6:49–51
DOI 10.1007/s12686-013-0073-8

TECHNICAL NOTE

SNP discovery in a reef-forming glass sponge, *Aphrocallistes vastus*, using the Ion Torrent next generation sequencing platform

Rachel R. Brown · Corey S. Davis · Sally P. Leys

Received: 4 October 2013 / Accepted: 14 October 2013 / Published online: 4 December 2013

Abstract Using next generation sequencing (NGS), we located 1,121 single nucleotide polymorphism (SNP) markers in a reef-forming glass sponge, *Aphrocallistes vastus*. 16 SNPs were validated using Sanger sequencing. SNPs developed here are available for use in describing the genetic structure of glass sponge reefs, and demonstrate that NGS methods overcome problems posed by genomic DNA in some invertebrates.

Keywords *Aphrocallistes vastus* · Glass sponge reefs · Single nucleotide polymorphism markers · Next generation sequencing · Ion Torrent

Glass sponge reefs form a globally unique habitat in the Northeast Pacific (Conway et al. 2001). Marine protected areas (MPAs) have been proposed to protect reefs from fishing activity, the choice of which to include will rely in part on knowledge of their genetic diversity. Microsatellites are commonly used population markers, but where genomic duplication exists finding sufficient unduplicated markers de novo is difficult (e.g. Ji and Zhang 2004; Baums et al. 2005). Here we adapt a method of selectively sampling identical genomic regions across several individuals (Poland et al. 2012) to the discovery of single nucleotide polymorphisms (SNPs), a high-density marker, on the Ion Torrent sequencing platform. This protocol allows duplicated loci to be removed from thousands of identified markers while retaining a high number of reliable SNPs.

R. R. Brown (✉) · C. S. Davis · S. P. Leys
Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2E9, Canada
e-mail: rrbrown@ualberta.ca

We collected tissue samples from *Aphrocallistes vastus* (Porifera: Hexactinellida), the dominant reef-building sponge in the Strait of Georgia, at eight discrete locations within four broad sample regions covering 900 km of the Northeast Pacific from 2007 to 2011. We extracted DNA from ten samples (taken from all sample regions) with Qiagen DNeasy spin-columns following the manufacturer's instructions with these modifications: more tissue (approximately 5 mm^3) was digested in twice the reagent volumes (proteinase K, buffers ATL and AL, EtOH, and RNase A); tissue was incubated in lysis buffer overnight at 56 °C and spin-columns were loaded in two steps. DNA was quantified using fluorometry.

A reduced representation library was developed using two restriction enzymes with paired linkers (Fig. 1a). Samples were digested with *PstI* [New England Biolabs (NEB)] in 10 separate 50 µl reactions (300 ng DNA, 5U *PstI*, 1 × BSA, 1 × Buffer 3) for 3 h at 37 °C. Modified Ion Xpress™ Barcode Adapters (Life Technologies) were ligated to 200 ng of *PstI* digested DNA in separate 50 µl reactions (0.25 pmol adapter, 100 U T4 DNA ligase (NEB), 10 pmol rATP). Barcoded samples were pooled, purified with the MinElute PCR Purification kit (Qiagen), and digested with *MspI* (NEB) in a single 50 µl reaction (5U *MspI,* 1 × Buffer 4) for 3 h at 37 °C. Following a second column purification, a reverse Y-adapter (P1Y) including the Ion Torrent P1 primer sequence was ligated to the pooled DNA library in a 50 µl reaction (0.25 pmol adapter, 100 U T4 DNA ligase, 1 × ligase buffer) at room temperature for 30 min. E-Gel SizeSelect Gels (Invitrogen) were used to extract bands at 200, 250, 300, and 350 bp both before and after PCR. Fragments were amplified in a 100 µl reaction using 2U Phusion High Fidelity DNA polymerase (NEB), 200 µM of each dNTP, 1 × Phusion HF buffer, 0.5 µM of forward and reverse

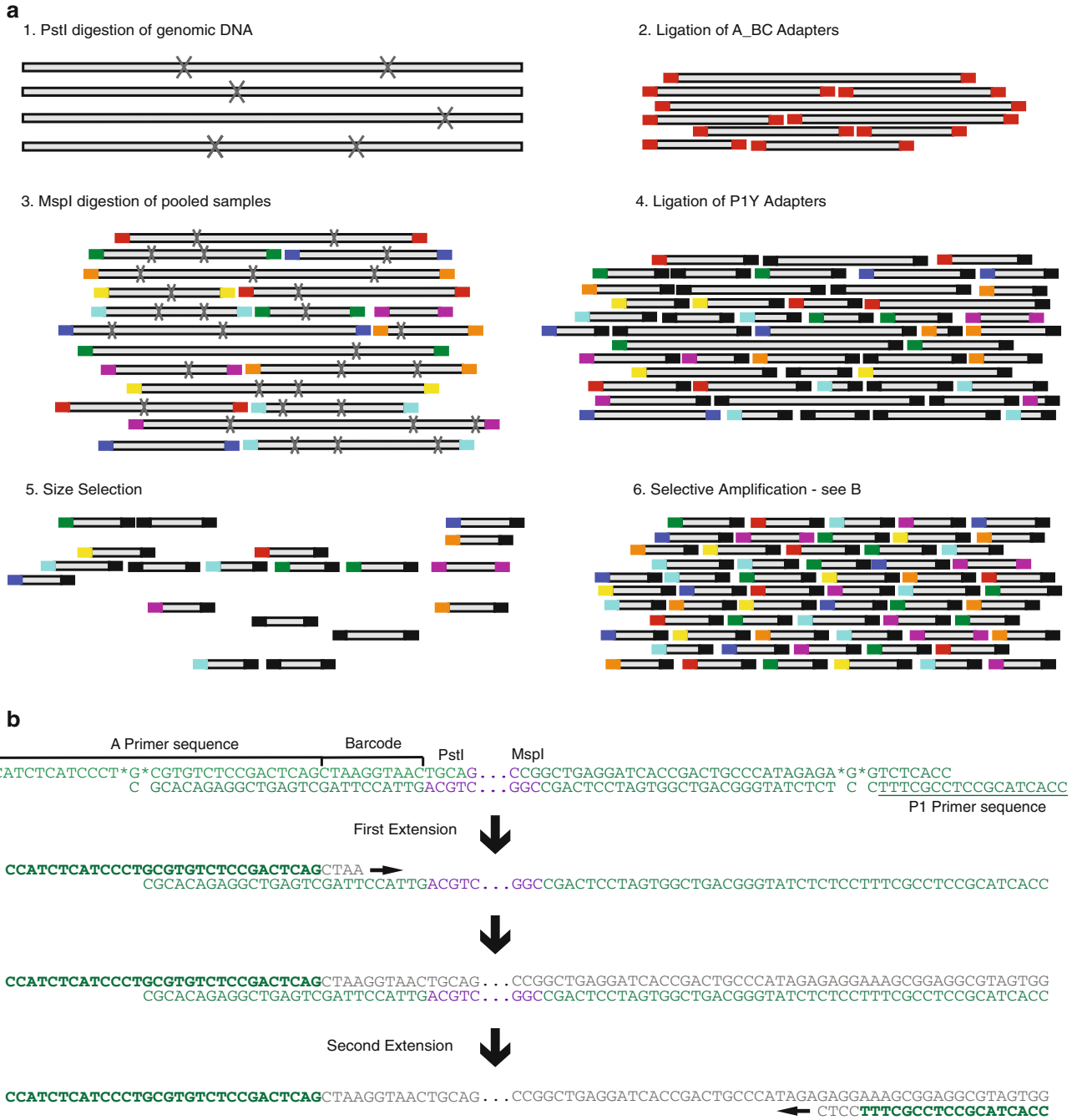

Fig. 1 **a** Reduced Representation Library construction *1* uncut genomic DNA is digested with a rare (6 bp) cutter RE, followed by *2* ligation of barcoded adapters; all barcoded samples are pooled, *3* digested with a common (4 bp) cutter, and *4* ligated to a Y shaped adapter; *5* Size selection is performed prior to *6* selective amplification of fragments with an A_BC adapter. **b** Illustration of the first two PCR cycles (*Primers are bold*): P1 primers do not have a binding site until a complimentary strand is created using the A primer, resulting in amplification of only the fragments which have both A_BC and P1Y adapters, or 2 A_BC adapters (rare—eliminated in the sequencing). 2B adapted from Poland et al. (2012)

