Design of microsatellite markers

Microsatellite markers specific to *Philophthalmus* sp. were designed from a genomic library prepared from a Roche 454 shotgun sequencing procedure (Abdelkrim *et al.*, 2009). A *Philophthalmus* sp. infected snail was collected from Lower Portobello Bay, Otago Harbour, South Island, New Zealand (45°52’ S, 170°42’ E) during September 2010. Small and large rediae were dissected from one snail and rinsed three times in autoclaved water to remove all snail tissue. DNA was extracted from pooled rediae. They were placed in 20 µl buffer (0.1M NaCl, 0.05 M Tris-HCl, 0.01 M Na_2EDTA, pH 8.0), 2 µL Tween [20] 2%, and 4 µL Proteinase K (20 mg/ml; Roche #03115879001). The sample was heated for two hours in a 65°C water bath (mixed and spun at 30 minute intervals) and boiled at 95°C for 10 minutes. Samples were then cooled to room temperature and kept frozen (Devlin *et al.*, 2004). The concentration of DNA was measured using a NanoDrop (ND-1000 spectrophotometer) (556.7 ng/ul).

A genomic library was prepared using a Roche 454 shotgun sequencing procedure. One sixteenth of a LR70 plate was run through the Genome Sequencer FLX System, producing 47,299 sequences between 21-713 base pairs long (Roche, Penzberg, Germany) (Margulies *et al.*, 2005). Reads were converted into two separate FASTA files. Sequences were analyzed and filtered for quality using a trial version of Geneious 5.4 (Drummond *et al.*, 2011). Sequences containing repeats were identified using MSATCOMMANDER (Faircloth, 2008), identifying 400 potentially useful repeat motifs. We looked for repeat regions of at least 6X dinucleotide repeats; 4X trinucleotide repeats, and 3X tetranucleotide repeats. From this library of sequences containing repeat regions, those nested within enough readable sequence to design primers were manually identified. Using Primer 3 plus (Untergasser *et al.*, 2007), primer pairs were designed around repeat motifs. All primers were designed to

amplify at similar conditions (melting temperatures between 55-62°C optimum 60°C; GC% 45-55 optimum 50; length 18-20bp). Net Primer (http://www.premierbiosoft.com/netprimer/) was used to determine if primers would form dimers, palindromes, or hairpins. To each forward primer, an 18 base-pair M13(-21) tail was added to the 5' end (TGTAAAACGACGGCCAGT). This, when combined with the fluorescently labeled universal M13(-21) primer (labeled with either FAM, PET, NED, or VIC fluorescent dye) in the PCR reaction, allows for fluorescent labeling of PCR products (Schuelke, 2000). Primers were dyed strategically to allow for multiplexing and analysis in one tube. To each reverse primer, a pigtail sequence was added to the 5' end (GTTTCTT) to inhibit addition of non-template nucleotides to the 3' end of the PCR product (Brownstein *et al.*, 1996). In total, 57 primer pairs were ordered from Sigma.

Initial primer amplification and optimal conditions were tested using DNA extractions from nineteen *Philophthalmus* sp. infections dissected out of snails collected from Lower Portobello Bay on 4 October, 2011. One redia from each infection was used for DNA extraction to ensure single genotypes were being observed. DNA was extracted according to the previously mentioned protocol. PCR reactions contained 15 ng DNA template, 0.45 U DNA Polymerase (Bioline), 800µM of each dNTP, 0.04 µM forward primer, 0.16µM reverse primer, 0.16 µM fluorescently dyed M13(-21) primer, 1.5 mM $MgCl_2$, in 10X NH_4 reaction buffer (160mM $(NH_4)_2SO_4$, 670 mM Tris-HCL (pH 8.8 at 25C), 0.1% Tween-20) and made up to a final volume of 10µL with milliQ water. Products were amplified using a Eppendorph Mastercycler ep gradient S thermocycler as follows: an initial 2 minutes at 94°C, 30 cycles of denaturation (94°C, 30 seconds), annealing (55°C, 45 seconds), and extension (72°C, 45 seconds), followed by 12 cycles to further amplify the dyed primer which included denaturation (94°C, 25 seconds), annealing (53°C, 45 seconds), and extension (72°C, 45 seconds); final extension time was 10 minutes at 72°C followed by 30 minutes at 60°C. To

ensure PCR amplification occurred successfully, products were run through a 1% agarose gel dyed with SYBER Safe DNA gel stain (Invitrogen). Electrophoresis of the amplified products was performed using the ABI 3730xl DNA Analyser (Applied Biosystems, Foster City, CA, USA). Loci were scored using GeneMarker (Softgenetics, LLC, State College, PA, USA) and were useful as markers if they were both polymorphic and highly allelic. Eight primer pairs amplified loci that were polymorphic and highly allelic (Supplementary Table 1). The number of alleles at each locus ranged from three to twelve.

