1 Primer Note

2

3 **Isolation and characterization of nine microsatellite markers for a galaxiid fish endemic to**

4 **Patagonia**

5

6 Tyler S. Zemlak[1, 2,*], Cecilia Carrea[1,*], Erica Newton[1, 3], Ian Paterson[1], Daniel E. Ruzzante[1]

7

8 **1:** Department of Biology, Dalhousie University, Halifax, NS, Canada B3H4R2

9 **2:** Present address**:** Department of Microbiology & Immunology, Dalhousie University, Halifax,

10 NS, Canada B3H4R2

11 **3:** Present address**:** Environmental and Life Sciences Department, Trent University,

12 Peterborough, ON, Canada K9J 7B8

13

14 *These authors contributed equally to this work.

15

16 **Keywords:** *Galaxias platei*, microsatellite markers, tetranucleotides, compound tetra-

17 dinuclotides.

18

19 Corresponding Author: Daniel Ruzzante (daniel.ruzzante@dal.ca)

20 phone: 1-902-494-1688 fax: 1-902-494-3736

21

22 **Running Title:** Microsatellite isolation for *Galaxias platei.*

23

24 **Abstract**

25 This study describes the isolation and characterization of nine polymorphic microsatellite
26 markers developed specifically for *Galaxias platei* (Pisces: Galaxiidae), a species endemic to
27 Patagonia. Seven tetranucleotide and two compound tetra-dinucleotide markers were tested in
28 126 individuals from two different localities. The observed and expected heterozygosities per
29 locus ranged between 0.23-0.94 and 0.27-0.92, respectively. The number of alleles per locus
30 varied from 3 to 23. These markers are currently being used in studies on landscape genetics of
31 *Galaxias platei* in Patagonia.

32 **Main text**

33 The family Galaxiidae (Pisces) is confined to the Southern Hemisphere. *Galaxias platei*
34 (Steindachner 1898) is a freshwater species endemic to the lakes and rivers of the Patagonian
35 Andes between latitude 38º S and the southern tip of South America (Cussac *et al.* 2004).
36 *Galaxias platei* is primarily found within deep benthic environments with adaptations for the
37 cold and turbid waters characterizing deep Andean lakes fed by glacial melt water (Cussac *et al.*
38 2004), but can also be found in the littoral shore environments of some lakes and rivers that are
39 free of salmonids (Evelyn Habit, personal communication).
40 Recent molecular studies show that Quaternary glacial cycles had a strong influence on the
41 phylogeography of this species (Ruzzante *et al.* 2008; Zemlak *et al.*2008). For instance, a range-
42 wide survey of genetic diversity using control region sequence data found evidence of four
43 genealogical lineages that are geographically clustered with minimal overlap (Zemlak *et al.*
44 2008). It is likely that this signal reflects the isolation of this species into four separate glacial
45 refugia by Andean orogeny and repeated Pleistocene glacial cycles and recent secondary contact
46 among lineages (Zemlak *et al.* 2008). Other studies have examined the phenotypic variability as

a function of latitude (Milano *et al.* 2006), predation risk and diet (Milano *et al.* 2002, 2006). More recent studies based on multilocus phylogepographic approaches have shown that the species also experienced severe genetic bottlenecks in the mid Pleistocene when glaciers covered vast areas of Patagonia (Zemlak *et al.* 2011). These studies provide a valuable historical and ecological context for the interpretation of the regional genetic diversity and ancestral population dynamics of *G. platei* based on mitochondrial and nuclear DNA sequence variation. Here, we describe 9 novel microsatellite loci developed specifically for *G. platei* and intended for their use in landscape genetics and phylogeographic studies.

Genomic DNA was extracted using the standard phenol-chloroform, isoamyl-alcohol technique (Sambrook *et al.* 1989) from one individual collected from Lake Llanquihue in 2006. Microsatellite enriched libraries were created following the Glenn and Schable (2005) protocol. Briefly, approximately 2 μg of genomic DNA was digested using *RsaI* and subsequently ligated to super SNX linkers (Glenn & Schable 2005). Biotinylated oligonucleotide probe mixtures with the following motifs were hybridized to the template DNA: Mix1[(CATC)4, (GACA)4, (GATA)4, (AG)9 (AC)9] and Mix2[(AAAG)7, (GGAT)5, (GTAT)5, ((GATA)7, (GACA)7]. Enriched fragments were then captured using streptavidin coated magnetic beads (Dynal, Invitrogen), PCR amplified, ligated into pDrive cloning vector (Qiagen PCR Cloning Kit), transformed into New England Biolabs 5-alpha high efficiency competent *E. coli* and plated on Invitrogen imMedia™ Amp Blue media. Positive clones (n=200) were PCR amplified using M13 primers under standard PCR conditions. PCR products were screened in 1% agarose gels and 92 suitably sized inserts (>500 bp) were sent for sequencing to Macrogen USA. A total of 30 sequences were aligned and edited using Sequencher 4.5 and searched for microsatellite repeats using Microsatellite Repeat Finder (http://sgdp.iop.kcl.ac.uk/nikammar/repeatfinder.html).

