PeerJ Preprints NOT PEER-REVIEWED

1 **Validation and characterization of thirteen microsatellite markers for queen conch,**
2 ***Lobatus gigas***

3 Nathan K. Truelove[1], Loong Fai Ho[1], Richard F. Preziosi[2], and Stephen J. Box[1]

4 [1]Smithsonian national Museum of Natural History, Smithsonian Marine Station, Fort Pierce, FL, USA
5 [2]Manchester Metropolitan University, School of Science and the Environment, Manchester, UK
6

7 **Abstract**

8 We report the development and characterization of 13 novel microsatellite loci for
9 the Caribbean queen conch, *Lobatus gigas*, an ecologically and commercially important
10 marine gastropod. Paired-end sequencing was carried out on genomic DNA from a single
11 queen conch using half a flow cell lane of an Illumina MiSeq. A total of 48 potentially
12 amplifiable loci containing microsatellites were tested on 45 individuals from the Florida
13 Keys and Bahamas. In total, 13 consistently amplifying and polymorphic microsatellite
14 loci were identified. The number of alleles ranged from 4 to 26 and observed
15 heterozygosities ranged from 0.340 to 1.00. There was no evidence of scoring error, large
16 allele dropout, or evidence of linkage disequilibrium at any locus. Four loci deviated
17 from Hardy-Weinberg equilibrium due to moderate levels of null alleles (null allele
18 frequencies ranged from 0.081 to 0.230). Although null alleles were detected at four
19 microsatellite loci, the high levels of polymorphism and moderate null allele frequencies
20 suggest that these 13 novel microsatellite markers will be useful for researchers carrying
21 out conservation genetic studies of *L. gigas*.

22

23

24

25

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2559v1 | CC BY 4.0 Open Access | rec: 28 Oct 2016, publ:

26 **1. Introduction**

27 The queen conch, *Lobatus gigas*, is a commercially important fishery resource for
28 many countries in the greater Caribbean and is part of a lucrative export market with the
29 United States and Europe (Acosta 2006). The high demand for queen conch has lead to
30 overfishing in several Caribbean nations (Hernandez-Lamb et al. 2012). The status of the
31 greater Caribbean queen conch fishery led to the listing of the species on Appendix II of
32 the Convention on International Trade in Endangered Species of Wild Fauna and Flora
33 (CITES) in 1990 and trade suspensions now exists in several Caribbean nations (Acosta
34 2006).

35 Genetic techniques are widely used to guide conservation strategies for overfished
36 species by identifying genetically unique stocks and levels of gene flow among
37 populations (Allendorf et al. 2012). Queen conch have a 2 - 4 week pelagic larval phase
38 during their life history that facilitates dispersal within the Caribbean seascape (Mitton et
39 al. 1989). Previous population genetics studies of queen conch, using mitochondrial DNA
40 and allozyme markers, found high levels of geneflow and no evidence of genetically
41 unique subpopulations (Mitton et al. 1989, Zamora-Bustillos et al. 2011, Pérez-Enriquez
42 et al. 2011). In contrast to previous genetics studies, biophysical modeling of queen
43 conch larval dispersal suggests that larval connectivity among Caribbean queen conch
44 populations may be more limited than previously thought, implying that genetic
45 differentiation may occur (Paris et al. 2006).

46 Large microsatellite datasets allow for inferences about contemporary levels of
47 gene flow in species with high levels of dispersal and allow for the detection of genetic
48 structure caused by ecological and evolutionary processes (Andersen et al. 2010; Selkoe

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2559v1 | CC BY 4.0 Open Access | rec: 28 Oct 2016, publ:

