Research Article

Development and characterisation of twenty polymorphic microsatellite loci in the deep-sea squat lobster, *Munida isos* Ahyong & Poore, 2004 and cross-amplification in two congeneric species

Ruo-Jin Yan[1,*], Kareen E. Schnabel[2], Xiang-Zhao Guo[1], Jonathan P. A. Gardner[1]

[1] School of Biological Sciences, Victoria University of Wellington, Wellington, New Zealand

[2]National Institute of Water and Atmospheric Research, Wellington, New Zealand

*For correspondence. E-mail: Ruojin.yan@vuw.ac.nz

Running Title: Development and characterisation of microsatellites

Unedited version

Abstract. *Munida isos* is a deep-sea squat lobster species that is widely distributed across the New Zealand and east Australian region and is often associated with deep-sea Vulnerable Marine Ecosystems. To investigate its population genetic structure and patterns of regional connectivity, microsatellite loci for *M. isos* were developed from two genomic libraries using the Illumina HiSeq 2500 sequencing platform. Twenty-six loci amplified consistently in *M. isos* from the Tasman Sea, amongst which twenty loci were polymorphic and selectively neutral. Evidence of null alleles was observed at eight loci. Most loci exhibited moderate to high levels of polymorphism, with an average Polymorphic Information Content value of 0.482. The mean number of alleles per locus was 7.45, with a mean expected heterozygosity of 0.520. Thirteen loci exhibited significant deviation from Hardy-Weinberg equilibrium, whilst only one locus pair was in linkage disequilibrium after False Discovery Rate correction for multiple testing ($p < 0.05$). Cross-species amplification tests revealed that the transferability of 14 loci (70%) was positive for the two congeners *M. endeavourae* and *M. gracilis*. The accessibility to new polymorphic microsatellite loci will facilitate population genetic studies and aid in developing conservation and management strategies for Vulnerable Marine Ecosystems.

Keywords. Genetic diversity; Genetic connectivity; Management; Conservation; South Pacific Ocean; Vulnerable Marine Ecosystems.

Introduction

Squat lobsters are a diverse group of decapod crustaceans, which are dominant at all depths throughout many of the world's oceans (Baba *et al*. 2008). *Munida isos* Ahyong & Poore, 2004 is an abundant squat lobster that is widespread in New Zealand (NZ) Exclusive Economic Zone (EEZ), and also in eastern Australia (Ahyong *et al*. 2015). Whilst there are a number of studies examining the connectivity of Vulnerable marine ecosystems (VMEs) indicator species (e.g., Zeng *et al*. 2017), relatively few studies have focussed on VME-associated taxa, including squat lobsters. Knowledge of the connectivity of marine populations is vital for designing effective marine protected areas (MPAs) networks which provide a means to protect critical natural habitats, such as VMEs (Jenkins and Stevens 2018).

Microsatellites have provided a population genetic framework for the evaluation of connectivity in many marine invertebrates (Frankham *et al*. 2002). Previous attempts to develop microsatellite loci from squat lobster species were mainly based on the construction of enriched genomic libraries (e.g., Molecular Ecology Resources Primer Development Consortium *et al*. 2010). However, using the enrichment library method has been problematic in some marine invertebrates (Selkoe and Toonen 2006). Next-generation sequencing (NGS) is now the prevailing experimental approach for the development of genetic markers (Bahassi and Stambrook 2014). To date, this technique has been employed to obtain a number of microsatellite loci for several squat lobster species from the southwest Pacific Ocean, North Atlantic Ocean, Southern Ocean and Northwest Pacific Ocean (Boyle *et al*. 2013; Coykendall and Morrison 2013; Nakajima *et al*. 2018; Roterman *et al*. 2013).

