Conservation Genet Resour (2010) 2:227–230
DOI 10.1007/s12686-010-9244-z

TECHNICAL NOTE

Characterization and evaluation of polymorphic microsatellite markers in the anadromous fish *Spirinchus thaleichthys*

Joshua A. Israel · Bernie May

Received: 23 April 2010 / Accepted: 5 May 2010 / Published online: 28 May 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract We describe seventeen microsatellite loci isolated from longfin smelt, *Spirinchus thaleichthys*, which is an anadromous fish inhabiting estuarine and nearshore marine waters along the western Pacific coast. Five to 16 alleles per locus were detected, the expected heterozygosity ranged from 0.33–0.93, and moderate levels of differentiation were found between collections. These polymorphic microsatellites will provide useful tools for studying population genetic structure at ecological scales within wild and propagated populations.

Keywords Longfin smelt · Conservation · Propagation · Pelagic Organism decline

While longfin smelt, *Spirinchus thaleichthys*, remain among one of the more common pelagic native species in the San Francisco Bay-Delta, CA, a recent significant population decline at the southern edge of its distribution has been observed (Sommer et al. 2007). This species was recently listed as Threatened under the California Endangered Species Act (CDFG 2009) although a petition for listing under the U.S. Endangered Species Act remains undetermined. Longfin smelt show a diversity of migratory patterns and anadromous and resident populations are native to the Pacific coast of North America. These fish are semelparous, typically undergo a two-year life cycle and spawn in freshwater and the low salinity zone before larval and juvenile fishes are transported downstream to nursery habitats (Hobbs et al. accepted). Juveniles move out of the San Francisco Bay into nearshore coastal waters, while others seem to reside in the Bay-Delta (Moyle 2002). Little is known about San Francisco longfin smelt spawning habitats and ecology, although adults presumably migrate back to the low salinity zone immediately prior to spawning (Rosenfield and Baxter 2007). A better-studied resident population inhabits Lake Washington, WA (Chigbu 2000; Chigbu and Sibley 1994; Chigbu et al. 1998). While longfin smelt are reported to be widespread in nearshore coastal regions and estuaries, few fish have been observed and fewer spawning populations have been identified between San Francisco Bay and Lake Washington (USFWS 1994, CDFG 2009).

Whole genomic DNA was extracted from fin tissue of longfin smelt from Suisun March, CA using the PureGene Genomic DNA Purification Kit. Six libraries enriched for one of four tetranucleotide repeat motifs [AAAC, CAGA, CATC, TAGA] were constructed, screened, and sequenced by Genetic Identification Services according to Börk et al. (2008). Sequences were analyzed with SEQUENCHER version 4.7 (Gene Code Corporation) to remove duplicate sequences. Repeat regions in sequences were identified with MREPS version 2.5 (Kolpakov et al. 2003) and primer pairs flanking these regions were created with PRIMER 3 (Rozen and Skaletsky 2000) for 200 of 227 novel sequences. For screening, one of four 5′ modifications (universal primers: T7T, T7P, M13, SP6) was added to the end of each forward primer using the method of Schuelke (2000), which incorporated a 5′ fluorophore during polymerase chain reaction (PCR). Each primer set was initially tested on eight longfin smelt samples to determine

J. A. Israel · B. May
Department of Animal Science, Genomic Variation Laboratory, 1 Shields Ave, Davis, CA 95616, USA

J. A. Israel (✉)
Applied Science Branch, Bureau of Reclamation, U.S. Department of Interior, 2800 Cottage Way, MP-151, Sacramento, CA 95825, USA
e-mail: jaisrael@ucdavis.edu; JAIsrael@usbr.gov

Table 1 Characterization of microsatellite loci: locus, motif, accession number, primer sequence, tail sequence, annealing temperature (T_a), sample size (*n*), number of alleles (N_A), size range of observed allelic variation, H_O, and H_E. Significant deviations from Hardy–Weinberg expectations are in bold

Locus	Sequence	Accession no.	Primer sequence	Primer tail sequence name	Ta (°C)	N_A	Size range (bp)	Lake Washington			Bay-Delta		
								H_O	H_E	*n*	H_O	H_E	*n*
SptB12	$(CTGT)_{15}$	GU946459	F: TGTCTTTGGGTCCATTTC	M13	54	12	157–272	0.75	0.69	12	0.50	0.79	12
			R: GAACAAGTCTGCCACCAG										
SptC3	$(GTAT)_2(GGAT)_4GTAT(GGAT)_3$	GU946466	F: GCCGTTGCTTGAGACTAA	T7P	55	10	287–373	0.60	0.56	15	**0.54**	0.85	13
			R: TGTCTGGGGAATTCAGTG										
SptC5	$(CATC)_2(ACTT(CATC)_2CACC(CATC)_5$	GU946467	F: TGACTCCCTCCTTTCCTC	T7P	55	5	238–258	0.47	0.38	15	**0.64**	0.66	14
			R: GGGTAGAGGGAACCATGT										
SptC6	$(CA)_4CC(CA)_7AA(CA)_2CC(CA)_4$	GU946468	F: GTTCGGCTGCTCCTCTAT	T7P	55	6	122–142	0.66	0.54	15	0.82	0.75	11
			R: TGGGTGTGTTTGTGTGTG										
SptC10	$(GGAT)_6(GAAT)_2GGACGAAT(GGAT)_3(GAAT(GGAT)_2)_2$	GU946463	F: ATTTACGGACGGATGGAT	T7P	55	9	202–232	0.53	0.63	15	0.79	0.87	14
			R: GCCATTTTCTTGCTTTCA										
SptB130	$(CC(TGTC)_2(TCGT)_2)_7$	GU946462	F: TGGACTGGTGTGACCATCTG	T7T	60	10	490–542	0.77	0.67	13	0.70	0.86	10
			R:GCCAGACAGATATAGACGAACAAA										
SptC114	$(CA)_4CC(CA)_{10}CC(CA)_4$	GU946465	F: TCGGCTGCTCCTCTATGACT	M13	60	6	95–130	0.53	0.39	15	0.71	0.68	14
			R: GGGTGTGTGTGTGTGTGTGT										
SptB104	$((GACA)_2A)_{17}$	GU946461	F: CAGCCTGCCAACTGATGATA	T7P	59	11	201–291	0.53	0.79	15	0.90	0.86	10
			R: GCTTTGCCTCCTTCATTGAC										
SptB108	$(ATGG)_{13}$	GU946464	F: TCCTTGTGAAAACGCTAATCC	M13	60	8	121–148	0.50	0.61	14	**0.43**	0.78	14
			R: GAATGTCACCCACCCCAGTA										
SptD9	$(GT)_{13}$	GU946474	R: CCCCACACACCCACTAC	M10	62	11	190–201	0.57	0.62	14	0.93	0.82	14
			R: CCCCACACACCCACTAC										
SptD3	$(ATCT)_2(CTATCTGT)_3CTAT(CTATCTGT)_2CTAT(CTATCTGT)_4(CTATCCGT)_2(CTAT)_3$	GU946473	F: TCGTTTTCCACAAAGGTC	M10	59	9	200–316	0.75	0.69	12	0.50	0.80	12
			R: CGGGAGAGGAATACAACA										
SptD269	$(GTCT)_9$	GU946469	F: TGCTCCTCTAAGTACAGGAAAGG	M13	62	10	124–174	**0.60**	0.76	15	0.50	0.61	15
			R: AGGCAGGCAGAGAGACAGAT										
SptD206	$((TGTC)_3TGCA)_4(TGTC)_6TGCA(TGTC)_6$	GU946460	F: TGCCTCGTCGTCTACCTACC	T7T	59	16	204–372	0.93	0.57	14	**0.57**	0.93	14
			R: AGCCAAGTGTCCAGACAAGG										
SptD220	$(CTGT)_3CTACAACCCTA(CTGT)_7$	GU946475	F: TACAACATCGCCAGACTGCT	M13	62	13	297–432	**0.31**	0.76	13	0.93	0.90	14
			R: TGTGACAGACAAGACGGAGAG										
SptD292	$(ACAG)_9$	GU946472	F: TGGGATGGGATTTTAGGATG	SP6	60	5	146–166	0.27	0.33	15	0.85	0.76	13