49 et al. 2008). A total of eight microsatellite markers have already been developed for
50 queen conch, however, six out of the eight are likely to have reduced statistical power for
51 detecting population structure due to a low numbers of alleles and the confounding
52 effects of five out of eight loci containing null alleles (Zamora-Bustillos et al. 2007,
53 Carlsson 2008). Larger microsatellite datasets for queen conch will help improve the
54 knowledge of genetic connectivity among queen conch populations, which is an
55 important component for conservation plans and fishing quotas.
56 The objective of our study was to develop novel polymorphic microsatellite loci
57 for *Lobatus gigas*, which will be useful for assessing genetic population structure, gene
58 flow, and genetic diversity in this CITES listed species. Next-generation sequencing
59 technology has greatly reduced the cost and time required to identify microsatellites in
60 species such as queen conch, which lack detailed genome-wide data. We applied this
61 technology to identify a panel of 13 microsatellites for queen conch.
62
63 **2. Methods**
64 **2.1 Tissue Sampling and DNA Extraction**
65 A research permit was not required to collect tissue samples from Florida since
66 the Florida Fish and Wildlife Conservation Commission assisted with sample collection.
67 The Bahamas Department of Marine Resources Research Permit Number MAMR/FIS/3
68 and CITES Export Permit 2015/258 were used to collect and export samples from the
69 Bahamas.
70 Sterile biopsy forceps were used to extract a 1.5cm^3 tissue sample from the mantle
71 of a single queen conch from the Long Key, Florida. The sample was soaked in RNA-

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2559v1 | CC BY 4.0 Open Access | rec: 28 Oct 2016, publ:

72 Later (Qiagen, Valencia, CA, USA) at 4°C for 1 day then stored at -80°C. DNA was
73 extracted using the E.Z.N.A. Mollusc DNA extraction kit (Omega Bio-Tek) following the
74 manufacturer's protocol. DNA concentration and quality was tested using 1) a NanoDrop
75 2000 spectrophotometer (Thermo Scientific), 2) Qubit® fluorometer (Life Technologies),
76 and 3) by running 5ul of DNA solution on a 2% agarose gel that was visualized using a
77 GelGreenDNA stain (Biotium Inc.). The DNA concentration was 20ng/µl, and the bands
78 on the gel were not fragmented or smeared, indicating high quality, non-degraded
79 genomic DNA (Wong et al. 2012). The DNA sample was shipped on dry ice to the
80 University of Manchester Genomics Core Facility for sequencing.
81
82 **2.2 Library Preparation and Sequencing**
83 A total of 50ng of DNA from one individual was used to create a paired-end
84 library using the Illumina Nextera® DNA Sample Preparation Kit according to the
85 manufacturer's protocol, with Illumina adaptors used for identification purposes. Two-
86 hundred and fifty cycle paired-end sequencing was carried out in half a flow cell lane of
87 the Illumina MiSeq platform at the Genomics Core Facility at the University of
88 Manchester. The left and right raw reads were separated into two files and converted to
89 the FASTQ format to conduct further quality control and filtering procedures.
90
91 **2.3 Quality Filtering of Sequence Data**
92 The FASTQ files containing raw MiSeq reads were imported into Galaxy-Golem,
93 a customized local installation of the Pennsylvania State University's Galaxy, run by the
94 University of Manchester's Bioinformatics Core Facility (Goecks et al. 2010). Galaxy is

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2559v1 | CC BY 4.0 Open Access | rec: 28 Oct 2016, publ:

PeerJ Preprints NOT PEER-REVIEWED

95 an open-access web-based bioinformatics tool that allows researchers to run a number of
96 analyses on large next-generation sequencing datasets. The quality of the raw Illumina
97 reads was checked in Galaxy using FastQC
98 (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/) and reads with low quality were
99 removed using Trimmomatic (Bolger et al. 2014). The sliding window trimming
100 operation was performed using four bases and an average Phred quality score (Ewing and
101 Green 1998) of 20 or higher.
102
103 **2.4 Microsatellite Identification and Primer Design**
104 Potentially amplifiable loci (PAL) containing microsatellites were identified with
105 program PAL_FINDER v0.02, which was incorporated into Galaxy (Castoe et al. 2012).
106 Primer development settings were adjusted to develop microsatellites PCR primers with
107 optimal melting temperatures for use with the Qiagen Type-it® Microsatellite PCR Kit.
108 Primer design parameters were set to: minimum melting temperature (Tm) 60°C;
109 optimum Tm 68°C; maximum difference in Tm between primers in pair 2°C; and primer
110 length 20-30bp. Primers meeting these criteria were then screened against all other
111 MiSeq reads to identify primers that may bind to > 1 location in the conch genome. Only
112 primers that bound to 1 location in the MiSeq dataset were selected for primer testing to
113 minimize the possibility of primers binding to multiple sites in a repetitive region of the
114 queen conch genome.
115
116
117