Sourcing suitable microsatellite loci from closely related species is a cost-effective and time-saving alternative. To date, cross-species amplification transferability has been successful in several squat lobster species (e.g., Boyle et al. 2013). For the present study, cross-amplification of microsatellite primers developed for *M*. *isos* is tested with *M*. *endeavourae* Ahyong & Poore, 2004 and *M*. *gracilis* Henderson, 1885. *Munida endeavourae* primarily inhabits seamounts that occur in the northern region of the NZ EEZ and the southeastern Australia region, whereas *M*. *gracilis* is generally found on both seamount (VME) and sediment (non-VME) habitats that are distributed across the entire NZ continental shelf (Schnabel 2009).

Here, we present a suite of novel microsatellite markers for the VME-associated species *M*. *isos,* as well as results for their cross-amplification in two congeners, *M*. *endeavourae* and *M*. *gracilis*. Microsatellites developed in this study will contribute to the ongoing study of connectivity amongst squat lobster populations within and beyond the NZ EEZ, and inform conservation and management strategies for the protection of VMEs.

Material and methods

Two specimens of *M. isos*, collected from the Chatham Rise (eastern region of NZ EEZ), were used for library development. Seventy-three specimens of *M. isos*, collected from the Tasman Sea (southeastern region of Australian EEZ), were utilised for the subsequent characterisation of the microsatellite loci. Ten specimens of *M*. *endeavourae* collected from the Kermadec Ridge (northeastern region of NZ EEZ) and ten specimens of *M*. *gracilis* collected from the Challenger Plateau (western region of NZ EEZ) were used for the cross-amplification test. Whole genomic DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. The integrity and molecular weight of the DNA samples were assessed in a 3% agarose gel and with a NanoPhotometer (Implen, Munich, Germany). Species identification was

Unedited version

confirmed by genetic screening with GenBank accession number MF457406 using the mitochondrial COI gene. Two paired-end DNA libraries were constructed from the two DNA samples of *M. isos*, using an Illumina HiSeq 2500 Genome Analyser (Sangon Biotech Co. Ltd., Shanghai, China). MISA software (http://pgrc.ipk-gatersleben.de/misa/) was used to search for di-, tri-, tetra-, and pentanucleotide motifs with a minimum of 6, 5, 5, 5 repeats, respectively. MISA was also employed to design primers.

Experimental screening was initially conducted against a subset of *M. isos* (n=10). Primers successfully optimised were subsequently scored for amplification success in a sample of 73 individuals of *M. isos*. Reactions were performed in a 15 μL total volume containing 1.5 μL genomic DNA (~30 ng/μL), 7.5 μL MyTaq Red Mix, 0.6 μL of each primer (10 mmol/L), and 4.8 μL ddH_2O. PCR amplifications were programmed as follows: denaturation at 94 °C for 4 min, followed by 40 cycles of denaturation at 94 °C for 30 s, annealing at primer-specific annealing temperature for 30 s, extension at 72 °C for 45 s, and a final extension at 72 °C for 10 min. Fragment analysis was achieved on an ABI 3730 Genetic Analyzer (Sangon Biotech Co. Ltd., Shanghai, China). Genotyping was performed using PeakScanner v2.0 (Applied Biosystems).

The genotype data were assessed using Lositan (Antao *et al*. 2008) to identify outlier loci. Microchecker v2.2.3 (van Oosterhout *et al*. 2004) was employed to detect the presence of null alleles, large allele dropout and scoring errors, as well as calculating null allele frequency (*r*). Genetic diversity statistics including the number of alleles (N_A), effective number of alleles (N_E), expected (H_E) and observed (H_O) heterozygosity, Polymorphism Information Content (PIC) values and inbreeding coefficient (F_{IS}) were calculated using GenAlEx v6.5 (Peakall and Smouse 2012) and Cervus v3.03 (Kalinowski *et al*. 2007). Deviations from Hardy–Weinberg equilibrium (HWE) (Guo and Thompson 1992) and linkage disequilibrium were generated with Arlequin v3.5 (Excoffier and Lischer 2010) and Genepop v4.2 (Raymond and Rousset 1995; Rousset 2008). The False Discovery Rate (FDR) correction (Benjamini and Hochberg 1995) was applied to assess statistical significance ($p < 0.05$).

