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 ($0.005 \leq \text{mean } H_E \leq 0.017$ 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 DpnII (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

© 2004 Blackwell Publishing Ltd
were probed with two cocktails of the following \( \gamma ^{32} P \)-labelled oligonucleotides [(CA)\(_5\) (GA)\(_6\) (TAA)\(_3\)] and (CAA)\(_6\) (CATA)\(_5\) (GATA)\(_5\) (GACA)\(_5\)]. 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 \( [\gamma ^{32} P] \)-dATP end-labeling (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 GENEPOL 3.4 (Raymond & Rousset 1995). Monte Carlo simulation was used to estimate the significance level for tests of linkage disequilibrium (25,000 permutations).

### Table 1: Characterization of microsatellites developed from the Pacific sardine (Sardinops sagax sagax).

<table>
<thead>
<tr>
<th>Microsat. (pooling)*</th>
<th>Primer sequence (5’–3’)</th>
<th>Repeat sequence</th>
<th>Annealing (°C)</th>
<th>Size range (bp)</th>
<th>( H_E )/( H_O )</th>
<th>( P_{HT} )</th>
<th>GenBank Accession of clone sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>SarB-A07 (group 1)</td>
<td>CCTCTACACTCACGGGCTCTAAGGA</td>
<td>(GA)(_{12})</td>
<td>65</td>
<td>30/18</td>
<td>68–136</td>
<td>0.920/0.700</td>
<td>&lt; 0.001 AY636114</td>
</tr>
<tr>
<td>SarB-A08 (group 2)</td>
<td>GGTTACATTCTGCTGGGTGAGGG</td>
<td>(CA)(_{26})</td>
<td>54</td>
<td>30/21</td>
<td>201–301</td>
<td>0.953/0.967</td>
<td>0.653 AY636115</td>
</tr>
<tr>
<td>Sar1-A11 (group 1)</td>
<td>GAACTGGCAATCTGCTGATTTAGTGA</td>
<td>(GATA)(_8)OCTA</td>
<td>63</td>
<td>30/19</td>
<td>257–313</td>
<td>0.934/0.833</td>
<td>0.189 AY636120</td>
</tr>
<tr>
<td>SarB-C05 (group 2)</td>
<td>GGCTGATGGTCTTCTGTTCTCTC</td>
<td>(TC)(<em>{5})T(TC)(</em>{4})</td>
<td>56</td>
<td>30/24</td>
<td>182–256</td>
<td>0.963/0.900</td>
<td>0.417 AY636116</td>
</tr>
<tr>
<td>Sar1-D01 (group 1)</td>
<td>GTCCTGGGATGAGGCTCTTATCTAC</td>
<td>(CA)(<em>{10})G(CA)(</em>{3})</td>
<td>61</td>
<td>30/24</td>
<td>182–256</td>
<td>0.963/0.900</td>
<td>0.417 AY636121</td>
</tr>
<tr>
<td>Sar1-D06(B) (group 1)</td>
<td>CAGCTCTCTGTTAGCTACTAGTG</td>
<td>(TG)(_{18})</td>
<td>50</td>
<td>30/14</td>
<td>120–158</td>
<td>0.794/0.700</td>
<td>0.137 AY636123</td>
</tr>
<tr>
<td>SarB-D09 (group 3)</td>
<td>CCCCCATACGATATGCAAGAGC</td>
<td>(CA)(<em>{6})(GA)(</em>{4})</td>
<td>59</td>
<td>30/15</td>
<td>269–303</td>
<td>0.905/0.967</td>
<td>0.633 AY636117</td>
</tr>
<tr>
<td>SarB-G09 (group 3)</td>
<td>GAGCGCTCTTTGCAACACTTCG</td>
<td>(GA)(<em>{8})(GT)(</em>{4})</td>
<td>61</td>
<td>30/4</td>
<td>149–163</td>
<td>0.606/0.667</td>
<td>0.421 AY636118</td>
</tr>
<tr>
<td>SarB-H04 (group 2)</td>
<td>GGATTTGTCGTCGACACGGAGG</td>
<td>(GT)(_{5})</td>
<td>52</td>
<td>30/25</td>
<td>181–271</td>
<td>0.959/0.933</td>
<td>0.745 AY636119</td>
</tr>
<tr>
<td>SarB-H04(F) (group 3)</td>
<td>CTTCCGCTCTTCGAGAGGAGGA</td>
<td>(TG)(_{18})</td>
<td>59</td>
<td>30/28</td>
<td>116–168</td>
<td>0.947/0.733</td>
<td>&lt; 0.001 AY636119</td>
</tr>
<tr>
<td>Sar1-H11(B) (group 1)</td>
<td>CAGCGGATACCTGCTTTTCAG</td>
<td>(TG)(<em>{12})T(A)(</em>{6})</td>
<td>65</td>
<td>30/21</td>
<td>150–200</td>
<td>0.937/0.633</td>
<td>&lt; 0.001 AY636122</td>
</tr>
</tbody>
</table>

*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