

Characterization of 36 additional microsatellite loci in splittail (*Pogonichthys macrolepidotus*) and cross-amplification in five other native Californian cyprinid species

B. Mahardja · B. May · M. R. Baerwald

Received: 3 May 2012 / Accepted: 14 May 2012
© Springer Science+Business Media B.V. 2012

Abstract We characterized 36 new microsatellite markers for splittail (*Pogonichthys macrolepidotus*), an estuarine fish species endemic to the San Francisco Estuary. Genetic variation was assessed using 25 individuals from the Sacramento River and 16 individuals from the Petaluma River. Number of alleles per locus varies considerably, ranging from 3 to 19. One locus deviated significantly from Hardy–Weinberg equilibrium and a single pair of loci was found to be in linkage disequilibrium. Twenty-four markers successfully cross-amplified and were polymorphic in at least one of the five additional California cyprinid species (*Ptychocheilus grandis*, *Siphateles bicolor*, *Lavinia exilicauda*, *Orthodon microlepidotus* & *Mylopharodon conocephalus*) examined in this study.

Keywords *Pogonichthys* · Cyprinidae · Microsatellites · Effective population size

Due to habitat alteration, overexploitation, and introduction of exotic species, many of the native fishes in the San Francisco Estuary have experienced severe population decline (Service 2007; Sommer et al. 2007a). Included among this group is the splittail (*Pogonichthys macrolepidotus*), an endemic cyprinid species. Splittail represents the sole extant member of its genus and it is currently listed as a California Species of Special Concern (Moyle 2002; Sommer et al. 2007b). Although a number of studies on the life history and habitat use of splittail exists, a majority of these operated under the assumption of a single panmictic

population. Thus, the relatively recent discovery of two genetically distinct splittail populations (Central Valley population and Petaluma & Napa River population) suggests that our understanding of the species' ecology is incomplete (Baerwald et al. 2007; Feyrer et al. 2010). The thirteen microsatellite markers used in this previous study were sufficient in detecting population structure and designating an individual to its population of origin (Baerwald et al. 2008). However, as further information is required on the spatiotemporal occupancy, density, and spawning site preference differences between the two populations, supplementary microsatellite markers are essential in order to estimate effective population sizes and provide more accurate population assignment tests. Moreover, little is known about the genetic variation or population structure of other native California cyprinid species and the discovery of molecular markers for these species will facilitate such studies.

Two microsatellite libraries enriched for tetranucleotide repeats (CAGA)_n and (TAGA)_n previously described in Baerwald and May (2004) were used to sequence an additional 384 clones. Comparison of sequences to identify duplicates was conducted in SEQUENCHER version 4.8 (Gene Codes Corporation). MREPS version 2.5 (Kolpakov et al. 2003) was used to identify microsatellite repeats and flanking primer pairs were designed using Primer 3 (Rozen and Skaletsky 2000). M13 primer sequence was added to the 5' end of each forward primer and ran with a complementary 5' FAM fluorophore during polymerase chain reaction (PCR) reaction based on Schuelke's (2000) method. Three species for which the library was designed were used in the initial PCR screening, with a total of seven splittail individuals from the Central Valley population, four Sacramento pikeminnows (*Ptychocheilus grandis*), and four Mohave tui chubs (*Siphateles bicolor*).

B. Mahardja (✉) · B. May · M. R. Baerwald
Department of Animal Science, University of California, Davis,
One Shields Avenue, Davis, CA 95616, USA
e-mail: bmahardja@ucdavis.edu

Table 1 Characterization of 36 microsatellite loci in splittail (*Pogonichthys macrolepidotus*)

Locus and genbank accession number	Repeat motif	Forward primer (5'-3') Reverse primer (5'-3')	N (Central valley population) (Petaluma-Napa population)	H _E	H _O	HWE P value
<i>Pmac01</i> JQ958329	(ATAG) ₉ (AGAC) ₈ (AGAT) ₁₃	TTTACCCAACACTTGTCTCTGC TGGCACATTACCCTTGTGAA	26 16	0.74 0.83	0.65 0.81	0.195 0.647
<i>Pmac02</i> JQ958330	(AGAC) ₇	CACAGTGACTTGGGTGGATG CTGACACAGCAGCATTGGTT	26 16	0.25 0.09	0.19 0.09	1.000 1.000
<i>Pmac03</i> JQ958331	(ATAG) ₄ (AGAC) ₇ (AGAT) ₇	AAACATAGCATGCTGGAACAGA CACTTTGCCAACACGATGAC	26 15	0.83 0.91	0.78 0.67	0.088 0.016
<i>Pmac04</i> JQ958332	(TCTA) ₇ (TCTG) ₇	CCCCTGTGAGCATCATAACC CCAGCTCCTTCTCAAACCTG	26 16	0.80 0.77	0.83 0.88	0.085 0.990
<i>Pmac05</i> JQ958333	(AGAC) ₈	GAGAAAGAACCATATTCATCAGCA ACTGTGTGATTGTCCCCTCA	20 16	0.86 0.80	0.47 0.81	0.000* 0.056
<i>Pmac06</i> JQ958334	(AGAT) ₁₄	CATGTTGAGAGGGTTCGTGATT GGAAGAGGACTTGCCACAAA	26 16	0.20 0.41	0.17 0.50	0.090 1.000
<i>Pmac07</i> JQ958335	(TAGA) ₁₉	TCCAATATATGCATCCTTAATTTAAATC CATTTCTAAACACCCACTCACAG	24 15	0.22 0.19	0.24 0.13	1.000 0.103
<i>Pmac08</i> JQ958336	(TTTC) ₃ (TCTA) ₁₀ (TTCT) ₄	TGCCAGCAGTTGCATAGTA TGTGTAACAAATAAGCACTCGGTTA	25 15	0.67 0.66	0.65 0.67	0.640 0.456
<i>Pmac09</i> JQ958337	(TATC) ₉	CATTTATTCAATGGCGGCTA TGACATGGTTTATGGCATGG	26 15	0.63 0.71	0.70 0.87	0.919 0.619
<i>Pmac10</i> JQ958338	(CTAT) ₅ (TCTG) ₄ (TCTA) ₁₂	TTGTGCAATTGGTAAACACTTTG ACTGCCATGTGATTGGTTGA	26 16	0.72 0.60	0.74 0.75	0.657 0.319
<i>Pmac11</i> JQ958339	(TCTA) ₇ (TCTG) ₈	CAGCCGACTGACTAACCTCA TGCTTACCAAGGCTGCATT	26 16	0.62 0.57	0.61 0.44	0.526 0.112
<i>Pmac12</i> JQ958340	(GATA) ₉ (ATAG) ₅	TATCTCCATGCCAAAATCG ACACATTTCCCCAAAAGA	25 16	0.66 0.57	0.55 0.69	0.130 0.795
<i>Pmac13</i> JQ958341	(TAGA) ₈ (GATG) ₈	TGCCAATACCAGATCTGTCATAAT CTTGTGTGGTGGTTCATTCA	25 16	0.82 0.84	0.77 0.81	0.363 0.371
<i>Pmac14</i> JQ958342	(GACA) ₆ (AGAT) ₁₀	AGCCCTGGTCCCTTAGTTGT CGTGTATTGCAGCCATCAAA	25 16	0.94 0.88	0.95 0.88	0.886 0.250
<i>Pmac15</i> JQ958343	(ATCT) ₁₁	TGTTTGAACGGTCTTGCTTG CAAACCAGACGAGCCAGAAT	26 16	0.40 0.28	0.35 0.06	0.457 0.003
<i>Pmac16</i> JQ958344	(AGAT) ₈	TGCCTTGACAAAGTTTTGTGT CCTCACACACTGGCGAGTAA	26 15	0.85 0.85	0.61 0.80	0.025 0.287
<i>Pmac17</i> JQ958345	(ATAG) ₈ (AGAC) ₇	GGGAAATTTCAACCCAGAGG CGGACTCAGTGTGAATGACC	25 16	0.85 0.81	0.91 0.88	0.737 0.830
<i>Pmac18</i> JQ958346	(TATC) ₉ (CTGT) ₄	AAAACACGTACACTGGTTTGTCA TGAGTGGTCTCTGTCCCTGTT	25 16	0.83 0.79	0.74 0.81	0.244 0.232
<i>Pmac19</i> JQ958347	(TATC) ₇	ACTGTGGCGTGTGTTTGTGT CAGTTTTTCACCGTAGCAGTTTT	24 16	0.38 0.27	0.27 0.19	0.282 0.306
<i>Pmac20</i> JQ958348	(GATA) ₂₂ (AC) ₁₃	GAACACAGCTACTCTGGAAAACAA GGAAAAAGACATTGCTTACACCTT	25 16	0.33 0.41	0.36 0.50	1.000 1.000
<i>Pmac21</i> JQ958349	(GATA) ₂₁	TGCACTAAAGTGACATCTTGTGG CCCGACTATCAGAGCGAAGA	26 16	0.76 0.70	0.78 0.81	0.941 0.851
<i>Pmac22</i> JQ958350	(ATCT) ₈	TATCAAACCTGTGCCAGTGC TTTGGACAGGAATGTTGTTC	26 16	0.47 0.51	0.48 0.38	1.000 0.083
<i>Pmac23</i> JQ958351	(AC) ₉ (TCA) ₁₂	TCCTGTGTGAACCTGTCTGC GCCTGATGTCGTTCTTACCG	26 16	0.62 0.65	0.70 0.69	0.401 1.000

Table 1 continued

Locus and genbank accession number	Repeat motif	Forward primer (5'-3') Reverse primer (5'-3')	N (Central valley population) (Petaluma-Napa population)	H _E	H _O	HWE P value
<i>Pmac24</i> JQ958352	(TG) ₁₆ (TCTA) ₁₁	GGCGAAAACGTCTCAGTTTAGG TCCCAAAGCTCTGAATCTGC	25 16	0.31 0.56	0.30 0.44	0.143 0.544
<i>Pmac25</i> JQ958353	(TAGA) ₇	TGGGGTTTTGTGTCATTCAG CAAATCAGTCCCCATCAAGG	26 16	0.84 0.74	0.65 0.69	0.318 0.232
<i>Pmac26</i> JQ958354	(TAGA) ₁₈	GAGCAGTGGAAATGCAAAGGT AACCCATTCGTTTAAAACAGC	25 16	0.40 0.63	0.41 0.69	0.623 0.616
<i>Pmac27</i> JQ958355	(ATAG) ₁₂	TCCATGAGACTCTTCTCTCTTCA CATTGTCAAACCGGAAAAG	24 16	0.85 0.89	0.86 0.81	0.985 0.228
<i>Pmac28</i> JQ958356	(TATC) ₈	TCGTACTTAACCAGTTGCACAGA GGCGACTCGCTGGTATAGAG	26 14	0.72 0.73	0.57 0.71	0.062 0.423
<i>Pmac29</i> JQ958357	(TCTA) ₈	ACACGACAGTTTCCGCTGTA GTGTGCAGGTGCAATCATT	26 15	0.65 0.63	0.61 0.73	0.279 1.000
<i>Pmac30</i> JQ958358	(AT) ₇ (ATAG) ₁₀	GCAATGGAGTTGCCTCATT AGCGCTGATAGCTGGAGGT	26 16	0.73 0.79	0.61 0.81	0.348 0.885
<i>Pmac31</i> JQ958359	(TCTA) ₂₀	TTGCATTGCACTCTGAGAAAA TGGGGAAATGGTGTGTTGTGT	25 15	0.61 0.49	0.32 0.20	0.0087 0.010
<i>Pmac32</i> JQ958360	(ATAG) ₁₀	TGTGGCTTACATAAGGTGGTCA TGGGAGACGCTATCATCAGA	24 15	0.68 0.64	0.71 0.67	0.841 0.667
<i>Pmac33</i> JQ958361	(GATA) ₁₅	AAATGATGGGGAGAAGACGA GAGATGGGTGTGAGCTGCTT	24 16	0.60 0.48	0.57 0.38	0.933 0.052
<i>Pmac34</i> JQ958362	(AGAT) ₄ (AGAC) ₅ (AGAT) ₉	TTACAGCTTAAATTCAGTCATACTCG GAGGGTTAGGGATGTGGAAA	25 15	0.84 0.84	0.90 0.80	0.886 0.477
<i>Pmac35</i> JQ958363	(TCTA) ₇ (ATCC) ₃	CGTGTGGAAAAGGAGGAAAC GAATTGCCGTGAGATCAGGT	26 16	0.70 0.72	0.78 0.63	0.715 0.225
<i>Pmac36</i> JQ958364	(ATCT) ₈ (TTTC) ₄	GACCGTGAACCCAAATTTATTC CATTCTGTGGGTAGGAAAAGC	25 16	0.61 0.55	0.55 0.57	0.442 0.844

The two populations are examined separately for departures from HWE due to their status as distinct populations (Baerwald et al. 2007). * indicates significant deviation ($P < 0.05$) after sequential Bonferroni correction

PCR reaction and thermal cycler conditions followed microsatellite marker protocol listed in Baerwald et al. (2011). Amplified products were separated by size using a 3730xl Genetic Analyzer (Applied Biosystems) with GeneScan 600 LIZ (Applied Biosystems) as size standard. Further screening with larger sample size (including 16 additional splittails from the Central Valley population and 16 from the Petaluma-Napa River population) and cross-species amplifications for 3 other cyprinid species (*Lavinia exilicauda*, *Orthodon microlepidotus* & *Mylopharodon conocephalus*) (Table 2) were then conducted under the same conditions.

