MOLECULAR ECOLOGY
RESOURCES

Isolation and Characterization of 27 Polymorphic Microsatellite loci for Maritrema novaezealandensis, a marine trematode from New Zealand

Journal:	*Molecular Ecology Resources*
Manuscript ID:	MER-09-0062.R1
Manuscript Type:	Permanent Genetic Resources Note
Date Submitted by the Author:	16-Mar-2009
Complete List of Authors:	Springer, Yuri; University of Otago, Zoology Koehler, Anson; University of Otago, Zoology Poulin, Robert; University of Otago, Zoology
Keywords:	Evolutionary Theory, Host Parasite Interactions, Invertebrates, Life History Evolution, Parasitology

TITLE: Isolation and Characterization of 27 Polymorphic Microsatellite loci for *Maritrema novaezealandensis*, a marine trematode from New Zealand

AUTHORS: Springer, YP; Koehler, AV; Poulin, R

ADDRESS: Zoology Department, University of Otago, PO Box 56, Dunedin 9054, New Zealand

KEYWORDS: *Maritrema novaezealandensis*, microsatellite, parasite, shotgun sequencing, trematode

CORRESPONDING AUTHOR: Yuri P. Springer, Zoology Department, University of Otago, PO Box 56, Dunedin 9054, New Zealand. FAX: 64.3.479.7584 EMAIL: yuripringer@gmail.com

RUNNING TITLE: *Maritrema novaezealandensis* microsatellite loci

Abstract:

In this report we describe the development of 27 polymorphic microsatellite markers for the parasitic marine trematode *Maritrema novaezealandensis*. Loci were identified using a shotgun sequencing approach and then screened for polymorphism using 50 parasite clones. The number of alleles per locus ranged from 2 to 14 (mean 5.6) with expected and observed heterozygosities ranging from 0.0586-0.8836 and 0.060-0.980, respectively. These markers collectively provide a robust means to identify host individuals infected with multiple parasite clones and will be used to test hypotheses about the relationship between parasite genetic diversity and host specificity.

Main Text:

Maritrema novaezealandensis (Microphallidae) is a parasitic marine digenean trematode found in nearshore marine ecosystems of New Zealand (Martorelli *et al.* 2004). It has a "complex," multi-host lifecycle in which it utilises the mudsnail *Zeacumantus subcarinatus*

as its first intermediate host, a variety of intertidal and shallow subtidal crustaceans as second intermediate hosts, and the red-billed gull *Larus novaehollandiae scopulinus* as its final host. As part of a study examining the evolution of host specificity of parasites we sought to identify a set of microsatellite markers that could collectively be used to quantitatively estimate genome-wide heterozygosity of individual parasite clones.

M. novaezealandensis tissue (~250 parasite sporocysts), was removed from a single infected *Z. subcarinatus* snail collected at Lower Portobello Bay, Otago Harbor, New Zealand in June, 2008. We isolated genomic DNA from this tissue using a modified CTAB/chloroform extraction (Murray & Thompson 1980). Tissue was rinsed in 1mL extraction buffer, centrifuged at 13000 rpm for 5 minutes, and then resuspended in a mixture of 75 µL of extraction buffer, 250 µL CTAB, 20 µL Proteinase K, and 5 µL RNase. We incubated this solution for 90 min at 65°C and then allowed it to cool to room temperature. A 250 µL aliquot of chloroform/isoamyl alcohol (24:1) was added, the solution was gently mixed for 2 min, centrifuged at 2000 rpm for 5 min, and the supernatant transferred to a new tube. We used an overnight sodium acetate/ethanol reprecipitation to concentrate the DNA, which was eluted in 50 µL of TE buffer to produce our purified template sample. Spectrophotometric analysis of the sample indicated a concentration of 85.6ng/µL and a 260/280 ratio of 1.96. To verify the taxonomic identity of the sample and the absence of contamination by tissue from other species of parasites we used polymerase chain reaction (PCR) to amplify the mitochondrial cytochrome oxidase I gene (COI) and compared the resulting sequence to those associated with other species of trematodes known to infect *Z. subcarinatus*. No parasites other than *M. novaezealandensis* were detected.

