Author/s:
Miller, AD;Van Rooyen, A;Sweeney, OF;Whiterod, NS;Weeks, AR

Title:
The development of 10 novel polymorphic microsatellite markers through next generation sequencing and a preliminary population genetic analysis for the endangered Glenelg spiny crayfish, Euastacus bispinosus

Date:
2013-01-01

Citation:
Miller, A. D., Van Rooyen, A., Sweeney, O. F., Whiterod, N. S. & Weeks, A. R. (2013). The development of 10 novel polymorphic microsatellite markers through next generation sequencing and a preliminary population genetic analysis for the endangered Glenelg spiny crayfish, Euastacus bispinosus. Molecular Biology Reports, 40 (7), pp.4415-4419. https://doi.org/10.1007/s11033-013-2531-5.

Persistent Link:
https://hdl.handle.net/11343/282644

The development of 10 novel polymorphic microsatellite markers through next generation sequencing and a preliminary population genetic analysis for the endangered Glenelg spiny crayfish, *Euastacus bispinosus*

Adam D. Miller[1, 2], Anthony Van Rooyen[1], Oisín F. Sweeney[4], Nick S. Whiterod[5], Andrew R. Weeks[1, 3]

[1]***cesar**, 293 Royal Parade, Parkville 3052, Australia*

[2]*Department of Zoology, The University of Melbourne, Parkville, Victoria, 3010 Australia*

[3]*Department of Genetics, The University of Melbourne, Parkville, Victoria, 3010 Australia*

[4]*Department of Environment, Water and Natural Resources, Mount Gambier, South Australia, 5290 Australia*

[5]*Aquasave - Nature Glenelg Trust, Goolwa Beach, South Australia, 5214 Australia*

*Corresponding author: Adam D. Miller (admiller@unimelb.edu.au)

Keywords: *Euastacus bispinosus*, microsatellite library development, Next Generation DNA sequencing, genetic diversity, species conservation

Abstract

The Glenelg spiny crayfish, *Euastacus bispinosus*, is an iconic freshwater invertebrate of south eastern Australia and listed as ‘endangered’ under the *Environment Protection and Biodiversity Conservation Act 1999*, and ‘vulnerable’ under the International Union for Conservation of Nature’s Red List. The species has suffered major population declines as a result of over-fishing, low environmental flows, the introduction of invasive fish species and habitat degradation. In order to develop an effective conservation strategy, patterns of gene flow, genetic structure and genetic diversity across the species distribution need to be clearly understood. In this study we develop a suite of polymorphic microsatellite markers by Next Generation Sequencing. A total of 15 polymorphic loci were identified and 10 characterized using 22 individuals from the lower Glenelg River. We observed low to moderate genetic variation across most loci (mean number of alleles per locus = 2.80; mean expected heterozygosity = 0.36) with no evidence of individual loci deviating significantly from Hardy-Weinberg equilibrium. Marker independence was confirmed with tests for linkage disequilibrium, and analyses indicated no evidence of null alleles across loci. Individuals from two additional sites (Crawford River, Victoria; Ewens Ponds Conservation Park, South Australia) were genotyped at all 10 loci and a preliminary investigation of genetic diversity and population structure was undertaken. Analyses indicate high levels of genetic differentiation among sample locations (F_{ST} = 0.49), while the Ewens Ponds population is genetically homogeneous, indicating a likely small founder group and ongoing inbreeding. Management actions will be needed to restore genetic diversity in this and possibly other at risk populations. These markers will provide a valuable resource for future population genetic assessments so that an

effective framework can be developed for implementing conservation strategies for *E. bispinosus*.

Introduction

The Glenelg Spiny Freshwater Crayfish, *Euastacus bispinosus* Clark, 1936, is an iconic long-lived freshwater parastacid crustacean species endemic to the Glenelg, and Fitzroy river catchments in south-western Victoria and several spring-fed coastal streams in south-eastern South Australia [10]. Population declines over the past century have been observed across the entire species distribution as a result of over-fishing, low environmental flows, the introduction of invasive fish species and habitat degradation [5]. *Euastacus bispinosus* is currently listed as 'endangered' under the *Environment Protection and Biodiversity Conservation Act 1999*, and 'vulnerable' under the International Union for Conservation of Nature's Red List. Species conservation is dependent on the restoration of critical habitat, and research that will enhance our understanding of the species' biology, ecology and genetics.