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2559v1 | CC BY 4.0 Open Access | rec: 28 Oct 2016, publ:

118 **2.5 Microsatellite Primer Testing**

119 The PCR amplification success of 48 PALs were tested on DNA from eight queen
120 conch filets purchased from the Pelican Bay Fish Market in Fort Pierce, Florida, USA.
121 DNA was extracted using the QIAGEN DNeasy kit following the manufacturer's
122 protocol. DNA concentrations for each sample were quantified using a Nanodrop ND-
123 1000 and they ranged from 15ng/µl to 100ng/µl. The quality of genomic DNA was tested
124 by visualizing 5µl of genomic DNA solution on a 2% agarose and visualized with a
125 GelGreenDNA stain (Biotium Inc.). All DNA samples showed a single high molecular
126 weight band with no smearing, indicating that the DNA used for primer testing was of
127 high quality (Wong et al. 2012). PCR was carried out using the Type-it® Microsatellite
128 PCR kit (Qiagen) in a reaction volume of 10µl, with a thermal cycle of: 95°C for 5
129 minutes, 28 x (95°C for 30 seconds, 60°C for 90 seconds, 72°C for 30 seconds) and 60°C
130 for 30 minutes. The final concentration of forward and reverse primers were 0.2 µM.
131 PCR products were checked on a 2.0 % agarose gel to determine if PCR amplification
132 was successful. A PAL was classed as successfully amplified if at least 7 of the 8 samples
133 tested showed one or two clear bands on the gel.

134 Primer pairs that successfully amplified 15 PALs were further tested on 45
135 individuals from Delta Shoal in the Florida Keys (n = 23) and Grand Cay, Bahamas (n =
136 22). Bahamas Department of Marine Resources Research Permit Number MAMR/FIS/3
137 and CITES Export Permit 2015/258 were used to collect samples and export samples
138 from the Bahamas. DNA extraction of these individuals followed the same methodology
139 as the eight individuals used for the first stage of primer testing. The quality of PALs to
140 produce reliable genotypes were tested by labeling the 5' end of the forward primer of

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2559v1 | CC BY 4.0 Open Access | rec: 28 Oct 2016, publ:

141 each PAL with either the 6-FAM or HEX fluorescent label (Table 1). PCRs conditions
142 were the same as those used for the first stage of primer testing. PCR products were sent
143 to the Smithsonian Laboratory for Analytical Biology where fragment analysis was
144 conducted on an ABI 3730xl Sequencer (Life Technologies). These data were manually
145 scored using GENMAPPER 3.7 (Life Technologies). GENODIVE (Meirmans and van
146 Tienderen 2004) was used to estimate observed (H_O) and expected (H_E) heterozygosity,
147 number of alleles (N_A), and deviations from Hardy–Weinberg equilibrium. The presence
148 of null alleles or scoring errors caused by excessive stuttering was analyzed with
149 MICROCHECKER (van Oosterhout et al. 2004). Linkage disequilibrium was tested for
150 by using the program GENEPOP (Raymond and Rousset 1995, Rousset 2008).
151
152 **3. Results**
153 The Illumina MiSeq run produced 3,872,724 (2 x 1,936,362) total reads with
154 852,641 reads containing microsatellites. Primer testing identified 13 primer pairs that
155 consistently amplified polymorphic microsatellite loci. The 13 microsatellite loci
156 produced allele sizes ranging from 144 base pairs to 476 base pairs. The number of
157 alleles ranged from 4 to 26 and observed heterozygosities ranged from 0.340 to 1.00.
158 There was no evidence of scoring error, large allele dropout, or evidence of linkage
159 disequilibrium at any locus. Four loci (Conch31, Conch37, ConchPR1, and ConchPR7)
160 deviated from Hardy-Weinberg equilibrium due to moderate levels of null alleles with
161 null allele frequencies ranging from 0.081 to 0.230 (Table 1).
162
163

