Molecular Ecology Notes (2004) **4**, 739–741 doi: 10.1111/j.1471-8286.2004.00801.x

PRIMER NOTE

Characterization of polymorphic microsatellites in the Pacific sardine *Sardinops sagax sagax* (Clupeidae)

R. T. PEREYRA,* E. SAILLANT,† C. L. PRUETT,† C. E. REXROAD III,‡ A. ROCHA-OLIVARES* and J. R. GOLD†

**Departamento de Oceanografía Biológica, Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Carr. Tijuana-Ensenada km 107, Ensenada, Baja California C.P. 22860, México,* †*Center for Biosystematics and Biodiversity, Texas A & M University, College Station, Texas 77843–2258, USA,* ‡*USDA/ARS National Center for Cold and Cool Water Aquaculture, 11876 Leetown Road, Kearneysville, West Virginia 25430, USA*

Abstract

We isolated 11 microsatellites from the Pacific sardine *Sardinops sagax sagax*. The number of alleles and H_E among 30 individuals from a single population ranged from four to 24 and from 0.606 to 0.959, respectively. Pacific sardines are a vital economic resource in the northeastern Pacific Ocean, but insufficiently polymorphic loci have limited inferences about its stock structure and genetic variation. The level of variability of these new markers will prove useful in testing hypotheses on the stock-structure and long-term genetic integrity of the species.

Keywords: genomic library, marine fish, microsatellites, Pacific sardine, PCR primers

Received 14 June 2004; revision accepted 9 August 2004

Pacific sardines (*Sardinops sagax sagax*) have sustained economically important fisheries in Mexico, USA, and Canada since the 1930s (McFarlane *et al*. 2002). Coast-wide population collapses have raised concerns about their long-term management and conservation (Nevárez-Martínez *et al*. 2001; McFarlane *et al*. 2002). Critical issues for management of the Pacific sardine fisheries are stock delineation and identification of latitudinal origins of mixed catches along the northeast Pacific. In light of taxonomic debate on *Sardinops* species recognition (Grant *et al*. 1998), the use of morphological variation to identify intraspecific biological groups has proven difficult. Life-history features such as age at first reproduction, maximum size, age-specific mortality and meristics and morphometrics have been used to define four distinct stocks (Radovich 1982; Parrish *et al*. 1989), however, the significance of these characters for stock delineation remains a source of debate due to their often continuous variation (Hedgecock *et al*. 1989; McFarlane *et al*. 2002). Another management issue is the long-term genetic integrity of the resource; recent studies (Turner *et al*. 2002) have shown that some exploited marine fishes have effective population sizes orders-of-magnitude smaller than census sizes. Small effective sizes may constrain the ability of a population to respond to changing or novel environmental pressures (Higgins & Lynch 2001). A previous genetic study based on allozyme electrophoresis reported very low levels of polymorphisms ($0.005 \leq$ mean $H_E \leq 0.017$ in only nine bialelic loci) limiting the power of inference about its stock structure (Hedgecock *et al*. 1989). No other nuclear marker has been used to assess the population structure of Pacific sardines. Here, we report novel PCR primers for 11 nuclear-encoded microsatellites isolated from a genomic library of Pacific sardine. These markers are being used in studies of both stock structure and effective size of this species.

Microsatellite isolation followed Broughton & Gold (1997). Genomic DNA was digested with *Dpn*II (New England Biolabs, NEB). ~500–2000 base pairs (bp) fragments excised from a 1.5% agarose gel were ligated (T4 DNA ligase, NEB) into the *LacZ* gene multicloning site of a P-Bluescript vector, heat-shock transformed into XL10-Gold Ultra-competent cells (Stratagene, CA), and plated onto LB/Ampicillin/IPTG/X-Gal plates. 19 500 colonies with inserts (white) were picked using a Q-Bot (Genetix Ltd, UK) and transferred to LB-filled microtiter plates for screening: cells from individual wells were used to inoculate a Hybond nylon membrane (Amersham, Newark, NJ) on top of a slab of LB agar plus

Correspondence: A. Rocha-Olivares. Fax: + 52(646)175 0545; E-mail: arocha@cicese.mx

ampicillin. Each colony was spotted twice to eliminate false positives. Following overnight incubation (37 °C), colonies were lysed and DNA hybridized to the membranes, which were probed with two cocktails of the following γ^{32}P-labelled oligonucleotides [$(CA)_{13}$ $(GA)_{13}$ $(GAA)_8$ $(TAA)_{13}$ and $(CAA)_8$ $(CATA)_8$ $(GATA)_9$ $(GACA)_8$]. We screened 192 of these 'positive' clones as follows. Frozen glycerol stocks arrayed in 96-well plates were used to inoculate 1 mL cultures of Luria Broth selective media (ampicillin) and incubated overnight at 37 °C. Plasmid DNA was isolated (alkaline lysis) with a BioRobot 8000 (QIAGEN, CA). Miniprep DNA was then quantified, normalized, and both strands sequenced, using M13 forward and reverse sequencing primers and ABI BigDye TERMINATOR v3.1. Products were purified and electrophoresed on an ABI 3100 DNA Analyser (Applied Biosystems, CA). Phred (CodonCode, MA) was used for DNA sequence base calling and vector trimming.