In order to develop an effective conservation strategy, patterns of gene flow, genetic structure, and genetic diversity across the species distribution need to be established. A comprehensive population genetic study will identify patterns of dispersal and recruitment and assist in identifying isolated, self-recruiting populations that require independent management consideration. Estimates of genetic diversity will also help identify potentially resilient and vulnerable populations, thereby assisting in the prioritisation and guidance of management investments (including translocation activities that improve genetic diversity and reduce the negative effects of random genetic drift and inbreeding). Contemporary estimates of genetic diversity at the population level will also provide valuable baseline data for monitoring the health of *E. bispinosus* populations across generations.

Here we report the development and characterization of novel microsatellite markers for *E. bispinosus*, and demonstrate their effective utility by conducting a preliminary assessment of genetic diversity in Victoria and South Australia. Microsatellite libraries have been developed previously for other freshwater crayfish species [4,6], including Australian natives [1,7,15], to assist conservation management. However this is the first marker set developed for the genus *Euastacus*, one of the largest, most diverse, and most threaten groups of crayfish species in Australia [5].

Methods and Results

The 454 next generation sequencing platform was used to identify microsatellite markers for *E. bispinosus*. Approximately 10 µg of genomic DNA was extracted from muscle tissue from a single *E. bispinosus* specimen using a QIAGEN DNA Easy kit (Qiagen). DNA was subsequently processed by the Australian Genome Research Facility (AGRF) where it was nebulized, ligated with 454 sequencing primers and tagged with a unique oligo sequence allowing sequences to be separated from pooled species DNA sequences using post-run bioinformatic tools. The DNA sample was analyzed using high throughput DNA sequencing on 1/16 of a 70 x 75 mm PicoTiterPlate using the Roche GS FLX (454) system [8]. A total of 103,385 reads were obtained from the analysis, from which 3207 unique sequence contigs possessing microsatellite motifs were identified using the software GDD [9]. Primer3

[13] was used to design optimal primer sets for each unique contig where possible, with a total of 2,385 contigs found to possess optimal priming sites. A selection of 40 contigs was used for subsequent analysis, 30 of which contained di-nucleotide repeats, 9 containing tri-nucleotide repeats, and 1 containing a tetra-nucleotide repeat.

Loci were screened for polymorphism using template DNA from eight individuals from the Lower Glenelg River in south western Victoria (-38.0544°S, 141.2711°E) and eight individuals from the Ewens Ponds Conservation Park in south eastern South Australia (-38.0264°S, 140.7903°E). Loci were pooled into ten groups of four, labelled with unique fluorophores (FAM, NED, VIC, PET) and co-amplified by multiplex PCR using a Qiagen multiplex kit (Qiagen) and an Eppendorf Mastercycler *S* gradient PCR machine following the protocol described by Blacket et al. [3]. Genotyping was subsequently performed using an Applied Biosystems 3730 capillary analyzer (AGRF, Melbourne, Australia) and product lengths were scored manually and assessed for polymorphisms using GeneMapper version 4.0 (Applied Biosystems). From a total of 40 loci, 15 were found to be polymorphic (Table 1), 16 were monomorphic and 9 failed to amplify. Interestingly, none of the 40 loci screened were polymorphic in Ewens Ponds individuals.

The 15 polymorphic loci were pooled into four groups for multiplexing based on observed locus specific allele size ranges (Table 1), and further characterized using 22 individuals from the lower Glenelg River. Microsatellite profiles were again examined using GeneMapper version 4.0 and alleles were scored manually. The *Excel Microsatellite Toolkit* [11] was then used to estimate expected (H_E) and observed (H_O) heterozygosities and number of alleles (N_A), while examination of conformation to

Hardy-Weinberg equilibrium (HWE), the inbreeding coefficient (F_{IS}) and linkage disequilibrium estimates between all pairs of loci were conducted using GENEPOP version 4 [12]. Significance values were adjusted for multiple comparisons using Bonferroni corrections where necessary [14]. Finally, all loci were assessed using MICRO-CHECKER to check for null alleles and scoring errors [17]. The frequency of null alleles per locus was obtained using the 'Brookfield 1' formula as evidence of null homozygotes across loci was not observed [2].

