Living with uncertainty: genetic imprints of climate shifts in East Pacific anchovy (Engraulis mordax) and sardine (Sardinops sagax)

F. Lecomte,* W. S. Grant,† J. J. Dodson,* R. Rodríguez-Sánchez‡ and B. W. Bowen§

*Québec-Océan, Département de Biologie, Université Laval, Ste Foy, Québec, Canada, †P.O.Box 240104, Anchorage, Alaska, USA, ‡CICIMAR, Apdo. Post 592, La Paz, Baja California Sur, Mexico, §Hawaii Institute of Marine Biology, University of Hawaii, PO Box 1346, Kaneohe Bay, Hawaii, HI 96744, USA

Abstract

In the upwelling zone of the northeastern Pacific, cold nutrient-rich conditions alternate with warm nutrient-poor intervals on timescales ranging from months to millennia. In this setting, the abundances of Pacific sardine (Sardinops sagax) and northern anchovy (Engraulis mordax) fluctuate by several orders of magnitude, with sardine dominating during warm conditions and anchovy dominating during cool conditions. Two population models can explain the response of these fishes to adverse conditions. Under the basin model, species distributions contract to a central (optimal) range during population crashes. Expectations of this model may include a single range-wide population with a decline in genetic diversity on both sides of a central refuge. In contrast, the self-recruitment model invokes a series of local oceanographic domains that maintain semi-isolated subpopulations. During adverse conditions, some subpopulations cannot complete the life cycle within the local environment and are extirpated. Expectations of this model include some degree of population genetic structure and no clear gradient in genetic diversity. We examined mitochondrial DNA cytochrome b sequences to assess these competing models for anchovy (N = 196; 539 bp) and sardine (N = 107; 425 bp). The mitochondrial DNA gene genealogies are shallow but diverse for both species. Haplotype frequencies are homogeneous among subpopulations, but genetic diversities peak for both species along Baja California and adjacent southern California. Mismatch distributions and Tajima’s D-values reveal distinctive signatures of population bottlenecks and expansions. Sardine haplotypes coalesce at ~241 000 years BP, with an initial female effective population size Nf0 = 0 followed by exponential growth to Nf1 = 115 million. Anchovy haplotypes coalesce at ~282 000 years BP, with an initial population size of N0 = 14 000, followed by exponential growth to Nf1 = 2.3 million. These results indicate a founder event for sardine and a severe population decline for anchovy in the California Current during the late Pleistocene. Overall, these data support the basin model on decadal timescales, although local recruitment may dominate on shorter timescales.

Keywords: basin model, climate change, coalescence, fish, mtDNA, recruitment

Received 10 January 2004; revision received 7 April 2004; accepted 7 April 2004

Introduction

In contemporary studies of biogeography and evolution, glaciation and other long-term climate changes are regarded as primary influences on genetic diversity and speciation (Hewitt 1996; Avise 2000). Such changes typically occur on a scale of 104–106 years. Higher frequency oscillations, such as decadal variations in temperature or productivity, are principally studied under the domain of population dynamics (Cushing 1975). Beyond the description of extirpations and recolonizations, few studies have examined the impact of decadal-scale regime changes on the genetic architectures of species (Jansson & Dynesius 2002). In highly variable environments, fluctuations in abundance and distribution...
should impose distinctive genetic signatures (Slatkin & Hudson 1991; Rogers & Harpending 1992).

Coastal upwelling areas are among the most variable environments on earth. In the eastern Pacific, highly productive upwelling conditions can be reversed in a matter of months by El Niño–Southern Oscillation (ENSO) events (Glynn 1988; McPhaden 1999). Somewhat longer climate cycles of ~50 years are also apparent (Chavez et al. 2003). Changes in upwelling intensity and corresponding currents vary on several timescales (Kennett & Ingram 1995; Kieler et al. 2001), as does the overall productivity of the coastal zone (Webb & Bartlein 1992).