Fragments were visually scored using GeneMapper 4.0 (Applied Biosystems). GDA (Lewis and Zaykin 2001) was used to calculate observed (H_O) and expected (H_E) heterozygosities. Exact tests for conformation to Hardy-Weinberg equilibrium (HWE) and for the presence of linkage disequilibrium were performed with GenePop

Version 4.0.10 (Raymond and Rousset 1995; Rousset 2008). HWE tests were conducted independently for each splittail population to avoid Wahlund effect deviations (Baerwald et al. 2007). Due to low sample size ($n = 7-8$), other non-splittail species were excluded from these analyses.

A substantial amount of genetic diversity was observed in the 36 loci with allele number per locus from 3 to 19 with a mean of 7.1. Similarly, expected heterozygosity (H_E) ranged from 0.076 to 0.916 with an overall mean of 0.638 for the species (Table 1). Five loci deviated from HWE at $P = 0.05$, however, only one (*Pmac05*) remained significant after Bonferroni correction (Rice 1989). Linkage disequilibrium was found after Bonferroni correction for a single pair of loci: *Pmac04-Pmac16*. Cross-amplification was successful in 25 loci for at least one of the five non-splittail cyprinid species, though *Pmac23* locus

Table 2 Cross-species amplification of 36 markers for splittail and five other Californian species in the Cyprinidae family

Locus	Splittail (<i>Pogonichthys macrolepidotus</i>) N = 36–42	Sacramento pikeminnow (<i>Ptychocheilus grandis</i>) N = 8	Tui chub (<i>Siphateles bicolor</i> spp.) N = 8	Hitch (<i>Lavinia exilicauda</i>) N = 6	Sacramento blackfish (<i>Orthodon microlepidotus</i>) N = 8	Hardhead minnow (<i>Mylopharodon conocephalus</i>) N = 8
<i>Pmac01</i>	9 (162–198)	7 (208–260)	2 (273–281)	0	3 (289–297)	0
<i>Pmac02</i>	3 (238–246)	Mono (237)	0	0	2 (206–214)	2 (253–257)
<i>Pmac03</i>	18 (259–354)	2 (255–259)	2 (247–251)	3 (267–275)	3 (184–192)	4 (246–269)
<i>Pmac04</i>	7 (204–224)	9 (199–272)	6 (197–246)	7 (216–278)	12 (214–328)	6 (194–224)
<i>Pmac05</i>	11 (208–252)	0	0	0	0	0
<i>Pmac06</i>	4 (227–243)	7 (254–279)	0	0	0	6 (282–330)
<i>Pmac07</i>	3 (312–320)	10 (234–286)	4 (226–250)	9 (279–371)	0	4 (219–256)
<i>Pmac08</i>	5 (378–394)	5 (390–410)	0	0	0	0
<i>Pmac09</i>	5 (228–243)	0	0	0	0	0
<i>Pmac10</i>	6 (213–281)	0	0	0	0	0
<i>Pmac11</i>	6 (249–269)	3 (281–305)	3 (276–288)	0	Mono (160)	8 (292–338)
<i>Pmac12</i>	4 (273–297)	0	0	0	0	0
<i>Pmac13</i>	19 (279–371)	0	Mono (312)	5 (432–492)	Mono (310)	8 (366–414)
<i>Pmac14</i>	19 (188–288)	3 (179–191)	0	8 (194–295)	0	11 (220–337)
<i>Pmac15</i>	4 (246–262)	6 (244–268)	3 (267–275)	0	7 (277–369)	0
<i>Pmac16</i>	13 (212–260)	0	0	0	0	0
<i>Pmac17</i>	7 (244–272)	0	0	0	0	0
<i>Pmac18</i>	10 (166–212)	10 (159–199)	Mono (126)	8 (194–286)	0	5 (129–165)
<i>Pmac19</i>	4 (150–170)	0	0	9 (282–346)	0	5 (462–486)
<i>Pmac20</i>	4 (195–207)	9 (225–283)	0	0	9 (268–338)	0
<i>Pmac21</i>	5 (351–371)	4 (230–318)	6 (234–269)	8 (286–354)	11 (232–316)	7 (226–266)
<i>Pmac22</i>	5 (190–212)	0	0	0	0	0
<i>Pmac23</i>	5 (236–251)	0	0	0	0	Mono (162)
<i>Pmac24</i>	5 (209–233)	6 (227–265)	0	6 (231–267)	8 (192–246)	4 (234–256)
<i>Pmac25</i>	7 (231–255)	0	0	0	0	0
<i>Pmac26</i>	5 (205–221)	7 (218–254)	2 (177–189)	10 (210–318)	9 (230–290)	10 (254–318)
<i>Pmac27</i>	10 (245–285)	0	0	0	0	0
<i>Pmac28</i>	6 (212–240)	2 (221–233)	Mono (238)	0	0	0
<i>Pmac29</i>	4 (237–249)	6 (332–364)	Mono (306)	0	10 (342–440)	0
<i>Pmac30</i>	8 (170–198)	8 (266–317)	0	0	Mono (205)	9 (241–329)
<i>Pmac31</i>	4 (207–219)	7 (217–273)	Mono (198)	0	9 (245–294)	7 (303–336)
<i>Pmac32</i>	4 (218–230)	8 (256–292)	4 (252–284)	9 (244–304)	Mono (372)	5 (222–254)
<i>Pmac33</i>	4 (170–182)	7 (242–270)	Mono (267)	9 (222–328)	0	7 (488–532)
<i>Pmac34</i>	12 (197–245)	5 (195–228)	6 (204–232)	0	9 (228–276)	3 (232–252)
<i>Pmac35</i>	7 (246–274)	0	0	0	0	0
<i>Pmac36</i>	7 (254–286)	0	0	0	0	0

