Isolation and characterization of microsatellite loci for the introduced broad-nosed weevil *Galapaganus howdenae howdenae* in the Galapágos archipelago

C. C. STEPIEN, M. E. KACZMAREK, H. F. MOK, A. J. STUCKERT, P. P. CHEN, S. J. DOWNER and A. S. SEQUEIRA

Department of Biological Sciences, Wellesley College, 106 Central Street, Wellesley, MA 02481

Keywords: Bottleneck, introduced species, invasive potential, endemic genus

Correspondence: Andrea Sequeira, Department of Biological Sciences, Wellesley College, 106 Central Street, Wellesley, MA 02481, USA, Fax: 781-283-3642; Email: asequeir@wellesley.edu

Running Title: Microsatellite loci for an introduced weevil

Abstract

Eight microsatellite loci were developed for *Galapaganus howdenae howdenae*, a winged broad-nosed weevil introduced to the agricultural zone of Santa Cruz Island, Galapágos archipelago. It has also recently been found outside the agricultural zone on endemic vegetation alongside wingless endemic close relatives. Between two to four alleles were documented per locus and observed and expected heterozygosities ranged from 0.034 to 0.932 and 0.130 to 0.735, respectively. Three out of eight loci deviated from Hardy-Weinberg equilibrium, possibly suggesting that populations of *G. h. howdenae* have experienced some demographic and population structuring effects as a result of their recent colonization history.

Galapaganus howdenae howdenae was first reported in the island of Santa Cruz in 1996 and has been formally listed as an invasive species in the Galápagos archipelago (Causton *et al.* 2005). The island invader's natural range encompasses lowlands of mainland Ecuador, and in contrast to its endemic counterparts which are flightless, has well-developed metathoracic wings (Lanteri 2004). Even though *G. h. howdenae* is thought to have been introduced to the lowland agricultural zone in Santa Cruz, it has now been repeatedly found outside of the boundaries of the disturbed area on endemic and introduced vegetation alongside its endemic close relatives.

Two microsatellite libraries were created using *MseI* and *AseI* restriction enzymes to generate fragments according to the FIASCO (Fast Isolation by AFLP of Sequences COntaining repeats) protocol (Zane *et al.* 2002). DNA was extracted from four specimens of *G. h. howdenae* from two localities in Santa Cruz Island using a DNeasy Blood & Tissue Kit (QIAGEN Inc.). Templates were digested and ligated to AFLP (amplified fragment length polymorphism) adaptors and amplified by polymerase chain reaction with the adaptor-specific primers. The mixture was enriched for microsatellite-containing fragments by magnetic bead selection (Streptavidin NEB) with biotinylated $(AC)_{17}$. Selected fragments were then amplified with the adaptor-specific primers and cloned using a TOPO-TA cloning kit (Invitrogen). A total of 334 clones were created in *MseI* and *AseI* libraries. Plasmids were extracted using QIAprep Spin Miniprep Kit (QIAGEN Inc.) and those bearing inserts greater than 400bp were sequenced and analyzed for repeats using SEQUENCHER version 4.7 (GeneCodes Corporation). Of the 271 plasmids screened, 223 were sequenced. From the 67 plasmids containing unique di-, tri- or tetra-nucleotide repeats, 35 contained suitable flanking regions for primer design. Primers were designed using PRIMER 3 (Rozen & Skaletsky 2000) and Invitrogen's web tool OligoPerfect.

PCR conditions were successfully optimized for eight microsatellite regions that were later found to be polymorphic. PCR reactions consisted of 1-2uL DNA template (20-120ng DNA), 1uL 10X ThermoPol Reaction Buffer containing 2mM $MgSO_4$ (New England Biolabs), 0.8mM dNTPs (0.2mM each) (Invitrogen), 0.5uM of each primer, 0.25ul *Taq* polymerase (New England Biolabs), and 0-3.0mM additional $MgCl_2$ in a 10uL final volume. Details of optimal cycling and reaction conditions for genotyping amplifications were not consistent across loci (Table 1). Microsatellite variation was assessed for the eight loci in 59 individuals of the introduced *G. h. howdenae* from Santa Cruz Island. Genotyping amplifications were performed with one fluorescently tagged primer (HEX, PET, 6FAM, NED from Applied Biosystems). Quartets of PCR products were combined with LIZ600 size standard (Applied Biosystems) and run on an ABI3100 Genetic Analyzer. GENEMAPPER version 4.3 was used to analyze the electropherograms and allele sizes were individually confirmed. Randomly selected individuals were genotyped two to three times and verified by colleagues through blind tests to ensure accuracy in allele calling.

Observed heterozygosity, expected heterozygosity, deviation from Hardy-Weinberg equilibrium and linkage disequilibrium were calculated using GENEPOP (Raymond & Rousset 1995) and are presented in Table 2. None of the 28 pairs of loci tested showed significant linkage disequilibrium. The average expected heterozygosity across loci was 0.510 and the number of alleles per locus ranged from two to four (Table 2). Three loci (O4, AC30 and J2-2) showed deviation from Hardy-Weinberg equilibrium ($p<0.01$) using the exact H-W test, option 1.3, Probability Test as implemented in GENEPOP. Additionally, loci AC30 and J2-2 displayed significant heterozygote deficits with one tailed tests, option 1.1 in GENEPOP, and potential for

the presence of null alleles when analyzed with the program Micro-Checker (van Oosterhout *et al.* 2004).