The impact of these regime shifts is exemplified by cycles of abundances in anchovies (Engraulis spp.) and sardines (Sardinops spp.) in temperate boundary currents (Lluch-Belda et al. 1989; Parrish et al. 1989). Anchovies reach peak abundance during the colder, highly productive regime, whereas sardines peak during warm, less productive conditions (Chavez et al. 2003). The changes in population abundance are dramatic: spawning biomass of anchovy in the California Current varied by three orders of magnitude in the latter half of the 20th century (Lo 1985). These fluctuations in species abundance drive coastal food webs (Chavez et al. 2003) and regional economies (Doucet & Einardsson 1966; Nevarez-Martinez et al. 2001). At peak abundances, sardines and anchovies account for about a quarter of fish landings worldwide (Whitehead 1985). The fishery for the Peruvian anchovy (Engraulis ringens) was once the largest harvest on the planet, but annual catches fell from over 12 million tons to < 2 million tons between 1976 and 1986, subsequently disappearing as a commercial fishery (Klyashtorin 2001). During the ensuing warm conditions, sardine abundance rose sharply (Chavez et al. 2003). Hence these fishes epitomize the uncertainty that unravels fishery management plans (Harwood & Stokes 2003).

Changes in abundance are accompanied by changes in distribution. Pacific sardine disappeared from the northern half of their range (northern California to southern Alaska; Fig. 1) in the middle of the 20th century, following years of low recruitment and intense harvesting (Murphy 1966), only to reappear in 1992 (McFarlane et al. 2000). Northern anchovy disappeared from the Gulf of California during this interval, reappearing in 1985 (Hammann & Cisneros-Mata 1989). A year later anchovy was an important local fishery with landings in excess of 2000 tons (Hammann & Cisneros-Mata 1989; Cisneros-Mata et al. 1995).

Two contrasting models have been proposed to explain changes in abundance and distribution as a function of environmental variability (Sinclair 1988; MacCall 1990). Under the local self-recruitment model, multiple semi-isolated subpopulations occur along the coastline, and each is adapted to proximal environmental conditions.

![Fig. 1 Previously defined stock distributions (boxes), sampling locations, principal spawning areas and egg-larval stage distribution for northern anchovy (Engraulis mordax) and Pacific sardine (Sardinops sagax).](image-url)
Subpopulations persist and succeed where local oceanographic settings allow completion of the life cycle. This model is consistent with the member–vagrant hypothesis (Iles & Sinclair 1982; Sinclair 1988). In this view, recent geographical expansions of sardine and anchovy would have been seeded by remnant subpopulations that have gone undetected. Subpopulation-level genetic differentiation may result, in which the level of genetic diversity at each location will largely depend on the effective size of the local deme (Kimura & Ohta 1971). In other words, self-recruitment predicts a discernible pattern in subpopulation structure, but no discernible pattern in range-wide genetic diversity.

Under the basin model, when abiotic and biotic factors reduce population size, a species’ geographical range contracts to locations where conditions are most favourable (MacCall 1990). Over longer (evolutionary) timescales and under the most severe environmental conditions, anchovies may be reduced to a single core habitat. When more favourable conditions return, the population expands into peripheral areas. Therefore, basin-model dynamics should produce little or no population structure. In contrast to self-recruitment, the basin model predicts a gradient in genetic diversity descending from optimal (refuge) habitat to the periphery of the range. This model is consistent with the source–sink hypothesis (Pulliam 1988).

To assess the genetic architectures of northern anchovy and Pacific sardine, sequences of mitochondrial DNA (mtDNA) cytochrome b were surveyed across the ranges of both species. If subpopulations are perpetuated by local recruitment, we expect to observe a pattern of population genetic differentiation across the entire range (although this pattern would not be detected with mtDNA under some conditions such as moderate gene flow or recent colonization). If a species experiences range-wide expansions from a source population, we expect to find a peak in genetic diversity at the putative source habitat. We used mtDNA sequence diversity and the coalescence approach (Hudson 1990) to evaluate postulated demographic factors such as bottlenecks, demographic explosions and effective population sizes before and after population growth (Tajima 1989; Slatkin & Hudson 1991; Rogers & Harpending 1992). Of particular interest are fish in the Gulf of California, isolated from the Pacific coast by tropical conditions at the tip of Baja California Sur (BCS; Fig. 1). Results can illuminate how sardines and anchovies cope with precipitous climate changes. Such information may assist fishery managers who are themselves operating in a regime of high uncertainty, and perhaps other conservation efforts during a period of rapid global warming.