Allele frequencies from this initial screen were used to estimate allele frequencies for the *Philophthalmus* sp. population in Lower Portobello Bay. Allele frequencies were estimated from the results of nineteen individually genotyped rediae from separate colonies. Frequencies ranged from 0.03 to 0.50. Rare alleles were not represented in the subset of individuals used to estimate allele frequencies (for example, alleles 304 and 307 for locus 44). They are included here (Supplementary Table 2) because they were observed in subsequent individuals genotyped with these markers.

The expected and observed heterozygosity of allele frequencies at each locus was determined using GDA version 1.1 (Lewis and Zaykin, 2001). Loci were tested for deviation from Hardy–Weinberg equilibrium (HWE) and genetic disequilibrium was tested for all loci pairs using Fstat version 2.9.3.2 (Goudet, 2002). *p* values were corrected for multiple simultaneous pair wise comparisons using the Bonferroni correction (Rice, 1989). Observed heterozygosity ranged from 0.48 – 0.93 and expected heterozygosity ranged from 0.10 – 0.66. The observed heterozygosity of five of the eight loci were significantly higher than expected, but only four of eight deviated significantly from HWE after using the Bonferroni correction (Supplementary Table 1). Genetic disequilibrium was not found to be significant for any pairs of loci.

Supplementary Table 1

Microsatellite loci for *Philophthalmus* sp. T_a indicates the optimal appealing temperature, H_O indicates the observed heterozygosity, H_E indicates the expected heterozygosity, and * indicates where the observed heterozygosity deviates significantly from the expected

Locus	Repeat Motif	Primer Sequence (5'-3')	T_a	Range (bp)	No. Alleles	$H_O(H_E)$	Dye used to Multiplex
Pbu7	$(TG)_{14}$	**F:** TGTAAAACGACGGCCAGT GATGAACGAGAACCGACACA **R:** GTTTCTT CGTGGAAAACAAACGAACAG	55°C	196-218	12	0.81(0.47)*	VIC
Pbu30	$(AC)_{12}$	**F:** TGTAAAACGACGGCCAGTGGCTTGTTCACCATAGTCGC **R:** GTTTCTTTCGTGTAGTTCTGATGCAATGTG	55°C	154-162	5	0.48(0.45)	VIC
Pbu32	$(AC)_{10}$	**F:** TGTAAAACGACGGCCAGTTGGTGGGCGGTTAGTACTTC **R:** GTTTCTTGCTGCCATGCTTACCAGATC	55°C	118-124	4	0.67(0.66)	PET
Pbu36	$(AC)_{12}$	**F:** TGTAAAACGACGGCCAGTCTGACTGTTCTTGCACACCG **R:** GTTTCTTGTGTGAAACGCTGCATTTCC	55°C	159-171	8	0.85(0.50)*	PET
Pbu43	$(AC)_8$	**F:** TGTAAAACGACGGCCAGTTGTTGCCAAGTCAAGACACC **R:** GTTTCTTGGGATTGTTTCGACCTGAGC	55°C	233-239	3	0.64(0.10)*	FAM
Pbu44	$(ATC)_9$	**F:** TGTAAAACGACGGCCAGTGGTCATGGATGGATGTTCGC **R:** GTTTCTTACGATGGGTTGATGATGCAG	55°C	289-307	8	0.66(0.53)	VIC
Pbu48	$(AC)_{10}$	**F:** TGTAAAACGACGGCCAGTTGAGGGTAGGGCATCAAACG **R:** GTTTCTTGGAATCCGTAGTGAATCAGTCG	55°C	295-305	6	0.77(0.60)	NED
Pbu57	$(AT)_8$	**F:** TGTAAAACGACGGCCAGTTGGCCCAAATATAGACCCGG **R:** GTTTCTTATCGACGGCATAAGGGAAAC	55°C	166-194	10	0.93(0.38)*	NED

Supplementary Table 2

Estimated allele frequencies for *Philophthalmus* sp. microsatellites

Locus	Allele	Frequency	Locus	Allele	Frequency
Pbu7	196	0.08	**Pbu43**	233	0.50
	198	0.04		235	0.35
	200	0.27		237	0.15
	202	0.12	**Pbu44**	289	0.09
	204	0.42		292	0.35
	206	0.04		295	0.47
	208	0.12		298	0.06
	210	0.00		301	0.03
	212	0.08		304	0.00
	214	0.00		307	0.00
	216	0.00	**Pbu48**	295	0.20
	218	0.00		297	0.10
Pbu30	154	0.80		299	0.27
	156	0.05		301	0.37
	158	0.10		303	0.00
	160	0.05		305	0.00
	162	0.05	**Pbu57**	166	0.06
Pbu32	118	0.50		168	0.13
	120	0.27		172	0.00
	122	0.17		174	0.13
	124	0.07		178	0.06
Pbu36	157	0.00		180	0.13
	159	0.04		182	0.00
	161	0.07		184	0.25
	163	0.25		188	0.06
	165	0.07		194	0.06
	167	0.21			
	169	0.11			
	171	0.00			

Supplementary Table 3

Probability that colonies contained mixed genotypes. If probability was less that 10^{-7}, the colony was classified as a single genotype infection.