Given that this is a relatively recent introduction, it is not unlikely that populations have suffered demographic and genetic bottlenecks. The effects of genetic drift on these populations could account for the small average number of alleles per locus through allele depletion and increased homozygosity. Additionally, little is known about the number of introductions or the continental sources of these introductions, whereby lack of conformation to H-W expectations may be caused by population genetic structure deriving from admixture, rather than due to the presence of null alleles. Mitochondrial markers indicate this recently introduced species has low genetic structure compared to established endemics (Sequeira *et al.*, *in prep*), however, preliminary analyses of large numbers of individuals across six microsatellite loci suggest subtle but significant structure between localities in the agricultural zone and the National Park area.

The development of the first microsatellite library within this genus will enable evaluation of the invasive potential of *G. h. howdenae* and possibly of the impact of this introduced species on endemic *Galapaganus* species sharing the same habitat. Additionally, if results point towards population admixture then future investigations should look into the geographic origins of that genetic variability across the native range.

Acknowledgements

We thank Dr. Chiara Papetti for advice on primer design; Manisha Sijapati for technical assistance; and Lazaro Roque Albelo, Gálapagos National Park and the Charles Darwin Research Station for logistics and support. This work was supported by NSF #0817978 to AS and a Howard Hughes Medical Institute Fellowship to CCS.

References

Causton CE, Peck SB, Sinclair BJ, *et al.* (2005) Alien insects: Threats and implications for the conservation of the Galápagos Islands. *Annals of the Entomological Society of America* **99**, 121-143.

Lanteri AA (2004) New taxonomic and biogeographic information on Galapaganus femoratus species group (Coleoptera: Entiminae). *Transactions of the American Entomological Society* **130**, 177-192.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds. Krawetz S, Misener S), pp. 365-386. Humana Press, Totowa, NJ.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology* **11**, 1-16.

Table 1 Microsatellite primer sequences, repeat motifs and optimized conditions for genotyping amplifications for each locus. $MgCl_2$ indicates the concentration (mM) of $MgCl_2$ added in addition to the 2mM $MgSO_4$ already in the buffer. Cycles consisted of an initial denaturing step of 5 min at 95°C followed by 8-20 cycles of the following: 94°C for 30 s, a primer-specific touchdown annealing temperature for 40 s decreasing in intervals as indicated from the upper bound to the lower bound, and 72°C for 30 s, followed by 30 cycles at the lower-bound annealing temperature unless otherwise specified, and a final extension at 72°C for 15 min.

Locus	Primer sequences	Repeat motif	Clone size (bp)	$MgCl_2$ (mM)	Touchdown range / interval (°C)	GenBank Accession no.
J3	F-TTGTTTCCATTAGCCCTTTTAGCG R-GCAGTTATTCTGTCACATGGTAGA	$(TG)_{42}$	170	3.0	49-43°C/0.5°C	GU595389
O4	F-CTTACATGATGCACAGATCCATC R-CATGTGTTGTTATGTATATTTGCTCTT	$(AC)_5$	131	2.5	49-43°C/0.5°C	GU595393
M1	F-GTAATAACTTTTCTCTTATTTTTTCGA R-ACAATCTGCTACTAATTCGTACACC	$(GT)_{10}$	142	1.5	55-50°C/0.5°C	GU595396
M2-1	F-TGTACGAATTAGTAGCAGATTGTAAGA R-TCGTTACTTTTAGAGAAGGATACATGT	$(TG)_7$	216	2.5	55-50°C/0.5°C	GU595394
AC30	F-AGGATGAGGTTGGGGTTAGT R-TTGATCGTGCAAAATAGATCC	$(AAT)_5$	263	1.0	55-50°C/0.5°C	GU595395
J2-2	F-TGCATGGTTACAAGACCAGG R-GTCGTTTACGATATGAAAAATTGTGA	$(GAA)_9$	200	3.0	65-58°C/1°C 36 cycles 58°C	GU595390
N1-2	F-AAAATTCAAGAAAGCCAATACATAAAA R-AACGAGTTTTCACCATTGTTGAT	$(TC)_2$ $(AC)_5$	310	1.5	65-55°C/0.5°C	GU595392
A6	F-TCTGAGATATGGCCGCGGAC R-TATGCATTCCCCACCAATTT	$(AC)_8$	240	0	70-63°C/1.0°C	GU595391

Table 2 Characterization of eight microsatellite loci isolated from *G. h. howdenae*. n: number of individuals genotyped. k: number of alleles. H_O and H_E: observed and expected heterozygosity per locus. F_{IS} estimates are global estimates over all alleles by Weir and Cockerham (1984). *p<0.01

Locus	Allele size range (bp)	n	k	H_O	H_E	F_{IS}
J3	84-190	59	2	0.492	0.476	-0.0332
O4	114-125	59	4	0.932	0.529	-0.7732*
M1	128-140	59	4	0.729	0.735	0.0087
M2-1	216-236	59	2	0.475	0.493	0.0368
AC30	260-334	59	3	0.288	0.565	0.4921*
J2-2	171-196	59	3	0.034	0.130	0.7402*
N1-2	262-318	59	3	0.576	0.659	0.1265
A6	235-237	59	2	0.492	0.490	-0.0036