Listed above are the number of alleles observed in each locus and its size range. Zero value indicates non-specific or failed amplification

appears to be monomorphic for *M. conocephalus* (hardhead minnow) and did not amplify in other species (Table 2). These markers will be an important genetic tool in our goal to better understand the ecology and evolution of these six California cyprinid species.

Acknowledgments This material is based upon work supported by the Delta Science Program under Grant No. 2037. We would like to thank Katja Waldron for the assistance with data collection and

Andrea D. Schreier for providing helpful comments on earlier versions of the manuscript.

References

Baerwald MR, May B (2004) Characterization of microsatellite loci for five members of the minnow family Cyprinidae found in the Sacramento-San Joaquin Delta and its tributaries. *Mol Ecol Notes* 4:385–390. doi:10.1111/j.1471-8286.2004.00661.x

- Baerwald M, Bien V, Feyrer F, May B (2007) Genetic analysis reveals two distinct Sacramento splittail (*Pogonichthys macrolepidotus*) populations. *Conserv Genet* 8:159–167. doi:[10.1007/s10592-006-9157-2](https://doi.org/10.1007/s10592-006-9157-2)
- Baerwald M, Feyrer F, May B (2008) Distribution of genetically differentiated splittail populations during the nonspawning season. *Trans Am Fish Soc* 137:1335–1345. doi:[10.1577/T07-097.1](https://doi.org/10.1577/T07-097.1)
- Baerwald MR, Petersen JL, Hedrick RP, Schisler GJ, May B (2011) A major effect quantitative trait locus for whirling disease resistance identified in rainbow trout (*Oncorhynchus mykiss*). *Heredity* 106:920–926. doi:[10.1038/hdy.2010.137](https://doi.org/10.1038/hdy.2010.137)
- Feyrer F, Hobbs J, Sommer T (2010) Salinity inhabited by age-0 splittail (*Pogonichthys macrolepidotus*) as determined by direct field observation and retrospective analyses with otolith chemistry. *San Franc Estuary W Sci* 8:1–10
- Kolpakov R, Bana G, Kucherov G (2003) Mreps: efficient and flexible detection of tandem repeats in DNA. *Nucleic Acid Res* 31:3672–3678
- Lewis PO, Zaykin D (2001) Genetic data analysis: computer program for the analysis of allelic data. http://lewis.eeb.uconn.edu/lewis_home/software.html
- Moyle PB (2002) *Inland fishes of California*. University of California Press, Berkeley
- Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered* 86:248–249
- Rice WR (1989) Analyzing tables of statistical tests. *Evolution* 43:223–225
- Rousset F (2008) Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour* 8:103–106
- 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, New Jersey, pp 365–386. <http://fokker.wi.mit.edu/primer3/input.htm>
- Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nat Biotechnol* 18:233–234
- Service RT (2007) Delta blues, California style. *Science* 317:442–445
- Sommer T, Armor C, Baxter R et al (2007a) The collapse of pelagic fishes in the upper San Francisco estuary: El colapso de los peces pelagicos en la cabecera del estuario San Francisco. *Fisheries* 32:270–277
- Sommer TR, Baxter RD, Feyrer F (2007b) Splittail “delisting”: a review of recent population trends and restoration activities. *Am Fish Soc* 53:25–38