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2559v1 | CC BY 4.0 Open Access | rec: 28 Oct 2016, publ:

164 **4. Discussion**

165 Although null alleles were detected at four microsatellite loci, the high levels of
166 polymorphism and moderate null allele frequencies suggest that these 13 novel
167 microsatellite markers will be useful for researchers carrying out conservation genetic
168 studies of queen conch. Such studies can be used to help further the development of
169 conservation and management plans of the Caribbean queen conch *Lobatus gigas*.

170

171 **5. Acknowledgements**

172 The Summit Foundation provided funding for S. J. Box. We thank Lester Gittens
173 at the Bahamas Department of Marine Resources for his assistance with research and
174 CITES export permits from the Bahamas. We are grateful for the fieldwork assistance
175 provided by Einat Sandbank, Gabriel Delgado and Bob Glazer from the Florida Fish and
176 Wildlife Conservation Commission collecting samples in the Florida Keys and grateful
177 for the help Sara Rosewall, Phaedra Geer, and Emily Dark provided in the lab.

178

179 **6. Literature Cited**

180 Acosta, C. A. (2006). Impending trade suspensions of Caribbean queen conch under
181 CITES: a case study on fishery impacts and potential for stock recovery.
182 Fisheries, 31(12), 601-606.
183
184 Allendorf, F. W., G. H.Luikart, and S. N. Aitken. 2009. Conservation and the genetics of
185 populations. John Wiley & Sons.
186
187 Anderson, C. D., B. K. Epperson, M.-J. Fortin, R. Holderegger, P. M. A. James, M. S.
188 Rosenberg, K. T. Scribner, and S. Spear. 2010. Considering spatial and temporal
189 scale in landscape-genetic studies of gene flow. Molecular Ecology 19(17):3565–
190 3575.
191
192 Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for
193 Illumina sequence data. Bioinformatics: btu170.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2559v1 | CC BY 4.0 Open Access | rec: 28 Oct 2016, publ:

194 Carlsson, J. 2008. Effects of Microsatellite Null Alleles on Assignment Testing.Journal
195 of Heredity, 99(6), 616-623.
196
197 Castoe, T. A., A. W. Poole, A. P. J. de Koning, K. L. Jones, D. F. Tomback, S. J. Oyler-
198 McCance, J. A. Fike, S. L. Lance, J. W. Streicher, E. N. Smith, and D. D. Pollock.
199 2012. Rapid Microsatellite Identification from Illumina Paired-End Genomic
200 Sequencing in Two Birds and a Snake. PloS one 7(2):e30953.
201
202 Ewing, B., and P. Green. 1998. Base-Calling of Automated Sequencer Traces Using
203 Phred. II. Error Probabilities. Genome Research 8(3), 186-194.
204
205 Goecks, J., A. Nekrutenko, and J. Taylor. 2010. Galaxy: a comprehensive approach for
206 supporting accessible, reproducible, and transparent computational research in the
207 life sciences. Genome Biology 11(8):R86.
208
209 Hernandez-Lamb, J., A. Dibello, S. Lewis, G. Mackin, K. Kirby, and C. Acosta. 2012.
210 Modelling the effects of reserve size and fishing mortality for Caribbean queen conch
211 Strombus gigas. Aquatic Conservation: Marine and Freshwater Ecosystems
212 22(6):721–730.
213
214 Meirmans, P. G., and P. H. van Tienderen. 2004. Genotype and Genodive: two programs
215 for the analysis of genetic diversity of asexual organisms. Molecular Ecology Notes
216 4(4):792–794.
217
218 Mitton, J. B., C. J. Berg, and K. S. Orr. 1989. Population structure, larval dispersal, and
219 gene flow in the queen conch, Strombus gigas, of the Caribbean. The Biological
220 Bulletin 177(3):356–362.
221
222 Paris, C. B., M. P. Perez, J. Kool, and D. A. Aranda. 2006. Segregation of Queen Conch,
223 Strombus gigas, Populations from the Yucatan Peninsula, Mexico. Caribbean
224 Connectivity: Implications for Marine Protected Area Management.Proceedings of
225 Special Symposium at the 59th Annual Meeting of the Gulf and Caribbean Fisheries
226 Institute. Belize City, Belize. pp71-89.
227
228 Pérez-Enriquez, R., F. J. Garcia-Rodriguez, G. Mendoza-Carrion, and C. Padilla. 2011.
229 Geographical variation in the genetic diversity and composition of the endangered
230 queen conch Strombus gigas (Mesogastropoda: Strombidae) from Yucatán, México.
231 Revista de biología tropical 59(3):1115–1126.
232
233 Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): population genetics
234 software for exact tests and ecumenicism. Journal of Heredity 86(3):248–249.
235
236 Rousset, F. 2008.genepop’007: a complete re-implementation of the genepop software
237 for Windows and Linux. Molecular Ecology Resources 8(1):103–106.
238
239