Colony	Number of homozygous loci	Probability mixed infection	Infection
1	5	9.00279E-12	single
2	4	2.91843E-17	single
3	3	2.04158E-10	single
4	3	3.21127E-06	unknown
5	5	1.13401E-16	single
6	6	2.73887E-13	single
7	2	4.28594E-06	unknown
8	0		unknown
9	4	1.68151E-15	single
10	4	4.49413E-12	single
11	4	3.1117E-18	single
12	0		unknown
13	1	0.0625	unknown
14	3	1.19539E-05	unknown
15	2	0.007676563	unknown
16	2	9.8345E-06	unknown
17	2	0.00000016	unknown
18	0		unknown
19	1	0.01500625	mixed
20	2	0.000937891	mixed
21	5	1.02905E-18	single

22	3	3.7481E-06	unknown
23	4	2.401E-13	single
24	3	3.78229E-10	single
25	4	1.72841E-12	single
26	3	1.75776E-05	unknown
27	3	5.0625E-08	single
28	2	0.019987173	unknown
29	3	3.11521E-07	single
30	0		unknown
31	1	0.4096	unknown
32	3	0.00001296	unknown
33	5	2.42856E-13	single
34	2	5.86182E-05	unknown
35	5	6.69059E-11	single
36	2	0.00614656	unknown
37	3	7.32094E-09	single
38	4	1.5753E-09	single
39	1	0.01500625	mixed
40	2	0.00194481	unknown
41	2	0.019987173	unknown
42	0		mixed
43	0		mixed
44	2	1.04858E-06	unknown
45	1	0.00028561	unknown
46	2	2.42891E-07	single
47	1	0.00028561	mixed
48	5	5.60202E-16	single

49	2	3.7481E-10	single
50	0		mixed
51	4	9.32955E-09	single
52	4	6.14656E-07	single
53	2	5.0625E-08	single
54	0		unknown
55	2	1.89747E-09	single
56	4	4.90184E-11	single
57	3	6.4597E-11	single
58	2	3.603E-07	single
59	0		unknown
60	1	0.0625	mixed
61	2	1.78506E-05	unknown
62	3	1.0972E-11	single
63	0		unknown
64	4	5.7648E-10	single

References

Abdelkrim, J., Robertson, B., Stanton, J. A. and Gemmell, N. (2009). Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques,* **46**, 185-192. doi: 10.2144/000113084.

Brownstein, M. J., Carpten, J. D. and Smith, J. R. (1996). Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *BioTechniques,* **20**, 1004-1006, 1008-1010.

Devlin, C. M., Diamond, A. W. and Saunders, G. W. (2004). Sexing arctic terns in the field and laboratory. *Waterbirds,* **27**, 314-320. doi: 10.1675/1524-4695(2004)027[0314:satitf]2.0.co;2.

Drummond, A. J., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., Heled, J., Kearse, M., Markowitz, S., Moir, R., Stones-Havas, S., Sturrock, S., Thierer, T. and Wilson, A. (2011). Geneious 5.4. http://www.geneious.com/.

Faircloth, B. C. (2008). msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources,* **8**, 92-94. doi: 10.1111/j.1471-8286.2007.01884.x.

Goudet, J. (2002). Fstat, a program to estimate and test gene diversities and fixation indices (version 2.9.3). (Updated from Goudet, J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. Journal of Heredity 86: 485e486). Available at http://www.unil.ch/izea/softwares/fstat.html.

Lewis, P. O. and Zaykin, D. (2001). Genetic Data Analysis: Computer program for the analysis of allelic data.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P.,

Jando, S. C., Alenquer, M. L. I., Jarvie, T. P., Jirage, K. B., Kim, J.-B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F. and Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature,* **437**, 376-380. doi: http://www.nature.com/nature/journal/v437/n7057/suppinfo/nature03959_S1.html.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution,* **43**, 223-225.

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology,* **18**, 233-234.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, R., Geurts, R. and Leunissen, J. A. M. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research,* **35**, W71-W74. doi: 10.1093/nar/gkm306.