To identify microsatellite loci in our purified template sample we analysed results of a shotgun sequencing procedure described by Abdelkrim *et al.* (2009). We performed a 1/8th sequencing run (LR70 plate format) using the Genome Sequencer FLX System (Roche, Penzberg, Germany) (Margulies *et al.* 2005) which yielded a total of 31120 sequences (mean length=229bp, stdev=52bp, max=323bp, min=20bp) that passed the system's quality control analyses. We screened these sequences for microsatellite repeat motifs using the software MSATCOMMANDER (Faircloth 2008), searching for dinucleotide motifs with a minimum of five repeats and tri-, tetra-, penta-, and hexanucleotide motifs with a minimum of four repeats. This yielded 296, 262, 118, 34, and 7 fragments with repeats, respectively (total=717). Using MSATCOMMANDER's inbuilt interface with Primer3 (Rozen & Skaletsky 2000) we were able to design primers to amplify the repeat regions in 208 fragments (64 di-, 104 tri-, 32 tetra-, 9 penta-, and 2 hexanucleotide repeats, with three primer pairs amplifying regions containing two different repeat motifs).

From a haphazardly selected subset of 144 of these 208 loci 27 successfully amplified and were polymorphic. To quantify levels of polymorphism we collected a sample of parasite cercariae from 69 infected *Z. subcarinatus* collected from Lower Portobello Bay, New Zealand and used a chelex extraction (Walsh *et al.* 1991) to obtain DNA from each sample. PCR amplifications were carried out in a mixture containing 5-10 ng of DNA template, 0.5 U *Taq* DNA polymerase (Bioline), 200 μM of each dNTP, 0.6 μM of each primer, 1 mM $MgCl_2$, in 10X NH_4 reaction buffer (160 mM $(NH_4)_2SO_4$, 670 mM Tris-HCL (pH 8.8 at 25°C), 0.1% Tween-20) and made up to a final volume of 10 μL with nanopure water. PCR amplification was performed on an Eppendorf Mastercycler thermal

cycler using the following temperature profile: initial 4 min of denaturation at 94°C, followed by 25 cycles of denaturation (94°C, 1 min), annealing (temperature specified in Table 1, 45 s), and extension (72°C, 45 s), with final extension time of 6 min at 72°C. PCR products were resolved on 8% nondenaturing polyacrylamide gels with a 10-bp DNA ladder. Genotyping results indicated that of the 69 parasite samples (infected host snails), 19 consisted of more than one parasite clone, and these were excluded from further analyses. Based on the genotypes of the remaining 50 single parasite clones, the number of alleles per locus ranged from 2 to 14 (mean 5.59) and expected and observed heterozygosities ranged from 0.0586-0.8836 and 0.060-0.980, respectively. Deviations from Hardy-Weinberg equilibrium (HWE) allele frequencies and tests of linkage disequilibrium were assessed statistically using GenePop version 4.0 (Rousset 2008) with significance levels adjusted using Bonferroni corrections for multiple tests. Seven of the 27 loci showed significant deviations from HWE, with heterozygote deficiencies at six loci and heterozygote excess at one (Table 1). No significant linkage disequilibrium was found in any pair of these loci. Null alleles were detected at nine loci using Micro-Checker (Table 1, Van Oosterhout *et al.* 2004) and were associated with all six loci for which a heterozygote deficiency (relative to HWE) was observed. Together with the seven loci identified in a previous study (Keeney *et al.* 2006), this brings the total number of known polymorphic microsatellite loci for *M. novaezealandensis* to 34.

Acknowledgements:

We thank J. Stanton for performing the sequencing, J. Abdelkrim for help with screening sequences, T. King and T. Leung for assistance in the lab, and L. Hegg and N. Jessen for help scoring genotypes. This research was supported financially by the Marsden fund.