We found 40 microsatellite arrays of suitable size (5–45 repeats) flanked by sequences of sufficient length for PCR primer design, which were designed for 30 arrays using NETPRIMER® (PREMIER Biosoft International, CA). Amplification was subsequently tested under different experimental conditions with an Omn-E (Hybaid, UK) thermal cycler and [γ^{32}P]-dATP end-labelling (T4 polynucleotide kinase). Reactions (10 μL) consisted of 1 μL (*c.* 5 ng) template DNA, 0.5 μM each primer, 200 μM each dNTP, 1 mM $MgCl_2$, 500 mM KCl, 100 mM Tris pH 9.0, 10% Triton-X 100 and 0.1 U *Taq* DNA polymerase (thermal cycling: 3 min 94 °C; 30 cycles of 94 °C 45 s, 50–65 °C 30 s, and 1 min 72 °C; final extension 10 min 72 °C). Products were electrophoresed in 6% polyacrylamide gels and visualized by autoradiography.

A subset of 13 PCR primer pairs were tested on eight fish for tractability (reproducibility, consistency, range in allele size, presence/frequency of 'stutter' bands, and polymorphisms). Subsequently, 11 of these were labelled to screen 22 additional fish from a single population. Products were pooled (see Table 1) and run on an ABI 377 automated sequencer, using 400 HD size standards. GENESCAN 3.1 and GENOTYPER 2.1 (Applied Biosystems, CA) were used for allele sizing and calling. Estimates of observed and expected heterozygosity (Nei 1987) were obtained using POPGENE (S. Piry and D. Bouget, http://www.ensam.inra.fr/URLB/). Exact tests of Hardy–Weinberg and linkage disequilibria were made following Weir & Cockerham (1984), as implemented in GENEPOP 3.4 (Raymond & Rousset 1995). Monte

Table 1 Characterization of microsatellites developed from the Pacific sardine (*Sardinops sagax sagax*). Repeat sequence indicates repeat motif of cloned allele; N is number of individuals assayed; N_A is number of alleles detected; H_E and H_O are expected and observed heterozygosity, respectively; P_{HW} is the significance of Hardy–Weinberg disequilibrium

Microsat. (pooling)*	Primer sequence (5′–3′)	Repeat sequence	Annealing (°C)	N/N_A	Size range (bp)	H_E/H_O	P_{HW}	GenBank Accesion of clone sequences
SarB-A07 (group 1)	CTCCTCACTCAGCCGCTAAGGA GGGTAACATTTCGGCAAGTGCT	$(GA)_{12}$	65	30/18	68–136	0.920/0.700	< 0.001	AY636114
SarB-A08 (group 2)	GTGATACTCTCTGCCTTGGA GCACTTTGTCCTAGTAAATAGC	$(CA)_{26}$	54	30/21	201–301	0.953/0.967	0.653	AY636115
Sar1-A11 (group 1)	GAGCTGGAAATCTGGTGATATTTAG CCTGTTCACAAGTTAGAGCATTC	$(GATA)_2GCTA(GATA)_5GCTA(GATA)_8$	63	30/19	257–313	0.934/0.833	0.189	AY636120
SarB-CO5 (group 2)	GAACGCAGACATAAAAGGGTC GGGTATGTGGTGATTATCGTTC	$(TC)_5TT(TC)_4$	56	30/22	120–182	0.953/0.833	0.168	AY636116
Sar1-D01 (group 1)	GCTCTGGTCGGAGGCTCTATC GGTGTTCACGTGGGCTGGTA	$(CA)_{29}GG(CA)_3$	61	30/24	182–256	0.963/0.900	0.417	AY636121
Sar1-D06(B) (group 1)	CGGCTATTCTTAGACTAGGTG CCCCATCAGCAATGAATAAG	$(TG)_{18}$	50	30/14	120–158	0.794/0.700	0.137	AY636123
SarB-D09 (group 3)	GGTCATCTGCTTCAACAACAC GCAGCCTGTCTGAAACTCTG	$(CA)_9(GA)_8$	59	30/15	269–303	0.905/0.967	0.633	AY636117
SarB-G09 (group 3)	GGTGGAAAGAACACTGGTCA GGTTCACTATGCAGGCTATG	$(GA)_6GT(GA)_{36}GT(GA)_3$	61	30/4	149–163	0.606/0.667	0.421	AY636118
SarB-H04 (group 2)	CGAGTTTGTCCCACACCTGGAG CTCCAAGCACCGAGAGCATC	$(GT)_9$	52	30/25	181–271	0.959/0.933	0.745	AY636119
SarB-H04(F) (group 3)	CTCTCGGTGCTTGGAGAGGAA GGAGGAGGGGAGAAAAGATG	$(TG)_{18}$	59	30/28	116–168	0.947/0.733	< 0.001	AY636119
Sar1-H11(B) (group 1)	CACGGCACGTTACGTTTCAG CCAGCGTGTCATGAAATGATG	$(TG)_{11}TA(TG)_6$	65	30/21	150–200	0.937/0.633	< 0.001	AY636122