### Materials and methods

Northern anchovy and Pacific sardine were collected between 1994 and 2001 throughout the geographical ranges of both species (Table 1, Fig. 1). Anchovy were sampled from the Washington coast (Grays Harbor), northern California (Eureka), southern California (San Diego in 1994 and 2001), Baja California (Ensenada), Upper Baja California Sur (South of Cedros Is.), Lower Baja California Sur (Bahía Magdalena) and the Gulf of California (Bahía Guaymas). Sardine were obtained from Canada (Vancouver Is.), southern California (San Diego), Baja California (Ensenada) and the Gulf of California (Bahía Guaymas). Tissues were preserved in saturated salt–DMSO buffer (Amos & Hoelzel 1991) or 95% ethanol prior to DNA extraction.

Total genomic DNA was extracted from 0.1 g of muscle or gonad tissue with a standard phenol–chloroform–isoamyl protocol (Hillis et al. 1996), precipitated with cold 95% ethanol and re-suspended in 200 µL TE buffer (10 mM Tris–HCl, 1 mM EDTA).

For anchovy, a 687 bp portion of cytochrome b (including 51 bp in the flanking tRNase) was amplified using the

### Table 1 Sampling sites, sampling date, coordinates [exact sampling sites (s) or estimated from fishing area (f)], number of fishes sequenced for northern anchovy (Engraulis mordax) and Pacific sardine (Sardinops sagax)

<table>
<thead>
<tr>
<th>Region</th>
<th>Sampling site</th>
<th>Date</th>
<th>Coordinates Latitude</th>
<th>Longitude</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Northern Anchovy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington</td>
<td>Grays Harbor</td>
<td>1995</td>
<td>46°59'07″-N</td>
<td>123°57'35″-W</td>
<td>17</td>
</tr>
<tr>
<td>North California</td>
<td>Eureka</td>
<td>1994</td>
<td>40°47'58″-N</td>
<td>124°08'35″-W</td>
<td>20</td>
</tr>
<tr>
<td>South California</td>
<td>San Diego</td>
<td>1994</td>
<td>32°49'28″-N</td>
<td>117°05'59″-W</td>
<td>14</td>
</tr>
<tr>
<td>South California</td>
<td>San Diego</td>
<td>2001</td>
<td>32°49'28″-N</td>
<td>117°05'59″-W</td>
<td>30</td>
</tr>
<tr>
<td>Baja California</td>
<td>Ensenada</td>
<td>1995</td>
<td>31°52'00″-N</td>
<td>116°37'00″-W</td>
<td>23</td>
</tr>
<tr>
<td>Upper Baja California Sur</td>
<td>South of Cedros Is.</td>
<td>1995</td>
<td>27°30'06″-N</td>
<td>114°59'35″-W</td>
<td>35</td>
</tr>
<tr>
<td>Lower Baja California Sur</td>
<td>Bahía Magdalena</td>
<td>1995</td>
<td>24°34'59″-N</td>
<td>112°00'00″-W</td>
<td>37</td>
</tr>
<tr>
<td>Gulf of California</td>
<td>Guaymas</td>
<td>1995</td>
<td>27°55'59″-N</td>
<td>110°54'00″-W</td>
<td>20</td>
</tr>
<tr>
<td><strong>Pacific sardine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>Vancouver Is.</td>
<td>1999</td>
<td>48°55'30″-N</td>
<td>126°27'24″-W</td>
<td>32</td>
</tr>
<tr>
<td>California</td>
<td>San Diego</td>
<td>2001</td>
<td>32°49'28″-N</td>
<td>117°05'59″-W</td>
<td>30</td>
</tr>
<tr>
<td>Baja California</td>
<td>Bahía de Todos Santos</td>
<td>1995</td>
<td>31°47'59″-N</td>
<td>116°41'59″-W</td>
<td>25</td>
</tr>
<tr>
<td>Gulf of California</td>
<td>Guaymas</td>
<td>1995</td>
<td>27°55'59″-N</td>
<td>110°54'00″-W</td>
<td>20</td>
</tr>
</tbody>
</table>
heavy strand degenerate primer H15149 (5'-TTGAGCCCT(C)GCTGGGTTA(G)TTAGAT; Grant & Bowen, unpublished) and a light strand primer L14724 (5'-CGAAGCTTGGATGAAAAACCATCGTTG; Irwin et al. 1991). For sardine, a 481 bp portion of cytochrome b was amplified with a heavy strand primer (5'-TGAACCTTGGAAAAACCGGTTC; Kocher & Wilson 1991) and a light strand primer (5'-AAACTGCAGCCCTCAGAATGATATTTTGCTCTCA; Song 1994). For sardine, polymerase chain reaction (PCR) conditions were as follows: initial denaturing step at 94 °C for 3 min, then 24 cycles of amplification (30 s of denaturation at 94 °C, 40 s of annealing at 54 °C and 40 s of extension at 72 °C), and a final extension of 3 min at 72 °C. Amplifications were carried out in a 25 μL reaction with 0.16 mm dNTPs, 3.5 mm MgCl2, 40 nM of each primer, and 2.5 units of Taq DNA polymerase in a 1 X reaction buffer (10 x: 100 mm Tris–HCl, 500 mm KCl, 1% Triton ×-100; Promega).