Literature Cited:

Abdelkrim J, Robertson B, Stanton J, Gemmell N (2009) Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques* 46, 185-192.

Faircloth BC (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8, 92-94.

Keeney DB, Waters JM, Poulin R (2006) Microsatellite loci for the New Zealand trematode *Maritrema novaezealandensis*. *Molecular Ecology Notes* 6, 1042-1044.

Margulies M, Egholm M, Altman WE*, et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376-380.

Martorelli SR, Fredensborg BL, Mouritsen KN, Poulin R (2004) Description and proposed life cycle of *Maritrema novaezealandensis* N. sp (Microphallidae) parasitic in red-billed gulls, *Larus novaehollandiae scopulinus*, from Otago Harbor, South Island, New Zealand. *Journal of Parasitology* 90, 272-277.

Murray MG, Thompson WF (1980) Rapid Isolation of High Molecular-Weight Plant DNA. *Nucleic Acids Research* 8, 4321-4325.

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8, 103-106.

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds. S Krawetz, S Misener), 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.

Walsh PS, Metzger DA, Higuchi R (1991) Chelex-100 as a Medium for Simple Extraction of DNA for Pcr-Based Typing from Forensic Material. *BioTechniques* 10, 506-513.

Table 1: Microsatellite loci developed for *Maritrema novaezealandensis*. Locus name is followed by repeat motif present in sequenced fragment, primer sequences, primer annealing temperature (T_A), size range of alleles in base pairs (Range), number of alleles (N_A), number of individuals screened for polymorphism measurements (N_I), observed (H_O) and expected (H_E) heterozygosities and GenBank Accession number (Accession no.). * denotes significant deviation from Hardy-Weinberg equilibrium (P<0.0019) as determined by Genepop. § denotes detection of null alleles by Micro-Checker