Primer pairs were designed using Primer3 software (Rozen & Skaletsky 2000) for 19 candidate loci, 9 of which proved useful when tested with 126 samples from two lakes, Belgrano Lake (N=93) and Laguna del Mié (N=33) in Perito Moreno National Park, Argentina (Table I). PCR reactions were performed in 5µL volumes using a fluorescent dye labeled M13-Fwd-21 primer (IR700 or IR800 dye), a 5’-M13 tailed forward primer and unmodified reverse primer (following Oetting *et al.* 1995). Reaction conditions included 20 mM Tris-HCl, 10 mM (NH_4)_2 SO_4, 10 mM KCl, 1.5mM MgCl_2 , 0.1% Triton X-100, 0.2 mM each dNTP, 0.2 uM labelled M13 and reverse primers, 0.02 uM forward tailed primer and 0.25U Tsg DNA polymerase (BioBasic Inc, Markham, ONT, Canada). The cycling conditions used in Eppendorf thermocyclers were: 95 ºC for 5 min, followed by 30 cycles of 95 ºC for 45 s, primer-specific annealing temperature for 60 s (see Table I), 72 ºC for 60 s, and a final extension at 72 ºC for 10 min. PCRs products were imaged in 6% denaturing polyacrylamide gels (Sequagel 6, National Diagnostics, Atlanta, GA, USA) on LI-COR DNA Analyzers (LI-COR, Lincoln, NB, USA). PCR images included a custom built DNA standard ladder and scored by eye.

All genotypes were examined with Microchecker (van Oosterhout *et al.* 2004) and were found to be void of null alleles and large allele drop out, except for individuals from Laguna del Mié at the marker Gpla-13. Hardy Weinberg Exact Tests (Guo & Thompson 1992) were performed for each locus and population using GenePop (Raymond & Rousset 1995). No evidence of departure from HWE could be detected for any of the loci (all P>0.05), except for locus Gpla-3-10 from Lake Belgrano and for locus Gpla-13from Laguna del Mié. However, only Gpla-13 from Laguna del Miė remained significant after a Bonferroni correction (Rice 1989), possibly as a result of a small sample size (N=33). A Gametic Disequilibrium Test was implemented with GenePop (Raymond & Rousset 1995) and the hypothesis of independence of genotypes across loci could

93 not be rejected for any of the loci pair combinations (all P > 0.05), except for the pairs Gpla-2-5
94 and Gpla-13, and Gpla-3-5 and Gpla-3-8. The significance disappeared after Bonferroni
95 correction for multiple tests. Observed and expected heterozygosities were calculated using
96 Genalex (Peakall & Smouse 2006) and are shown in Table I. These markers are currently being
97 used to examine population genetics in Patagonian populations.

98 **References**

99 Cussac V, Ortubay S, Iglesias G, Milano D, Lattuca ME, Barriga JP, Battini MA, Gross M
100 (2004) The distribution of South American galaxiid fishes: the role of biological traits and post-
101 glacial history. *Journal of Biogeography*, **31**, 103–121.

102 Glenn TC, Schable NA (2005) Isolating microsatellite DNA loci. *Methods in Enzymology*, **395**,
103 202-222.

104 Guo SW, Thompson EA (1992) Performing the exact test for Hardy-Weinberg proportion for
105 multiple alleles. *Biometrics*, **48**, 361–372.

106 Milano D, Cussac VE, Macchi PJ, Alonso MF, Vigliano PH, Denegri MA (2002) Predator
107 associated morphology in *Galaxias platei* in Patagonian lakes. *Journal of Fish Biology*, **61**, 138-
108 156.

109 Milano D, Ruzzante DE, Cussac VE, Macchi P, Ferriz R, Barriga J, Aigo J, Lattuca ME, Walde
110 SJ (2006) Latitudinal and ecological correlates of morphological variation in *Galaxias platei*
111 (Pisces, Galaxiidae) in Patagonia. *Biological Journal of the Linnean Society*, **87**, 69–82.

112 Oetting WS, Lee HK, Flanders DJ, Weisner GL, Sellers TA, King RA (1995) Linkage analysis
113 with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed
114 primers. *Genomics*, **30**, 450-458.