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2559v1 | CC BY 4.0 Open Access | rec: 28 Oct 2016, publ:

240 Selkoe, K. A., C. M. Henzler, and S. D. Gaines. 2008. Seascape genetics and the spatial
241 ecology of marine populations. Fish and Fisheries 9(4):363–377.
242
243 Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. Micro-
244 checker: software for identifying and correcting genotyping errors in microsatellite
245 data. Molecular Ecology Notes 4(3):535–538.
246
247 Waples, R. S. 1998. Separating the wheat from the chaff: patterns of genetic
248 differentiation in high gene flow species. Journal of Heredity 89(5):438–450.
249
250 Wong, P. B., E. O. Wiley, W. E. Johnson, O. A. Ryder, S. J. O'Brien, D. Haussler, K.-P.
251 Koepfli, M. L. Houck, P. Perelman, G. Mastromonaco, A. C. Bentley, B. Venkatesh,
252 Y.-P. Zhang, R. W. Murphy, G10KCOS. 2012. Tissue sampling methods and
253 standards for vertebrate genomics. GigaScience 1(1):8.
254
255 Zamora-Bustillos, R., R. Rodríguez-Canul, and F. J. G. De León. 2007. Isolation and
256 characterization of eight polymorphic microsatellite markers from pink conch
257 (Strombus gigas). Molecular Ecology Notes 7(4):597–599.
258
259 Zamora-Bustillos, R., R. Rodríguez-Canul, F. J. García de León, and J. TelloCetina.
260 2011. Diversidadgenética de dos poblaciones del caracol Strombus gigas
261 (Gastropoda: Strombidae) en Yucatán, México, con microsatélite. Revista de biología
262 tropical 59(3):1127–1134.
263

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2559v1 | CC BY 4.0 Open Access | rec: 28 Oct 2016, publ:

PeerJ Preprints NOT PEER-REVIEWED

Table 1. Primer sequence, repeat motifs, fluorescently labeled dyes and characteristics of 13 microsatellite loci: Label (fluorescent label attached to forward primer), GenBank (GenBank Accession Number), T_A (annealing temperature), Na (number of alleles), Ho (observed heterozygosity), He (expected heterozygosity), Fis (fixation index), *P* (*P*-value for deviation from Hardy-Weinberg equilibrium), and F_{NA} (null allele frequency). Significant values after the false discovery rate correction for multiple comparisons (Benjamini & Hochberg, 1995) are in bold.