primers, and 3 % DMSO, under the following conditions: 96 °C for 5 min, 30 cycles of 96 °C (45 s), 60 °C (60 s), 72 °C (60 s), and a final extension at 72 °C for 5 min (Fig. 1b). Sequencing was performed on the Ion Torrent PGM (Life Technologies) using a 316 chip, resulting in 3,286,118 read fragments (416 Mbp, 126 bp mean read length).

In the absence of a reference genome, short contigs were assembled in CLC Genomics Workbench v6.0 (CLCbio) from 3,203,077 unsorted sequences after trimming to

remove adapters, low quality sequences and short (<25 bp) reads, using a minimum contig length of 100 bp, word size of 21 and bubble size of 50. 1,899,488 reads were matched into 42,066 contigs averaging 149 bp long (minimum 18 bp, maximum 478) for a total assembled length of 6,257,020 bp.

1,885,052 demultiplexed reads (mean read length 125.93 bp) were aligned to the reference contigs with stringent similarity fractions (0.95) after trimming adapters and low-quality regions. Potential SNPs were identified with quality-based variant calls and filtered for variability and locus duplication. Symptoms of duplication included contigs with 100 % heterozygosity and more than two alleles per locus. Variable loci required a minimum of one heterozygote and one homozygote genotype. 1,121 filtered SNPs were obtained from 7,393 variants called by CLC .

Ion Torrent genotype data was validated for 16 loci using Sanger sequencing on the same 10 samples. Paired forward and reverse primers were designed using Primer3 (Rozen and Skaletsky 2000) with product lengths ranging from 60 to 120 bp. Samples were amplified using touch-down PCR conditions: 16 cycles of decreasing annealing temperature (56–48 °C in 0.5° steps) followed by 14 cycles at 48 °C in 10 µl reactions (0.03 U colourless GoTaq Flexi DNA polymerase (Promega Corp.), 1 × GoTaq buffer, 0.2 mM of each dNTP, 2 mM $MgCl_2$, 20 ng DNA template and 0.5 µM of each forward and reverse primer). Products were sequenced with Big Dye Terminator v1.1 chemistry on a 3,730 DNA Analyzer (Applied Biosystems). Resulting sequences matched 93 % of the Ion Torrent genotypes.

Using the methods described here, thousands of molecular markers can be developed in-house for species with unknown, complex, and duplicated genomes. The SNP markers developed in this study will be used to describe the population connectivity and diversity of the glass sponge reefs.

Acknowledgments This work was funded by an NSERC Strategic Network grant (CHONe) and an NSERC Ship Time grant to SPL. We thank D Coltman and the Molecular Biology Service Unit (University of Alberta) for assistance and use of facilities.

References

Baums I, Hughes C, Hellberg M (2005) Mendelian microsatellite loci for the Caribbean coral *Acropora palmata*. Mar Ecol Prog Ser 288:115–127. doi:10.3354/meps288115

Conway KW, Krautter M, Barrie JV, Neuweiler M (2001) Hexactinellid Sponge Reefs on the Canadian Continental Shelf: a Unique "Living Fossil". Geosci Canada 28:71–78

Ji Y-J, Zhang D (2004) Characteristics of microsatellite DNA in lepidopteran genomes and implications for their isolation. Acta Zool Sin 50:608–614

Poland J a, Brown PJ, Sorrells ME, Jannink J-L (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE 7:e32253. doi:10.1371/journal.pone.0032253

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: Misener S, Krawetz SA (eds) Methods Mol. Biol. vol. 132 Bioinforma Methods Protoc. Humana Press, Totowa, pp 365–386