The utility of a final set of loci was assessed by means of cross-amplification across two congeneric species, *M. endeavourae* (n=10) and *M. gracilis* (n=10), using the same PCR amplification protocol as described above.

Results and Discussion

Two genomic libraries produced a total of 156.99 M and 198.40 M raw reads, amongst which 1.32 M and 1.43 M sequences containing putative microsatellite loci were screened, respectively. Dinucleotide repeat motifs (~71%) were the most common, followed by tri- (~18%), tetra- (~9%) and pentanucleotides (~2%). Sixty-four primer pairs containing di-, tri- and tetranucleotide motifs in length from 18 to 27 bp were selected and tested in ten individuals of *M. isos*. Twenty-six microsatellite loci displayed seemingly successful amplification for the initial evaluation. Further polymorphism screening of 73 individuals revealed that six of the 26 loci were monomorphic, the remaining 20 loci were polymorphic with optimal annealing temperature from 50 °C to 58 °C. However, it is presumed that evaluation in larger sample sizes and/or of individuals from populations from a larger spatial scale may reveal additional polymorphic loci. A homology search using the BLAST program showed that none of the 26 sequences was similar to any GenBank sequences (Benson *et al*. 2017). The sequences were then deposited in the National Centre for Biotechnology Information (NCBI) Sequence Read Archive under accession numbers MH649295 to MH649320. Descriptive information for the 20 polymorphic loci is reported in Table 1.

Unedited version

The 20 microsatellite loci showed moderate to high levels of allelic diversity (detailed characteristics summarised in Table 1). No locus was identified as experiencing positive selection according to Lositan. Analysis with Microchecker revealed that eight of the 20 microsatellite loci showed evidence of null alleles, with four loci exhibiting stuttering but no large allele drop out. Microsatellite null alleles are very common in marine invertebrates in general, and crustaceans in particular exhibit relatively high null allele frequencies (e.g., Boyle *et al*. 2013; Nakajima *et al*. 2018). The maximum null allele frequency was relatively low ($r = 0.141$), therefore all loci were retained for subsequent analysis. Ten microsatellite loci showed evidence of heterozygote deficiencies as revealed by positive inbreeding coefficients ($F_{IS} > 0$), indicating a potential departure from random mating. Seven loci had significant homozygote excesses after FDR correction ($p < 0.05$) and were flagged by Microchecker for the presence of null alleles. A complementary test for HWE showed significant deviation at 13 loci after FDR correction ($p < 0.05$). Eight (61.5%) and five (38.5%) of the deviations occurred in the form of deficits and excesses of heterozygotes, respectively. Deviation from HWE can be explained by a variety of different causes, including self-fertilisation, positive assortative mating, null alleles, the Wahlund effect, population history, selection and sex linkage (Waples 2015). However, all of these described causes of heterozygote deficit/excess should affect all loci, rather than just one or a few (Selkoe and Toonen 2006; Waples 2015). Of the 190 locus-pairs tested, only one pair (0.5%) of loci (MI_23-2 and MI_29) showed significant linkage disequilibrium after applying the FDR correction ($p < 0.05$).

The cross-amplification test of the 20 microsatellite loci from *M. isos* in two congeners, *M. endeavourae* and *M. gracilis*, revealed high amplification success (70%) and polymorphism across the loci (60%) (Table 2). Fewer alleles in total were observed in the cross-amplification tests compared with the target species *M. isos*, with only 37 in *M. endeavourae* and 28 in *M. gracilis*. This discrepancy may be due to the small sample sizes, highlighting the fact that larger sample sizes are beneficial for obtaining more genetic variety.

The microsatellite loci described here will be used to assess population genetic structure and genetic connectivity amongst VME-associated populations of NZ squat lobsters. In turn, this information will be used to inform management decisions about the future protection of NZ VMEs.

Acknowledgements Squat lobster specimens were supplied for genetic work by the National Institute of Water and Atmospheric (NIWA) Invertebrate Collection, Wellington, New Zealand and Museum Victoria, Melbourne, Australia. Particular special thanks to Ms. Sadie Mills and Ms. Diana Macpherson of the NIWA Invertebrate Collection and Dr. Anna McCallum of the Museum Victoria, for their diligent assistance with loans. Sample collections were supported by funding from the former New Zealand Foundation for Research, Science and Technology, former New Zealand Ministry of Fisheries, Land Information New Zealand, Department of Conservation (New Zealand), GNS Science (New Zealand), Auckland University and Woods Hole Oceanographic Institute (USA). Dr. Ashley Rowden of NIWA is thanked for his comments on the manuscript. This work was supported by funding from Victoria University of Wellington to J.P.A.G (SB80802).

References

Ahyong S. T., Schnabel K. E., Baba K. 2015 Southern high latitude squat lobsters: Galatheoidea and Chirostyloidea from Macquarie Ridge with description of a new species of *Uroptychus*. *Rec. Aust. Mus.* **67**, 109–128.

Unedited version

Antao T., Lopes A., Lopes R. J., Beja-Pereira A., Luikart G. 2008 LOSITAN: A workbench to detect molecular adaptation based on a *Fst*-outlier method. *BMC Bioinformatics* **9**, 323.
Baba K., Macpherson E., Poore G. C. B., Ahyong S. T., Bermudez A., Cabezas P. *et al*. 2008 Catalogue of squat lobsters of the world (Crustacea: Decapoda: Anomura-families Chirostylidae, Galatheidae and Kiwaidae). *Zootaxa* **1905**, 1–220.
Bahassi E. M., Stambrook P. J. 2014 Next-generation sequencing technologies: breaking the sound barrier of human genetics. *Mutagenesis* **29**, 303–310.
Benjamini Y. and Hochberg Y. 1995 Controlling the false discovery rate, a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series. B Stat. Methodol*. **57**, 289–300.
Benson D. A., Cavanaugh M., Clark K., Karsch-Mizrachi I., Ostell J., Pruitt K. D. *et al*. 2017 GenBank. *Nucleic. Acids. Res*. **46**, D41–D47.
Boyle E. A., Thaler A. D., Jacobson A., Plouviez S., Van Dover C. L. 2013 Characterization of 10 polymorphic microsatellite loci in *Munidopsis lauensis*, a squat-lobster from the southwestern Pacific. *Conserv. Genet. Resour*. **5**, 647–649.
Coykendall D. and Morrison C. 2013 Polymorphic microsatellite markers developed from 454 pyro-sequencing in the cold water coral-associated squat lobster species, *Eumunida picta* (Chirostylidae: Eumunididae). *Conserv. Genet. Resour*. **5**, 495–498.
Excoffier L. and Lischer H. E. 2010 Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour*. **10**, 564–567.
Frankham R., Ballou J. D., Briscoe D. A. 2002 *Introduction to conservation genetics*. Cambridge University Press, Cambridge.
Guo S. W., Thompson E. A. 1992 Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**, 361–372.
Jenkins T. L. and Stevens J. R. 2018 Assessing connectivity between MPAs: Selecting taxa and translating genetic data to inform policy. *Mar. Policy*. **94**, 165–173.
Kalinowski S. T., Taper M. L., Marshall T. C. 2007 Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol*. **16**, 1099–1106.
Molecular Ecology Resources Primer Development Consortium: Andris M., Aradottir G. I., Arnau G., Audzijonyte A., Bess E. C, Bonadonna F. *et al.* 2010 Permanent genetic resources added to molecular ecology resources database 1 June 2010 - 31 July 2010. *Mol. Ecol. Resour*. **10**, 1106–1108.
Nakajima Y., Shinzato C., Khalturina M., Nakamura M., Watanabe H. K., Nakagawa S. *et al*. 2018 Isolation and characterization of novel polymorphic microsatellite loci for the deep-sea hydrothermal vent limpet, *Lepetodrilus nux*, and the vent-associated squat lobster, *Shinkaia crosnieri*. *Mar. Biodivers*. **48**, 677–684.
Peakall R. and Smouse P. E. 2012 GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537–2539.
Raymond M. and Rousset F. 1995 GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered*. **86**, 248–249.
Roterman C. N., Copley J. T., Linse K. T., Tyler P. A., Rogers A. D. 2013 Development of polymorphic microsatellite loci for three species of vent-endemic megafauna from deep-sea hydrothermal vents in the Scotia Sea, Southern Ocean. *Conserv. Genet. Resour*. **5**, 835–839.
Rousset F. 2008 Genepop'007: a complete re-implementation of the Genepop software for Windows and Linux. *Mol. Ecol. Resour*. **8**, 103–106.
Schnabel K. E. 2009 *Squat Lobsters (Crustacea: Anomura: Galatheidae and Chirostylidae) of New Zealand: Diversity, Phylogeny and Biogeography*. Dissertation. University of Otago.
Selkoe K. A. and Toonen R. J. 2006 Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol. Lett*. **9**, 615–629.
van Oosterhout C., Hutchinson W. F., Wills D. P., Shipley P. 2004 MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Resour*. **4**, 535–538.
Waples R. S. 2015 Testing for Hardy-Weinberg proportions: have we lost the plot? *J. Hered*. **106**, 1–19.
Zeng C., Rowden A. A., Clark M. R., Gardner J. P. A. 2017 Population genetic structure and connectivity of deep-sea stony corals (Order Scleractinia) in the New Zealand region: Implications for the conservation and management of vulnerable marine ecosystems. *Evol. Appl*. **10**, 1040–1054.

Unedited version

Table 1 Characterisation of twenty microsatellite loci from 73 specimens of *Munida isos*.

Locus	GenBank Accession No.	Primer sequences (5'–3')	5' Dye	Repeat motif	Size range (bp)	N_A / N_E	H_O / H_E	PIC	F_{IS}	r
MI_01	MH649295	F: AAAACACACACAACAATCACAGAA R: AAAATGACGCTCATGATCCC	FAM	$(CA)_6$	94–112	5/3.372	0.469/0.722	0.65	0.330*	0.141
MI_02	MH649296	F: ACGTATTTGTGCGAGATGCT R: TCAAACGCACACACAGATGA	HEX	$(TG)_6$	89–121	12/4.095	0.616/0.777	0.753	0.181*	0.092
MI_04	MH649297	F: ATACACGATCACCTGCACGA R: TTCATGTGAATGCGTGTGTG	FAM	$(AC)_8$	132–156	10/4.494	0.725/0.801	0.764	0.062	0.037
MI_05	MH649298	F: ATGGCTGTCTGACGCTACCT R: TTGTAGTGGACTTTTCCGGC	HEX	$(CT)_9$	88–132	15/5.46	0.729/0.836	0.810	0.097*	0.058
MI_07	MH649300	F: CCACCACGGAAAGGTTCATA R: GAGATACCGTAGTGGACTGCC	TAMRA	$(CT)_8$	89–121	16/7.086	0.75/0.887	0.871	0.127*	0.068
MI_09	MH649301	F: CGGCACTGAATGCAGTACTTT R: ATTTAACCCGGGGACGAATC	TAMRA	$(AC)_9$	97–105	5/1.946	0.519/0.492	0.440	−0.100	−0.016
MI_10	MH649302	F: GACGTATGTATACGCCACAGAGA R: CGTGTGTAGGCTGTCCCTCT	FAM	$(TA)_7$	90–104	6/2.523	0.623/0.608	0.572	−0.073*	−0.018
MI_14	MH649303	F: GTCAAGGCATTGCAAGTCAA R: TGGTGGTGGTGTGTCTGC	FAM	$(CA)_9$	102–116	8/2.351	0.579/0.59	0.525	−0.006*	−0.011
MI_17	MH649304	F: TCCTTGAGGCTCACTTGGAT R: GGCAGGATCAGATTGGACAT	FAM	$(TG)_6$	182–194	4/2.039	0.969/0.526	0.394	−0.904*	−0.296
MI_19–2	MH649305	F: TAAGTGGGTAAATAAGGACG R: CCATCTACAGAAATATATACCATC	HEX	$(ATG)_6$	81–108	10/2.604	0.652/0.635	0.574	−0.067*	0.011
MI_20	MH649306	F: CAATAATGGTGTGTGAGTTGTGA R: TGATGAGGATGACGAGGTGA	FAM	$(TGG)_7$	88–115	8/3.05	0.739/0.69	0.626	−0.104	−0.015
MI_21	MH649307	F: CAATGTTATTTCAAAGTCTTCTCCAA R: TGCGTAAAGCAGGGTTATGA	HEX	$(CAA)_6$	89–110	3/1.016	0.016/0.016	0.027	−0.024	0.000
MI_23–2	MH649308	F: CTCCTTACATAATTTACTTCCTTCC R: TTGATGGTTGCTAACAGACTACA	TAMRA	$(ATT)_7$	100–121	7/1.621	0.261/0.384	0.385	0.254*	0.093