Locus	Primer sequence (5'-3')	Repeat motif	Label	GenBank	Size range	Ta	Florida (n = 23)						Bahamas (n = 22)					
							Na	Ho	He	Fis	*p*	F_{NA}	Na	Ho	He	Fis	*p*	F_{NA}
Conch3	F: GGAAATGTGAAGCATCAGCG	(AG)19 (TC)23	Hex	KP985701	211-309	63.16	26	1.000	0.962	-0.039	0.445	---	21	0.889	0.956	0.070	0.189	---
	R: GACGCTGTAGGAACAAGGGG																	
Conch4	F: CTTGGCGTAGCTACCTGACC	(TG)8 (AG)18 (TC)29	6fam	KP985702	198-242	65.25	10	0.957	0.879	-0.088	0.203	---	16	0.889	0.936	0.051	0.355	---
	R: GGAGAGAGCCAGAATGAGGG																	
Conch17	F: GAGGGATATGACGATGGTGG	(ATC)12 (ATG)14 (ACC)3	6fam	KP985703	426-459	63.02	7	0.826	0.831	0.006	0.566	---	9	0.762	0.804	0.052	0.394	---
	R: TGAGGAACAAGGTCTGGTCG																	
Conch21	F: CAGGTCTGGCTGGGAAGC	(TTC)12 (AAG)68	6fam	KP985704	248-280	64.36	9	0.870	0.873	0.003	0.565	---	9	0.778	0.845	0.079	0.309	---
	R: AAGTTGCCGTCAGTTTTGAGG																	
Conch23	F: AACATTTGTCACTGCTGCTGG	(TGC)15 (AGG)3 (ACT)3 (AGT)4 (ACC)3	Hex	KP985705	361-391	62.98	4	0.696	0.594	-0.171	0.193	---	7	0.350	0.499	0.298	0.037	---
	R: TGATAGGAGTGGTCAGGATGG																	
Conch29	F: TCAGTCTGTCTGTTTTGGCG	(ACAG)6 (TCTG)9 (TC)8	6fam	KP985706	259-294	62.71	12	0.696	0.870	0.200	0.02	---	12	0.818	0.871	0.061	0.310	---
	R: ATAGAGAATTGGAGACAGTAAGCG																	
Conch31	F: GCCTCCCACAGAATACCAGC	(AGG)27 (AAG)9 (TTC)18 (ATG)19 (TCC)7	Hex	KP985707	276-300	66.36	6	0.609	0.639	0.048	0.486	---	4	0.286	0.596	0.521	**0.005**	0.230
	R: CAAGACTGCACCACTGTCCC																	
Conch37	F: CAGTTTCCAGTCAAGAATACCCC	(TC)90	Hex	KP985708	234-263	61.93	12	0.696	0.912	0.237	**0.002**	0.107	11	0.476	0.923	0.484	**0.001**	0.229
	R: GCATCCACAATGAATCAAAATCC																	
Conch39	F: ATGTATCCACCCATCCACGG	(AG)42	6fam	KP985709	257-321	64.9	23	1.000	0.965	-0.036	0.448	---	20	0.867	0.969	0.106	0.077	---
	R: GTACTTCAGCCGCCAGATCC																	
Conch44	F: TGTGTATGTGTGTCACTCTGCG	(TG)50 (TTTG)3	6fam	KP985710	230-264	62.98	12	1.000	0.854	-0.171	0.02	---	16	1.000	0.942	-0.062	0.275	---
	R: TCTTGATGGCTTGTGGTTGG																	
ConchPR1	F: GCGAAGCTTGATTCAAAATGG	(ATCT)38	Hex	KP985711	364-476	61.81	17	0.696	0.950	0.267	**0.002**	0.124	15	0.773	0.921	0.161	0.029	---
	R: AAAGAGGGTGCTTCTGTGGC																	
ConchPR7	F: TGAAAGAGTGGCCAATACGC	(TTTG)29 (AAAC)17	Hex	KP985712	168-261	62.92	22	0.913	0.956	0.045	0.273	---	18	0.773	0.950	0.187	**0.002**	0.081
	R: GCTTTCCCTTGATCCCAGC																	
ConchPR11	F: CACTACGATAGATTGTGGCAGC	(AT)72	Hex	KP985713	144-209	64.21	19	0.826	0.937	0.118	0.056	---	16	0.952	0.937	-0.017	0.592	---
	R: TGTCAGTACAGGTTTGAGTGACC																	

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2559v1 | CC BY 4.0 Open Access | rec: 28 Oct 2016, publ: