Conservation Genet Resour
DOI 10.1007/s12686-011-9521-5

TECHNICAL NOTE

New tetranucleotide microsatellite loci in pink abalone (*Haliotis corrugata*) isolated via 454 pyrosequencing

A. P. Greenley · A. Muguia-Vega · A. Saenz-Arroyo · F. Micheli

Received: 26 August 2011 / Accepted: 28 August 2011

Abstract Microsatellite loci were developed for the pink abalone (*Haliotis corrugata*) via 454 high-throughput sequencing. From 193 tetranucleotide repeats identified from 77.5 Mb of sequence, we tested 80 loci and successfully amplified and scored 18 microsatellite markers. All loci were polymorphic with number of alleles ranging from 5 to 21. Average observed and expected heterozygosities were 0.745 and 0.844, respectively. Three loci deviated from Hardy–Weinberg equilibrium, two of which had significant deficits of heterozygotes and only one displayed statistical evidence of a null allele. None of the loci exhibited linkage disequilibrium. These loci are a valuable asset for fine-scale population genetic and paternity studies centered on the conservation and management of pink abalone.

Keywords Baja California · Fisheries · *Haliotis corrugata* · Abalone · Microsatellites · Paternity

On the Pacific coast of Baja California, Mexico, pink abalone (*Haliotis corrugata*) are extremely valuable to local fisheries and economies. Yet throughout their range, populations of pink abalone have severely declined (Karpov et al. 2000; Morales-Bojórquez et al. 2008). In efforts to replenish abalone populations, fishermen in central Baja California are beginning to incorporate marine reserves into their management strategies. In theory, reserves can augment fisheries yield via export of larvae to fished areas (Gell and Roberts 2003), yet empirical studies demonstrating reserve "spillover" are still lacking (Palumbi 2003; Sale et al. 2005). In recent years, microsatellite markers have increasingly been utilized to identify potential larval sources through parental analyses (Planes et al. 2009; Christie et al. 2010). However, such techniques often require numerous polymorphic microsatellite markers (Jones and Ardren 2003; Selkoe and Toonen 2006). Here we describe 18 novel, polymorphic, tetranucleotide microsatellite loci derived from 454 pyrosequencing (Malausa et al. 2011) of *Haliotis corrugata*. These loci supplement existing microsatellite markers developed for *H. corrugata* (Díaz-Viloria et al. 2008) and increase the statistical power and capabilities of future genetic analyses.

Epipodial tissue samples of pink abalone harvested at Isla Natividad, Baja California Sur, were collected and preserved in 95% ethanol. Genomic DNA was extracted using Nucleospin column extraction kits (Machery-Nagel). Approximately 5 µg of genomic DNA from a single *H. corrugata* individual was treated with RNAse, and submitted to the University of Arizona Genetic Core (UAGC) where the sample was used to construct a random library that was sequenced in the equivalent to 1/8th of a plate using the GS FLX Titanium chemistry and multiplex identifiers (Roche Applied Science). After applying a Q20 criteria over a 10 bp window to trim the ends of the sequences and removing the library tag, we obtained 77.5 Mb distributed over 207,874 reads with an average

A. P. Greenley (✉) · F. Micheli
Hopkins Marine Station, Stanford University,
120 Oceanview Blvd, Pacific Grove, CA 93950, USA
e-mail: greenley@stanford.edu

A. Muguia-Vega
Conservation Genetics Laboratory, School of Natural Resources and the Environment, The University of Arizona,
BioSciences East 317, Tucson, AZ 85721, USA

A. Saenz-Arroyo
Comunidad y Biodiversidad A.C., Blvd. Agua Marina #297,
Entre Jaiba y Tiburón, Colonia Delicias, 85420 Guaymas,
Sonora, Mexico

Published online: 10 September 2011

Table 1 Characterization of 18 microsatellite loci isolated from *Haliotis corrugata*

Locus GenBank	Repeat motif	Primer sequences (5′–3′)	Size range (bp)	N_a	H_o	H_e
HCOR01	$(tatc)_{26}$	F: TTCAGTTTTATCGTTCGCCC	265–342	16	1.000	0.914
JN619387		R: ATAGCCGTTGTCTTTGGCCT				
HCOR02	$(ctaa)_{24}$	F: GTGCATCCGACAAAAAGTGA	136–222	18	0.889	0.928
JN619388		R: ACCTACCAAGTTAAGATCACTCTGA				
HCOR03	$(tggt)_{23}$	F: AGTGGGTTCAGTATGGCGAC	98–179	12	0.241*	0.670
JN619389		R: ATGAGTGCTGCGTCAAAAGA				
HCOR05	$(tgag)_{21}$	F: TCTCATGTTATCAGTCACATTGGAT	296–372	18	0.815	0.922
JN619390		R: GGTTCAACATCAGATGCACG				
HCOR07	$(gata)_{20}$	F: AACGGCACTTGTTGTTACCC	239–318	15	0.840	0.901
JN619391		R: CACCCCCAATTCATGTTAAA				
HCOR11	$(agtg)_{19}$	F: CAGCCTATTTGAGGATCTGGA	179–268	19	0.889	0.928
JN619392		R: ACCCATTTCATGTAGGCTCC				
HCOR13	$(tgag)_{18}$	F: CGTTCGCATGTATGAGTTGTTT	140–211	14	0.731	0.882
JN619393		R: GTCCTTTTCAGGACCACCAA				
HCOR18	$(cact)_{17}$	F: TGTTATCAGTCCTTTCTGTTGAAAT	126–178	13	0.889	0.880
JN619394		R: TTGAATGAATGAGTGCGATG				
HCOR19	$(taaa)_{17}$	F: CTCCCACCATCCTTGAACAC	91–148	13	0.833	0.875
JN619395		R: TTTGCAACATGACTAAGGCG				
HCOR27	$(aagt)_{16}$	F: TCCTGCGTGTTAAATCTCCA	144–222	16	0.750	0.905
JN619396		R: TAAGGAGTTAGGTGCCGGTG				
HCOR28	$(agtg)_{16}$	F: TGAGTCTCAGTGTCGCCTAA	103–154	12	0.815	0.888
JN619397		R: AACATGATCTGACAACATCAAAA				
HCOR46	$(ataa)_{13}$	F: CGCGATGTGTAGAAAAGCGT	118–154	9	0.742	0.683
JN619398		R: CAGTTTGACAAAAACAAAAACGA				
HCOR54	$(ttgt)_{13}$	F: GAATGGCAGTTTTGGCTTGT	263–304	11	0.750	0.851
JN619399		R: AACCCATGCTTGTCGTAAGG				
HCOR65	$(ctca)_{12}$	F: GTGATCTAGTGGCAATGGGG	374–430	11	0.621	0.829
JN619400		R: CGCCTGAATGTTTCTGGAAT				
HCOR66	$(gtga)_{12}$	F: TGGACTCATTGGAATGGAAGA	175–257	20	0.833*	0.929
JN619403		R: CCTTTGCAGTATCACCTGTTCA				
HCOR72	$(actc)_{11}$	F: AACGCCGACGTTTACCATAG	175–217	11	0.480*	0.838
JN619404		R: GCGTTATCACCAAGGAGACAG				
HCOR75	$(tgag)_{11}$	F: ACCGATTGGATGTGAGTGGT	360–458	21	0.867	0.927
JN619401		R: TCCGTCTTCTCTTTTTGCGT				
HCOR80	$(cact)_{11}$	F: CCACCATGATATTGCTGGAA	153–185	5	0.421	0.436
JN619402		R: AACACCTGCAACAGCAACAG				

Locus/GenBank accession number, repeat motif, primer sequences, size range of allelic variation, number of different alleles (N_a), and observed (H_o) and expected (H_e) heterozygosities

Loci were successfully amplified and scored on thirty-one individuals from Isla Natividad, Baja California Sur