115 Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic
116 software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295.
117 Raymond M, Rousset F (1995) GENEPOP 1.2: population genetics software for exact tests and
118 ecumenicism. *Journal of Heredity*, **86**, 248–249.
119 Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.
120 Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist
121 programmers. Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana
122 Press, Totowa, NJ. 365-386.
123 Ruzzante DE, Walde SJ, Gosse J, Cussac VE, Habit EM, Zemlak TS, Adams EDM (2008)
124 Climate control on ancestral population dynamics: insight from Patagonian fish phylogeography.
125 *Molecular Ecology*, **17**, 2234–2244.
126 Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual Cold
127 Spring. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
128 Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) micro-checker: software for
129 identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**,
130 535–538.
131 Zemlak TS, Habit EM, Walde SJ, Battini MA, Adams EDM, Ruzzante DE (2008) Across the
132 southern Andes on fin: Glacial refugia, drainage reversals and a secondary contact zone revealed
133 by the phylogeographic signal of *Galaxias platei* in Patagonia. *Molecular Ecology*,**17**, 5049–
134 5061.
135 Zemlak TS, Walde SJ, Habit E, Ruzzante DE (2011) Climate induced changes to the ancestral
136 population size of two Patagonian galaxiids: the influence of glacial cycling. *Molecular Ecology*,
137 **20,** 5280-5294.

138 .

139 **Acknowledgements**

140 We thank Abby van der Jagt for assistance in the laboratory. The work was funded by an
141 NSERC Discovery grant and a Special Research Opportunities award (SROPJ⁄326493-06) to
142 DER.

143 **Table I.** Characterization and primer sequences for 9 novel microsatellite markers for *G. platei*. Observed (Ho) and expected (He)

144 heterozygosities are showed for each population: Lake Belgrano (LB) and Laguna del Mié (LM).

Locus name	GenBank Accesion number	Primer sequence (5'-3')	Repeat Motif	NA/ S[§]	Allele range	TA	Ho (LB)	He (LB)	P-value (LB)	Ho (LM)	Ho (LM)	P-value (LM)
Gpla-2	JQ043430	F:GAGGAAGGCCATTAGGTGAATGA R: ACGGCCACAACATCACCATC	(ATGA)9atgcattaa (ATGG)5	7/126	248-272	61º	0.59	0.68	0.20	0.76	0.71	0.66
Gpla-13	JQ043431	F: TGCTCCACTTGGCAACCATA R: CACAGGGTATGGCAGCTTGA	(AGAC)37	21/125	183-311	61º	0.91	0.91	0.71	0.78	0.89	0.001*
Gpla-2-5	JQ043432	F: AGGTTGTTGAAGGGCGTAAG R: GGGAGAAGGACGTCTGGACT	(CATC)8	5/125	171-193	60º	0.39	0.38	0.86	0.62	0.51	0.48
Gpla-2-6	JQ043433	F: GGAACGACACTTTGGCATTA R: GATCATGTTGCTGTCCGATG	(GACA)17 (CAGG)13 (CAGA)4	23/126	226-326	60º	0.76	0.77	0.34	0.88	0.86	0.73
Gpla-2-11	JQ043434	F:CAGACATGGAGGGAGACAGC R:AGTCATCCGTCCGACAAAAC	(AGAC)3 (AC)3 (GGCA)7 (CAGA)21	20/126	218-326	60º	0.90	0.91	0.21	0.94	0.86	0.93
Gpla-3-5	JQ043435	F: TTTCCAGAGTTTTTCGGTCCT R: GGAAAACGGATGGTGAAGAA	(TCTT)12	17/126	220-294	60º	0.8	0.85	0.19	0.94	0.81	0.27
Gpla-3-8	JQ043436	F: ACCACCACGTCAGGAATAGC R:CCTGACAGCAAACCCTCAGT	(CAGG)7 (AGAA)23 (AT)3	16/126	196-256	62º	0.88	0.90	0.73	0.94	0.89	0.11
Gpla-3-10	JQ043437	F:CCCCCTGGTAACCTTCTATTT R: CCTCCTGGCTGATAGACTCG	(TTTG)4 (TTTG)3 (TTTG)3	3/126	208-232	60º	0.23	0.27	0.01*	0.18	0.16	1.0
Gpla-3-11	JQ043438	F: GAGAGGCTTCAATGGCTCAA R: TCGCTTGCTGTGAAAATGAC	(TACC)23 (ACCT)3	17/126	175-247	58º	0.84	0.89	0.12	0.82	0.85	0.5

145 [§] NA/S= Number of Alleles/Genotyped samples.

146 *Null hypothesis was rejected with the "exact HW test" (Guo and Thompson 1992).