For anchovy, the conditions were similar except 34 cycles of amplification were carried out and annealing temperature was set to 50 °C. PCR products were purified with 40 000 MW filters (Millipore Corp.) and sequenced with dye-primer methods, following protocols for automated DNA sequencing (Applied Biosystems, model 310). Sequences were aligned with sequencher (version 3.0, Gene Codes Corp.). We used positive and negative PCR controls to detect contamination. Dubious sequences were reamplified and resequenced in the forward direction to assure accuracy of nucleotide designations.

We assessed population genetic structure using conventional F-statistics, \( F_{ST} \) (analogous to \( \theta \); Cockerham 1973) using 50 000 Markov chain bootstraps to estimate significance (calculated with arlequin 2.0; Schneider et al. 2000). An AMOVA was used to test for differences between the conventional stocks defined for fishery management (Fig. 1). Estimates of sequence divergence were based on the Nei (1987) and haplotype diversities, and their corresponding variances were calculated after Nei (1987) as implemented in arlequin 2.0 (Schneider et al. 2000). We used mismatch distributions for each sample to distinguish between models involving past exponential growth and historical population stasis (Slatkin & Hudson 1991; Rogers & Harpending 1992). A unimodal distribution reflects rapid growth from a small population size, while a multimodal distribution reflects long-term population stability. When a sample of haplotypes exhibited a unimodal distribution, we used summary variables defined by the distribution to estimate the onset of exponential population growth and effective population size before and after exponential growth. Using the approach developed by Li (1977) and Rogers & Harpending (1992), we fitted estimates of \( \tau, \theta_0, \theta_1 \) to the observed mismatch distribution with:

\[
\chi^2 = \sum_{i=1}^{n} \left( \frac{\bar{x}_i - \bar{x}}{V(\bar{x})} \right)^2
\]

where \( V(\bar{x}) \) is the variance, \( n \) the number of sites \( w_i \), the weighted contribution of the variance and \( \bar{x}_i \) the mean number of lineages identified for site \( i \). Samples exhibiting significantly high or low levels of lineage diversity were compared with the global mean. A jackknife approach was used to compare observed values to the distribution of values obtained by counting the number of lineages in samples of randomly selected individuals. We performed 1000 replicates and generated frequency histograms with corresponding P-values.

Haplotype mutational networks were developed using tcs version 1.13 (Clement et al. 2000) without assuming a specific evolutionary model.

Tajima’s D-test was used to identify past bottlenecks in population size or population expansions (Tajima 1989, 1996; Rand 1996). This test is a comparison of nucleotide polymorphism estimators \( \pi \) and \( \theta \) (based on the number of variable positions in the entire sample). Originally created to detect selection, negative D-values (\( \pi < \theta \)) can also indicate population bottlenecks or population expansions. Significance of D was estimated with the distribution of random samples generated using a coalescent algorithm assuming neutrality and population equilibrium (Tajima 1989), as implemented in arlequin 2.0 (Schneider et al. 2000).

We used mismatch distributions for each sample to distinguish between models involving past exponential growth and historical population stasis (Slatkin & Hudson 1991; Rogers & Harpending 1992). A unimodal distribution reflects rapid growth from a small population size, while a multimodal distribution reflects long-term population stability. When a sample of haplotypes exhibited a unimodal distribution, we used summary variables defined by the distribution to estimate the onset of exponential population growth and effective population size before and after exponential growth. Using the approach developed by Li (1977) and Rogers & Harpending (1992), we fitted estimates of \( \tau, \theta_0, \theta_1 \) to the observed mismatch distribution with:
A 539 bp segment of cytochrome \( b \) was aligned for 196 specimens of anchovy, and a 425 bp segment was aligned for 107 specimens of sardine. Sequences were aligned with cytochrome \( b \) fragments in GenBank (Benson \textit{et al.} 2003) for \textit{Engraulis} spp. (Accession nos AB040676/AF472579) and \textit{Sardinops} spp. (Accession nos AF472585/AB032554/AF472586). In total, 66 haplotypes were identified in 196 anchovy, including 51 haplotypes observed in single specimens. Forty-six of 53 (86.8\%) variable sites occurred in third codon positions. In sardine, 41 haplotypes appeared in 107 fish, including 29 haplotypes observed in single specimens. Twenty-eight of 32 (87.5\%) variable sites occurred in third codon positions.

Analyses of haplotype frequencies revealed no substantial genetic structure across the range of either species. Estimates of \( \Phi_{ST} \) for anchovy were not significant for either species, but estimates of \( F_{ST} \) revealed one significant value (\( F_{ST} = 0.027, P = 0.049 \)) for anchovy, between Magdalena Bay and northern California (opposite ends of the range). The chi square analyses with and without the pooling of rare haplotypes, \( K_{ST} \) and \( H_{ST} \) (Hudson \textit{et al.} 1992) also indicated no genetic structure (results not shown). AMOVA indicated that differences among putative stocks (see Fig. 1) explained only \( -0.33 \) and \( 1.21\% \) (both values not significant) of the total variation for anchovy and sardine, respectively. In contrast, latitudinal gradients in haplotype diversity were observed in both species (Fig. 2).

Anchovy haplotype diversity was lower (\( h = 0.74, SD = 0.09 \)) near the northern edge of its range (Grays Harbor, Washington) but significantly higher (\( h = 0.95, SD = 0.03 \)) near the southernmost boundary (Bahia Magdalena, Lower BCS). This gradient is robust to adjustments for sample size. The rate of increase in the number of mtDNA haplotypes as a function of the number of fish sampled levelled off for northern sites after bootstrap re-sampling of only 17 specimens (Grays Harbor \( \Delta_{16-17} = 0.17 \)). However, at the southern portion of its range, new lineages still appeared at a high rate at 37 specimens; in the Magdalena sample one new lineage appeared for every two additional sequences sampled from the population (\( \Delta_{36-37} = 0.46 \)) (Fig. 3a). Cochran’s G-test (Shoukri & Pause 1999) indicated significant differences between the number of haplotypes identified for anchovy after fixing sample sizes at 13 fish for each site (\( \chi^2 = 19.57, d.f. = 6, P < 0.005 \) (SD from 1000 re-samplings for each site; San Diego 1994 omitted due to small sample size). The Magdalena sample exhibited a significantly larger number of haplotypes (\( I_{37} = 26 \)) than expected (\( I_{37} = 15.9, P < 0.001 \)), and the Grays Harbor sample had a significantly smaller number of haplotypes (\( I_{17} = 6 \)) than expected (\( I_{17} = 10.2, P < 0.025 \)).
In contrast, sardine haplotype diversity peaked in the middle of the geographical range, reaching the highest value \( h = 0.92, \text{SD} = 0.04 \) at San Diego and the lowest at Guaymas \( h = 0.83, \text{SD} = 0.07 \). This trend was also robust to adjustments for sample sizes (Figs 2 and 3b). Significant differences were observed among locations in the number of haplotypes \( \chi^2 = 10.61, \text{d.f.} = 3, P = 0.01 \) appearing in samples after fixing sample sizes at 17 fish for each location. One new lineage appeared in every five fish sampled at Guaymas, but three new lineages appeared in every five fish at San Diego (Fig. 3b). The number of haplotypes in the sample from Guaymas was significantly smaller \( \sim 9 \) than expected \( \sim 13.6, P = 0.010 \).

In both species, Tajima D’s were negative at all sites (Table 2). These values were highly significant in pooled samples for each species (Table 2). Simulations revealed that D-values were heterogeneous over the geographical range of anchovy, and only two of the nine samples were significant. In contrast, sardine D-values were homogeneous and uniformly significant.