Table 1 continued

Locus	Sequence	Accession no.	Primer sequence	Primer tail sequence name	Ta (°C)	N_A	Size range (bp)	Lake Washington			Bay-Delta		
								H_O	H_E	n	H_O	H_E	n
			R: CTCCTCTGCAACCCTGTGAT										
SptD277	$(GGGAGAGACAGA(C/G)A)_{13}$	GU946470	F: TTCGATTGGGAGAGGAGAGA	SP6	59	16	218-397	0.40	0.67	15	**0.71**	0.93	14
			R: TGAGACATGATCCCCTCCTT										
SptD289	$(TC)_{24}$	GU946471	F: GACGCCCTATTTCTGGTCTTC	SP6	60	14	182–296	0.60	0.76	15	0.64	0.91	14
			R: TGGACGGACAAACTAGACAGG										

microsatellite amplification and polymorphism. Based on allelic diversity and clean scoring, nineteen primer sets were selected for amplification in a total of 12–15 individuals per collection from the San Francisco Delta, CA and Lake Washington, WA.

Polymerase chain reaction was performed with the following conditions: 20 ng DNA template, 1× GoTaq® Flexi Buffer, 1.875 mM MgCl2, 0.2 mM of each dTNP, 0.1 μM of fluorescently labeled primer and reverse primer, 30 nM tailed forward primer, and 0.4U of Taq DNA polymerase (all reagents from Promega), for a total reaction volume of 10 ul. PCR was performed using a Bio-Rad DNA engine Dyad thermal cycler under the following conditions: 95°C for 5 min; 5 cycles at 95°C for 45 s, 68°C for 5 min with 2°C decrease at each subsequent cycle, followed by 72°C for 1 min; 5 cycles at 95°C for 45 s, 58°C for 5 min with 2°C decrease at each subsequent cycle, followed by 72°C for 1 min; 30 cycles at 94°C for 45 s,°C for 2 min, followed by 72°C for 1 min; final extension at 72°C for 10 min; and held at 8°C. One microliter of fluorescently-labeled amplified product was run undiluted on an ABI 3130xl Genetic Analyzer with a LIZ600 size standard (ABI). GENEMAPPER version 4.0 (ABI) was used to analyze the electropherograms. All analyses were computed with the GENALEX software (Peakall and Smouse 2006) were estimated. This program calculated number of alleles, observed (H_o) and expected (H_E) heterozygosities, deviations from Hardy–Weinberg equilibrium (HWE), a pairwise F_{st} value, and genetic distance was calculated for Principal Component Analysis (PCA) to assess differentiation between collections.

Seventeen of the nineteen primers optimized in longfin smelt were consistently scoreable (Table 1). The number of alleles per locus ranged from 5 to 16 with an average of 13.8 alleles per locus, and the observed and expected heterosygosities ranged between 0.267–0.929 and