Locus	Repeat motif	Primer Sequence	T_A (C°)	Range (bp)	N_A	N_I	H_O	H_E	GenBank Accession Number
Mar13	$(AG)_5$	F: ATACCCTTCGCCCTTCGTC	58	219-209	6	50	0.6200	0.5812	FJ766537
		R: GGCAACCCTTCGTGTTGTG							
Mar19	$(CT)_5$	F: TCCCTCTCGCTTCTTACGC	50	162-158	3	50	0.0600	0.0586	FJ766538
		R: GGTGGGTCGTAATAGCCAATC							
Mar21	$(CA)_6$	F: AGTAGTGGAGTGGCTGACC	50	182-166	6§	50	0.2800	0.5763	FJ766539
		R: GAGCCCTGATGCAAAGCAG							
Mar22	$(GT)_8$	F: CGTCCACTGTTCGAGGTTG	50	194-164	8	49	0.7551	0.8047	FJ766540
		R: GTCCACTGAAACACACTCGTC							
Mar26	$(GT)_5$	F: GGGCCTCATTTCCGTTTGC	50	169-167	2	50	0.2600	0.2822	FJ766541
		R: GCCTCATTGTGAACAGCCC							
Mar34	$(CT)_7$	F: CGTCGTTCTGATCTTGGCTG	50	220-208	4§	50	0.3200	0.5646*	FJ766542
		R: AGACGATGGGGTGACCTTC							
Mar39	$(GT)_8$	F: TTGGAGCGCTAATGGACTG	50	197-171	11	50	0.7600	0.7814	FJ766543
		R: ACAGGAGAAGTGACCACGG							
Mar42	$(CT)_5$	F: GTCTGTAGGTCACTCAGCCG	50	151-143	5§	50	0.4000	0.5436	FJ766544
		R: ACGCGACCACAAACATGAC							
Mar47	$(TTCAAG)_4$	F: CAGTCCTAGGCGTTCGAGG	50	217-211	3§	50	0.2200	0.4486*	FJ766545
		R: TGTACCAAGCGCCTACTCC							
Mar50	$(AGGAC)_5$	F: GTTGCGACCTCGGATCAAG	50	233-183	9	50	0.7800	0.6878	FJ766546
		R: CTCATGTGAACCCTTGCGG							
Mar55	$(GAAA)_4$	F: TCGGTCAACTAAACAGGATTCG	55	222-214	3	50	0.5200	0.4984	FJ766547
		R: AGGAAGAGAAAGCGGCCAG							
Mar64	$(TAGG)_4$	F: TCATAACGGCGTACCCACC	50	209-189	6§	50	0.2400	0.4734*	FJ766548
		R: ACCGGGAGAATCCTTGTGC							
Mar65	$(GGTT)_6$	F: GCCGGCCTCACACAATTC	58	166-158	4	50	0.6200	0.6750	FJ766549
		R: AAAAGGGCGTAAGCAAGCC							
Mar70	$(TGAG)_5$	F: CTGTCAATTCCTTGAGCCCG	50	196-172	7	50	0.5800	0.5252	FJ766550
		R: TCGTCCCAATTCGACTCTTG							
Mar73	$(TAAG)_3$	F: TTTGCGCTGATTTGGCCTC	50	242-186	5	50	0.4600	0.4276	FJ766551
		R: TGTACCACAGTAGCCTCCTTC							
Mar74	$(TAAT)_5$	F: GTGCATTCATTGGTTGAACAGG	50	180-136	9§	50	0.5000	0.63*	FJ766552
		R: ATAAGTGGGGTGAGCTGCG							
Mar80	$(TAA)_3$	F: CATCAACCCACTGCTTCGG	50	196-163	5	50	0.2200	0.2368	FJ766553
		R: TGTCATCACTATGGATCATGCG							
Mar88	$(CAA)_6$	F: AACCAATCGGAAACGCTGG	50	227-197	10§	50	0.5800	0.7792*	FJ766554
		R: GCCTCAAGGTTGGATTCTTGG							
Mar93	$(GAA)_4$	F: CAAGCAACGAGCAGAGTGG	50	169-163	3	50	0.3800	0.4136	FJ766555
		R: TACGCCCGGTTCACGTATG							
Mar102	$(TAA)_3$	F: ACAACAGGGGCACTTGACG	50	155-149	4	50	0.2800	0.2790	FJ766556
		R: CCCATCATATCACTCATGCTTACC							
Mar107	$(AAG)_6$	F: GGAAGCAGATCATAAAATTCACAAGC	50	126-114	5	50	0.5800	0.5570	FJ766557
		R: GGCCTTGGACGGATTTGAC							
Mar111	$(AAG)_6$	F: TGGCGCTTACCTCCAACAC	50	196-181	6	50	0.7200	0.6088	FJ766558
		R: ACTGAAAGAGCCGGTTGTC							
Mar115	$(TTC)_6$	F: CGGAAACAAGTATGATTCGGGG	50	176-173	2	50	0.9800	0.4998*	FJ766559
		R: AGCGAAATGCTTTCAACTGAC							
Mar125	$(TTC)_7$	F: CGGGTCTTCTTCCGGGG	50	211-202	4	50	0.3200	0.3508	FJ766560
		R: GCAATATCCGTCGCCTAGC							
Mar126	$(CTT)_5$	F: AGAATGCTTCGAGGATGTAGG	58	237-177	14§	50	0.6000	0.8836*	FJ766561
		R: GCCTGCAGACAGCGTTTC							
Mar135	$(TAA)_5$	F: TCAGTGGTTTCAGCCTGGG	50	181-175	3§	50	0.2200	0.3694	FJ766562
		R: TGATCATGTACGCATGTCTTGG							
Mar156	$(CA)_6$	F: ACAGTGTTGAGGTCATGGTG	50	150-144	4	50	0.4800	0.4468	FJ766563
		R: TGCTTGATCGTCAAAGTGGC							