The majority of loci were characterized by low to moderate genetic variation, with an average of 2.80 alleles per locus (range = 2 – 4 alleles) and heterozygosity estimates ranging between 0.17 and 0.69 (mean = 0.36). Linkage disequilibrium analyses confirmed maker independence indicating no evidence of significant linkage between loci, and 10 loci were found to conform with HWE expectations with estimates of F_{IS} indicating no significant evidence of heterozygote excess or deficit (Tables 2 and 3). Those loci found deviating significantly from HWE expectations included markers EB2, EB14, EB16, EB28 and EB35. MICRO-CHECKER analyses confirmed these loci are potentially influenced by null alleles.

For comparative purposes we genotyped an additional 24 individuals from the Crawford River in south western Victoria (-37.9331°S, 141.5150°E), and an additional 22 from Ewens Ponds Conservation Park. Population statistics were performed for all 15 polymorphic loci following the methods outlined above and summary statistics are provided in Tables 2 and 3. Again, all individuals from Ewens Ponds were monomorphic at each locus, however genotypes from the Crawford River provided consistent results suggesting that markers EB2, EB14, EB16, EB28 and EB35 are problematic (significant HWE and F_{IS} estimates, and evidence of null alleles). To avoid possible biases these loci were excluded from further genetic analysis.

To demonstrate the utility of the microsatellite loci, a preliminary population genetic analysis was conducted. FSTAT was used to calculate global and population pairwise F_{ST} (with 95% confidence limits) (Weir & Cockerham 1984) as a preliminary measure of gene flow between each of the three sample locations. Estimates of F_{ST} across all loci were high and significantly different from zero (global $F_{ST} = 0.49$; 95% CI = 0.34 - 0.59; pairwise $F_{ST} = 0.13$ (Lower Glenelg/Crawford), 0.65 (Lower Glenelg/Ewens Ponds), 0.60 (Crawford/Ewens Ponds)) indicating gene flow and genetic structuring is likely to be limited amongst sample locales. Genetic diversity estimates indicate that the Victoria populations are characterized by low to moderate genetic diversity while the Ewens Ponds population is genetically invariable across all loci indicating that this population has undergone a severe founder event or bottleneck, followed by continued inbreeding.

Discussion

The 10 characterized microsatellite markers described in this study provide a valuable resource for future population genetic assessments of *E. bispinosus* in Australia. Estimates of population structure, gene flow, and genetic diversity will provide an effective framework for implementing management strategies that will maximize the environmental resilience of wild populations. Estimates of genetic diversity within populations will enable at risk populations to be identified (populations with low genetic diversity and/or significant inbreeding) so that appropriate management actions can be instigated. These estimates will also provide managers with baseline data for ongoing monitoring of *E. bispinosus* populations and assessments of the efficacy of management strategies.

Our preliminary population genetic investigation demonstrates the effective utility of these genetic markers for future studies. Genetic estimates indicate individuals from each of the three sample locations are highly differentiated, although more comprehensive sampling is required in order to draw definite conclusions about patterns of gene flow and population structuring. Loci across populations appear to be generally characterized by low to moderate genetic diversity, comparatively lower than estimates derived for some freshwater crayfish species [1,4,7,15], but highly consistent with others [6]. This suggests that perhaps some crayfish species are inherently less polymorphic due to biological effects, as has been observed in other animal groups [16]. However, drastic population declines are likely to have had some influence on levels of genetic diversity in *E. bispinosus*. Our results indicate the Ewens Ponds Conservation Park population is extremely inbred and at high risk of extinction. Population resilience ultimately relies on adequate population sizes to maintain and promote genetic diversity in order to adapt to environmental change. Therefore management actions, such as translocations and captive rearing, may be required to restore genetic diversity and adaptive potential [18] in this and other at risk populations.

Acknowledgements

The fieldwork for this project was partially funded by the Norman Wettenhall Foundation and grants from the federal governments Caring for Our Country program, awarded to the Friends of Mt Gambier Area Parks. Thanks to Dave Mossop, Lauren Kivisalu and volunteers from the Friends of Mount Gambier Area Parks and the wider community for assistance whilst collecting samples. All sample collection was conducted under DENR Wildlife Research permit E25963-1; PIRSA SA Fisheries Exemption permit 9902414; Victorian Fisheries Act 1995 permit RP1092; Victorian Natural Parks Act 1975 and Flora and Flora Guarantee Act 1988 permit 10006431; and Victorian DPI Animal Ethics permit 21.12. Thanks to the Department of Environment, Water and Natural Resources South Australia, the South East Water Conservation and Drainage Board, Parks Victoria and several private landholders in South East South Australia for access to sites.

References

1. Baker N, Byrne K, Moore S, Mather P (2000) Characterization of microsatellite loci in the redclaw crayfish, *Cherax quadricarinatus*. Molecular Ecology 9(4): 494-495

2. Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Mol Ecol 5: 453-455

3. Blacket MJ, Robin C, Good RT, Lee SF, Miller AD (2012) Universal primers for fluorescent labeling of PCR fragments--an efficient and cost-effective approach to genotyping by fluorescence. Mol Ecol Res. 12: 456-63

4. Dong-Ha A, Mi-Hyun P, Jae-Ho J, Mi-Jin O, Sanghee K, Jongwoo J, Gi-Sik M (2011) Isolation and characterization of microsatellite loci in the Korean crayfish, *Cambaroides similis* and application to natural population analysis. Animal Cells and Systems 15(1): 37-43

5. Furse JM, Coughran J (2011) An assessment of the distribution, biology, threatening processes and conservation status of the freshwater crayfish, genus *Euastacus* (Decapoda: Parastacidae) in continental Australia. II. Threats, conservation assessments and key findings. *Crustaceana Monographs: New Frontiers in Crustacean Research* 253-263

6. Gouin N, Grandjean F, Souty-Grosset C (2006) Population genetic structure of the endangered crayfish *Austropotamobius pallipes* in France based on microsatellite variation: biogeographical inferences and conservation implications Freshwater Biology 51: 1369-1387

7. Imgrund J, Groth D, Wetherall, J (1997) Genetic analysis of the freshwater crayfish *Cherax tenuimanus*. Electrophoresis 18(9): 1660-1665

8. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437: 376–380

9. Meglécz E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, Martin JF (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. Bioinformatics 26(3): 403-404

10. Morgan G J (1986) Freshwater crayfish of the genus *Euastacus* Clark (Decapoda, Parastacidae) from Victoria. Memoirs of the Museum of Victoria 47(1): 1-57

11. Park SDE (2001) *Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection*. PhD Thesis, University of Dublin

12. Raymond M, Rousset F (1995) An exact test for population differentiation. Evolution 49: 1280-1283

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

14. Rice WR (1989) Analyzing tables of statistical tests. Evolution 43(1): 223–225

15. Sherman CDH, Ierodiaconou D, Stanley AM, Weston K, Gardner MG, Schultz MB (2012) Development of twenty-four novel microsatellite markers for the

freshwater crayfish, *Geocharax gracilis*, using next generation sequencing. Conservation genetics resources 4(3): 555-558

16. Sivasunda A, Hey J (2003) Population genetics of *Caenorhabditis elegans*: the paradox of low polymorphism in a widespread species. Genetics 163(1): 147-157.

17. 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(3): 535–538

18. Weeks AR, Sgro CM, Young AG, Frankham R, Mitchell NJ, Miller KA, Byrne M, Coates DJ, Eldridge MDB, Sunnucks P, Breed MF, James EA and Hoffmann AA (2011) Assessing the benefits and risks of translocations in changing environments: a genetic perspective. Evol. Appl. (4): 709-725

Table 1. Primers sequences and characteristics of 15 polymorphic microsatellite loci isolated from *E. bispinosus*. Primer combinations for each multiplex reaction are provided. N_A = number of alleles. * loci were problematic and were not characterized further.