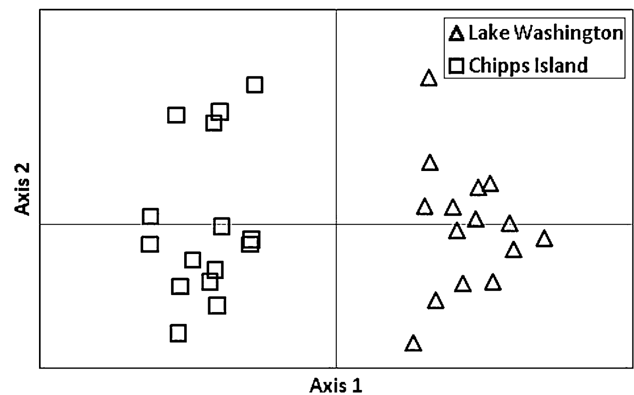

Fig. 1 Principal Component Analysis of two collections of longfin smelt ($n = 15$ each). Axis 1 accounts for 33.1% of the genetic covariance, while Axis 2 explains an additional 16.2% of the total variation

0.331–0.926, respectively. Seven loci deviated significantly from Hardy–Weinberg expectations. A pairwise F_{st} value of 0.091 was calculated between the two collections, and visualization of genetic distances using Principal Component Analysis supports moderate levels of differentiation between these two collection sites (Fig. 1). Genetic analysis of these collections suggested genetic isolation between locations, which are separated by over 800 miles of coast. These results are concordant with a previous allozyme study (Stanley et al. 1995), and support the distinctiveness and genetic isolation of the longfin smelt populations in these locations. These genetic markers will be useful for further evaluation of spatial and temporal genetic variation in *S. thaleichthys* and characterizing relationships among individuals in populations.

Acknowledgments This study was supported by the Interagency Ecological Program (DWR Award# 4600008137). The authors are indebted to Denice Lin, John Pedroia, and Emily Ringelman for laboratory assistance. Tom Quinn, Kathy Hieb, Jim Hobbs, and Randy Baxter are greatly appreciated for providing samples.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

Börk K, Drauch A, Israel JA, Pedroia J, Rodzen J, May B (2008) Development of new microsatellite primers for green and white sturgeon. Conserv Genet 9:973–979

CDFG (2009) A status review of the longfin smelt (Spirinchus thaleichthys) in California. Report to the fish and game commission. January 23, 2009, pp 45

Chigbu P (2000) Population biology of longfin smelt and aspects of the ecology of other major planktivorous fishes in Lake Washington. J Freshwater Ecol 15:543–557

Chigbu P, Sibley TH (1994) Relationship between abundance, growth, egg size, and fecundity in a landlocked population of longfin smelt, *Spirinchus thaleichthys*. J Fish Biol 1:1–15

Chigbu P, Sibley TH, Beauchamp DA (1998) Abundance and distribution of Neomysis mercedis and a major predator, longfin smelt (Spirinchus thaleichthys) in Lake Washington. Acta Hydrobiol 386:167–182

Hobbs JA, Lewis LS, Ikemiyagi N, Sommer T, Baxter R (accepted) The use of otolith strontium isotopes (87Sr/86Sr) to identify nursery habitat for a threatened estuarine fish. Environ Biol Fishes

Kolpakov R, Bana G, Kucherov G (2003) mreps: efficient and flexible detection of tandem repeats in DNA. Nucleic Acids Re 31:2678–3672

Moyle PB (2002) Inland fishes of California, 2nd edn. University of California Press, Berkeley, CA

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Rosenfield JA, Baxter R (2007) Population dynamics and distribution patterns of Longfin smelt in the San Francisco estuary. Trans Am Fish Soc 136:1577–1592

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Methods in molecular biology. Humana Press, Totowa, p 365

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234

Sommer T, Armor C, Baxter R et al (2007) The collapse of pelagic fishes in the upper San Francisco Estuary. Fisheries 32:270–277

Stanley SE, Moyle PB, Shafer HB (1995) Allozyme analysis of delta smelt, Hypomesus transpacificus and Longfin smelt, Spirinchus thaleichthys in the Sacramento-San Joaquin Estuary, California. Copeia 1995(2):390–396

USFWS (1994) Endangered and Threatened Wildlife and Plants; Notice of 1-Year Finding on a Petition to List the Longfin Smelt. U.S. Fish and Wildlife Service