*Groups (1, 2 and 3) refer to those in which PCR products were pooled for simultaneous electrophoretic runs.

Carlo simulation significances were adjusted by sequential Bonferroni correction (Rice 1989).

Five microsatellites showed 'perfect' dinucleotide repeats, five 'imperfect' motifs (mixtures of di- or tetranucleotide repeats), and one showed a compound repeat (Table 1). All loci were polymorphic in the 30 individuals assayed and the number of alleles per locus varied from four to 24, with allele sizes from 68 to 136 (*SarB-A07*) to 257–313 (*Sar1-A11*). Observed and expected heterozygosity ranged between 0.667 and 0.967 and 0.606–0.959, respectively. Genotype proportions at three loci [*SarB-A07*, *SarB-H04*(*F*) and *Sar1-H11*(*B*)] deviated significantly from Hardy–Weinberg equilibrium after sequential Bonferroni correction ($P < 0.05$). Estimated F_{IS} values of three microsatellites were positive (0.243 *SarB-A07*, 0.229 *SarB-H04*(*F*), and 0.328 *Sar1-H11*(*B*)), possibly suggesting null alleles. No significant pairwise linkage disequilibrium was found among loci after sequential Bonferroni correction (final $P = 0.004$), even though all loci pairs except those involving *Sar1-D01*, *SarB-D06(B)*, and *SarB-G09* were individually significant ($P < 0.05$).

Acknowledgements

We thank C. Bradfield for technical assistance and V. Taylor, L. Laughlin, J. McCrae and M. Robinson for providing samples. Funding was provided by a collaborative research grant from Mexican CONACYT and TAMU to ARO and JRG and by the Texas Agricultural Experiment Station under Project H-6703 to JRG.

References

Broughton RE, Gold JR (1997) Microsatellite development and survey of variation in northern bluefin tuna (*Thunnus thynnus*). *Molecular Marine Biology and Biotechnology*, **6**, 308–314.

Grant WS, Clark AM, Bowen BW (1998) Why restriction fragment length polymorphism analysis of mitochondrial DNA failed to resolve sardine (*Sardinops*) biogeography: insights from mitochondrial DNA cytochrome b sequences. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 2539–2547.

Hedgecock D, Hutchinson ES, Li G, Sly FL, Nelson K (1989) Genetic and morphometric variation in the Pacific Sardine, *Sardinops sagax caerulea*: comparisons and contrasts with historical data and with variability in the Northern Anchovy, *Engraulis mordax*. *U.S. Fishery Bulletin*, **87**, 653–671.

Higgins K, Lynch M (2001) Metapopulation extinction caused by mutation accumulation. *Proceedings of the National Academy of Science USA*, **98**, 2928–2933.

McFarlane GA, Smith PE, Baumgartner TR, Hunter JR (2002) Climate variability and Pacific sardine populations and fisheries. In: *Fisheries in a Changing Climate* (ed. McGinn NA), pp. 195–214. American Fisheries Society Symposium 32, Maryland, USA.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nevárez-Martínez MO, Lluch-Belda D, Cisneros-Mata MA *et al.* (2001) Distribution and abundance of the Pacific sardine (*Sardinops sagax*) in the Gulf of California and their relation with the environment. *Progress in Oceanography*, **49**, 565–580.

Parrish RH, Serra R, Grant WS (1989) the monotypic sardine, *Sardina* and *Sardinops*: their taxonomy, distribution, stock structure, and zoogeography. *Canadian Journal of Fisheries and Aquatic Sciences*, **46**, 2019–2036.

Radovich J (1982) The collapse of Californian sardine fishery. What have we learned? *California Cooperative Oceanic Fisheries Investigations Reports*, **23**, 56–79.

Raymond M, Rousset F (1995) GENEPOP Version 1.2.: Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Turner TF, Wares JP, Gold JR (2002) Genetic effective size is three orders of magnitude smaller than census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). *Genetics*, **162**, 1329–1339.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.