MI_25	MH649309	F: AAGCAGGGAAGAGCACCAT R: CTCTCTGTCTGCCCTCCCAG	FAM	$(AACA)_5$	90–114	4/2.113	0.884/0.543	0.420	−0.684*	−0.248
MI_29	MH649310	F: ATCACCTCCGTTTAGCGTGT R: TTTGGTAGCAAAAGGTTTAAAATG	HEX	$(GTCT)_5$	90–106	5/1.189	0.103/0.157	0.174	0.205*	0.059
MI_31	MH649311	F: CCTTAAAATCGGGAAAACTCC R: TGTACGGTAGAGTCTGGAATAAAGG	FAM	$(GATA)_5$	90–130	9/1.835	0.378/0.437	0.409	0.088	0.050
MI_39	MH649312	F: AAAATAATTCAGTATCCTGGAGCG R: ATAAAGGCAAACACCCCACA	TAMRA	$(GTA)_7$	93–111	5/1.169	0.151/0.143	0.114	−0.074	−0.005
MI_40	MH649313	F: AACACTTTGTTCTTCGTGTGGA R: TCTCCTCCTCCTACTCCTTTCA	FAM	$(GAG)_5$	101–107	3/1.086	0.078/0.073	0.053	−0.088	−0.001
MI_51	MH649317	F: ATGATCCCTTCGACCCTTG R: TGGTTAGACTGAAGAGTTGGCA	HEX	$(CAT)_6$	87–108	8/1.771	0.353/0.422	0.445	0.111*	0.064
MI_52	MH649318	F: ATTACAGACTCGCCGCTCAC R: TTCTCACCATCACCACTATTTCA	TAMRA	$(TTG)_5$	110–125	6/2.896	0.515/0.669	0.638	0.205*	0.095

N_A number of alleles; N_E number of effective alleles; H_O observed heterozygosity; H_E expected heterozygosity; PIC polymorphism information content; F_{IS} inbreeding coefficient; r null allele frequency.

Unedited version

* Significant departure from Hardy–Weinberg equilibrium, after FDR correction ($p < 0.05$).

Table 2 Cross-amplification of fourteen microsatellite loci in 10 specimens of *Munida endeavourae* and 10 specimens of *Munida gracilis*.

Locus	*M. endeavourae*			*M. gracilis*		
	Successful amplifications (n)	N_A	Size range (bp)	Successful amplifications (n)	N_A	Size range (bp)
MI_01	10	4	92–98	10	3	90–96
MI_02	10	5	99–113	10	3	87–101
MI_04	—	—	—	10	3	100–106
MI_09	10	3	98–102	—	—	—
MI_10	10	1	93	10	1	93
MI_14	—	—	—	10	1	119
MI_17	10	1	192	10	5	188–200
MI_20	10	3	87–93	10	2	84–87
MI_23–2	10	4	127–151	—	—	—
MI_25	—	—	—	10	4	95–115
MI_29	10	4	122–146	10	3	100–108
MI_39	10	2	99–105	—	—	—
MI_40	10	4	98–107	10	2	101–104
MI_51	10	6	90–108	10	1	99

N_A number of alleles.

Unedited version