* Loci that significantly deviated from HWE after Bonferroni correction (adjusted critical $P \leq 0.0028$)

length of 373 bp. We searched for microsatellite loci having a minimum of 10 repeats for dinucleotides, and eight repeats for tri and tetranucleotides, using the software iQDD (Meglécz et al. 2010), that implements CLUSTALw2 (Larkin et al. 2007), BLAST (ftp://ftp.ncbi.nih.gov/blast/executables/) and PRIMER 3 (Rozen and Skaletsky 2000) in a Perl environment (http://www.activestate.com/activeperl/). Sequence reads from duplicated loci and mobile elements were identified as sequence similarity groups (<95% similarity in flanking regions ≥30 bp) and excluded. Primer sequences were then designed based on unique reads and consensus sequences (≥95% similarity in

flanking regions), avoiding primers in flanking regions with short repeats (nanosatellites, ≤6 repetitions of homopolimers and ≤3 di-hexabase repetitions). The method is considered a conservative way to eliminate null alleles, redundancy and problematic loci for microsatellite amplification (Meglécz et al. 2010).

We found a total of 47 dinucleotides, 161 trinucleotides and 193 tetranucleotides for which primers were successfully designed. From these, we selected 80 tetranucleotide loci with the largest numbers of repeats for primer synthesis. To allow fluorescent labeling, the universal M13 primer was added at the 5′ end of all forward primers (Schuelke 2000). PCR conditions and protocols were conducted in 15 µl volumes with ~40 ng genomic DNA, 1× PCR buffer, 0.2 mM each dNTP, 1.5 mM $MgCl_2$, 0.2% BSA, 0.5 U Taq DNA polymerase (Fermentas), 0.02 µM of the unlabeled M13-tailed forward primer, 0.2 µM of the fluorescently-labeled M13 primer, and 0.2 µM of the reverse primer (Munguia-Vega et al. 2010). For thermocycling we used a touchdown protocol consisting of 94°C for 5 min, 15 cycles of 94°C for 30 s, 65-50°C for 30 s (1°C decrease each cycle), 72°C for 30 s, followed by 40 cycles at 94°C for 30 s, 55°C for 30 s, 72°C for 30 s, and a final extension of 72°C for 5 min. PCR products were genotyped using an Applied Biosystems' 3730 Genetic Analyzer (Molecular Genetics Core Facility, Children's Hospital Boston) and alleles were scored using GENEMAPPER 4.0 (Applied Biosystems). Allele sizes were assigned bins using FLEXIBIN (Amos et al. 2007), observed and expected heterozygosities were calculated using GENALEX 6.2 (Peakall and Smouse 2006), and deviations from Hardy–Weinberg equilibrium (HWE), linkage disequilibrium (LD) and heterozygote deficiencies were estimated using GENEPOP 3.4 (Raymond and Rousset 1995). We used MICROCHECKER 2.2.3 (Van Oosterhout et al. 2004) to test for genotyping errors and presence of null alleles. A sequential Bonferroni test for multiple comparisons was used to adjust p-values to a predetermined experiment error rate of 0.05.

We successfully amplified and scored eighteen polymorphic, tetranucleotide loci on 31 individuals from Isla Natividad, Baja California. The loci showed high polymorphism. The mean number of alleles per locus was 14.1 (range 5–21) and average observed and expected heterozygosities were 0.745 and 0.844, respectively. Three loci (Hco03, Hco66, Hco72) significantly deviated from HWE ($P < 0.0028$) due to heterozygote deficiencies that were significant only for Hco03 and Hco72.

MICROCHECKER detected null alleles for Hco03, Hco66, and Hco72 at estimated frequencies of 0.2887, 0.0524 and 0.2131, respectively (according to the method of Van Oosterhout et al. 2004) but only the null allele at Hco72 was statistically significant ($P < 0.001$). We did not find evidence of LD among loci pairs (adjusted P values > 0.0003). The estimated combined probability of identity was 7.4×10^{-29} and the non-exclusion probability of paternity assignment when both parents are unknown was 8.05×10^{-20}. These markers will be useful to estimate paternity and fine-scale population genetic structure in *H. corrugata* (Table 1).