Locus	Primer sequences (5’-3’)	Repeat motif	N_A	Size range (bp)	GenBank Accession
Multiplex 1					
EB5	ACCAACATCTGTCGGCTCAT TGTTCTAGCCCACATCCTGG	AG	5	160-170	KC291356
EB6	TTTCGTGCGCATAAACACC GGTTCCAGATACTGGAGTGTCAA	AAT	4	101-116	KC291357
EB14*	TATTCATCCGCCAAATGCTT CGGTTTAGTGAAGTCGGTGC	ACC	8	114-141	
EB25	AGTGTACGGAAGGAGGGTCG TAAGCCATCCATCACCTTCG	AC	2	157-159	KC291361
EB31	GTGCTTGGATAAGCTCGCAT GGTAGGTGGCGGTGATGAG	ACC	2	186-198	KC291364
Multiplex 2					
EB2*	TCACAAGAGGCGAGGAGATG ATGGCGTAGTCTCCGACAGC	AG	5	178-188	
EB13	CGTCTTCCTCCCTACGGGT AACCTTCAGAAACCCTTAGCTCC	AAT	5	110-122	KC291358
EB16*	GCTGTTGACGCTTCCTCTGT GGCGTTGACCACGTTGAT	ACG	3	198-227	
Multiplex 3					
EB17	GTTCCTGCAAACCATTCCTG CCTCTTCCTCCTCATCCACC	AGG	3	161-176	KC291359
EB27	AACTTGTAGCCTCAGCAGCC TTCGAAAGTTGAGGTCCGAG	AGC	2	116-122	KC291362
EB35*	TTTGGGAAACAATGAGAGCA TGGACACTGACGCTGCTTTA	AG	2	141-156	
Multiplex 4					
EB18	CAATTCCCTTAAACCCGACC TCTTCTCCTCGGTTATCGTCC	AG	3	237-241	KC291360
EB28*	AGCGAAATGCTGTCCATGTT CAGCCTGGTAAATGCAGGAT	ACAT	6	149-169	
EB30	TTACGGGCTTAGTGCTTCCC CTAGCCAAGCGAGCCATTTA	AAT	2	134-137	KC291363
EB40	CAAACTCAGGAAGATAATTCAAATAGG ATTTGTAATACAGGTGCAGAGCTA	AC	3	110-118	KC291365

Table 2. Statistics for *Euastacus bispinosus* populations screened with 10 polymorphic microsatellite loci. Mean values over loci are presented for number of alleles (N_A), expected (H_E) and observed (H_O) heterozygosities, Hardy-Weinberg equilibrium P values, and the inbreeding coefficient (F_{IS})

Population	N	Loci	N_A	H_E	H_O	HW P-value	F_{IS}
Lower Glenelg	22	10	2.80	0.37	0.39	0.79	-0.02
Crawford	24	10	2.50	0.36	0.34	0.34	0.04
Ewens Ponds	30	10	1	0	0	-	-

Table 3. Statistics for *Euastacus bispinosus* populations screened with 10 polymorphic microsatellite loci. Mean values per locus are presented for number of alleles (N_A), expected (H_E) and observed (H_O) heterozygosities, Hardy-Weinberg equilibrium *P* values, and the inbreeding coefficient (F_{IS}).

		EB25	EB31	EB5	EB6	EB13	EB17	EB27	EB18	EB30	EB40
Lower Glenelg	N_A	2	3	3	4	4	3	2	2	2	3
	H_E	0.17	0.37	0.44	0.38	0.69	0.53	0.30	0.17	0.46	0.17
	H_O	0.18	0.45	0.50	0.32	0.68	0.59	0.36	0.18	0.50	0.18
	HW *P*-value	1.00	0.62	1.00	0.17	0.48	0.52	1.00	1.00	1.00	1.00
	F_{IS}	-0.07	-0.23	-0.13	0.06	0.02	-0.12	-0.20	-0.07	-0.09	-0.06
Crawford	N_A	2	2	3	2	5	2	2	3	2	2
	H_E	0.47	0.28	0.46	0.16	0.64	0.42	0.49	0.23	0.34	0.12
	H_O	0.29	0.33	0.38	0.08	0.65	0.50	0.46	0.21	0.42	0.12
	HW *P*-value	0.08	1.00	0.23	0.13	0.79	0.62	1.00	0.11	0.54	1.00
	F_{IS}	0.38	-0.18	0.20	0.47	-0.02	-0.19	0.06	0.09	-0.24	-0.05
Ewens Ponds	N_A	1	1	1	1	1	1	1	1	1	1
	H_E	0	0	0	0	0	0	0	0	0	0
	H_O	0	0	0	0	0	0	0	0	0	0
	HW *P*-value	-	-	-	-	-	-	-	-	-	-
	F_{IS}	-	-	-	-	-	-	-	-	-	-