Acknowledgments This research was funded by the Walton Family Foundation. We wish to thank the fishing cooperative S.C.P.P. Buzos and Pescadores at Isla Natividad, S. Palumbi for laboratory support, and L. Vazquez, A. Haupt, and A. Hernandez for sample collection. Research was conducted under permit no. DGOPA.03185. 030510.1597, Secretaria de Agricultura, Ganaderia, Desarrollo Rural, Pesca, Y Alimentacion (SAGARPA). This is a scientific contribution of Stanford University and Comunidad Y Biodiversidad: http://www.cobi.org.mx/.

References

Amos W, Hoffman JI, Frodsham A, Zhang L, Best S, Hill AVS (2007) Automated binning of microsatellite alleles: problems and solutions. Mol Ecol Notes 7:10–14

Christie MR, Tissot BN, Albins Ma, Beets JP, Jia Y, Ortiz DM, Thompson SE, Ma Hixon (2010) Larval connectivity in an effective network of marine protected areas. PLoS ONE 5:e15715

Díaz-Viloria N, Pérez-Enríquez R, Fiore-Amaral G, Burton RS, Cruz P (2008) Isolation and cross-amplification of microsatellites in pink abalone (*Haliotis corrugata*). Mol Ecol Resour 8:701–703

Gell FR, Roberts CM (2003) Benefits beyond boundaries: the fishery effects of marine reserves. Trends Ecol Evol 18:448–455

Jones AG, Ardren WR (2003) Methods of parentage analysis in natural populations. Mol Ecol 12:2511–2523

Karpov KA, Haaker PL, Taniguchi IK, Rogers-Bennett L (2000) Serial depletion and the collapse of the California abalone (*Haliotis* spp.) fishery. Fisheries (Bethesda) 200:11–24

Larkin MA, Blackshields G, Brown NP et al (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948

Malausa T, Gilles A, Meglécz E, Blanquart H, Duthoy S, Costedoat C, Dubut V, Pech N, Castagnone-Sereno P, Délye C, Feau N, Frey P, Gauthier P, Guillemaud T, Hazard L, Le Corre V, Lung-Escarmant B, Malé P-JG, Ferreira S, Martin J-F (2011) High-throughput microsatellite isolation through 454 GS-FLX titanium pyrosequencing of enriched DNA libraries. Mol Ecol Resour 11:638–644

Meglécz E, Costedoat C, Dubut V et al (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. Bioinformatics 26:403–404

Morales-Bojórquez E, Muciño-Díaz MO, Vélez-Barajas JA (2008) Analysis of the decline of the abalone fishery (*Haliotis fulgens* and *H. corrugata*) along the Westcentral coast of the Baja California Peninsula, Mexico. J Shellfish Res 27:865–870

Munguia-Vega A, Soria G, Pfister T, Cudney-Bueno R (2010) Isolation and characterization of microsatellite loci in the rock scallop (*Spondylus calcifer*) (Bivalvia: Spondylidae) from the Northern Gulf California, Mexico. Conserv Genet Resour 2:51–54

Palumbi SR (2003) Population genetics, demographic connectivity, and the design of marine reserves. Ecol Appl 13:146–158

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in excel population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Planes S, Jones GP, Thorrold SR (2009) Larval dispersal connects fish populations in a network of marine protected areas. Proc Natl Acad Sci USA 106:5693–5697

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Rozen S, Skaletsky H (2000) Primer3 on the www for general users and for biologist programmers. In: Misener S, Krawetz SA (eds) Methods in molecular biology. Humana Press Inc., Totowa

Sale PF, Cowen RK, Danilowicz BS, Jones GP, Kritzer JP, Lindeman KC, Planes S, Polunin NVC, Russ GR, Sadovy YJ, Steneck RS (2005) Critical science gaps impede use of no-take fishery reserves. Trends Ecol Evol 20:74–80

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecol Lett 9:615–629

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538