Sequence mismatch distributions were unimodal in pooled samples of both species, consistent with a model of rapid population growth from a small number of fish. Individual samples were unimodal for sardine, but the pattern was more complex in anchovy (Fig. 4). First, the crests of the distribution for anchovy were closer to the origin at northern locations, indicating more recent population disturbances than in the southern regions. Second, the distributions in northern areas tended to be bimodal, possibly reflecting a haphazard representation of haplotype lineages among colonizers to these areas. Because of the ephemeral nature of northern habitat (as indicated by historical records), we focus the description of coalescence results on the southern–central part of each range, which are likely to be the oldest segments for both species. We used the conventional calibration of 2% sequence divergence per million years between lineages, and generation times of 1.9 years for anchovy (Butler et al. 1993) and 4.4 years for sardine (Murphy 1967; Butler et al. 1993) to estimate the beginning of rapid population growth. For the Magdalena sample of anchovy, this estimate is \( \sim 282 \text{,000 years} \) (95% CI between 144,000 and 447,000), and for sardine at San Diego, it is \( \sim 241 \text{,000 years} \) (95% CI between 200,000 and 373,000). The estimated initial anchovy female effective population size \( N_0 \) from the Magdalena sample is \( \sim 14 \text{,000 individuals} \) (95% CI: 0 to \( \sim 241 \text{,000} \)) and the estimated effective population size following exponential growth \( N_1 \) is \( \sim 2.3 \text{ million} \) (95% CI: 472,000 to \( \sim 364 \text{,000,000} \)). In contrast, \( N_0 \) for sardine at San Diego is \( 0.0 \) (95% CI: 0 to \( \sim 6000 \)), and \( N_1 \) is \( \sim 115 \text{ million} \) (95% CI: 286,000 to \( \sim 247 \text{,000,000} \)).

Anchovy and sardine have shallow gene genealogies along the west coast of North America (Fig. 5), with maximum sequence divergences between haplotypes of \( d = 1.88\% \) for anchovy and \( d = 1.67\% \) for sardine, and with no discernible phylogeographical structure. Parsimony networks for each species consisted of two closely related lineages, surrounded by low-frequency haplotypes (Fig. 5), typical of the starburst pattern of young growing populations. None of the common haplotypes were geographically restricted, a pattern consistent with high levels of gene flow or recent geographical expansion.

**Discussion**

In this study we assess the impact of climate fluctuations on the genetic architectures of northern anchovy and Pacific sardine. Our results generally support the basin model of range contraction and expansion, at least over the timeframe that can be resolved with mtDNA sequence analyses. Notably, Dawson et al. (2001) reached a similar conclusion for the tidewater goby, a euryhaline fish that inhabits lagoons along the California coast.
Prior to dissecting these results, we temper the conclusions with four prominent caveats:

1. The sampling regime is somewhat lopsided, with nine samples for anchovy, and four for sardine. Hence conclusions for Pacific sardine must be considered provisional.

2. We support the basin model for population trends over decades and centuries, but do not rule out a role for self-recruitment on a scale of years or generations. Larger samples and microsatellite studies may be profitably employed here (see Toonen 2001).

3. In fishes that number in the billions at peak abundance, drastic reductions of three orders of magnitude will not produce the bottleneck effect of classic population genetic theory (Nei et al. 1975). However, these crashes will reduce overall diversity and corresponding estimates of effective population size. Estimates of effective population size are based on a harmonic mean of reproductive individuals in each generation, and therefore are especially sensitive to population crashes (Maruyama & Kimura 1980).

4. The ratio of contemporary census size to effective population size of females ($N_{ef}$ estimated here by mtDNA diversity) may range from 6–8 in mammals with relatively uniform reproductive success (Roman & Palumbi 2003) to at least $10^2$ in fishes with pelagic eggs and larvae (Bowen & Avise 1990). The latter condition (or a more extreme ratio) clearly applies here.

The answers to several questions about the evolution of marine species in upwelling zones (and their management

© 2004 Blackwell Publishing Ltd, Molecular Ecology, 13, 2169–2182

---

Table 2 Percentage of rare haplotypes (< 1%), haplotype diversity ($h$), nucleotide diversity ($π$), diversity in segregating sites ($θ$), Tajima $D$-value, corresponding $P$-value, and mismatch distribution parameter estimates

<table>
<thead>
<tr>
<th>Sampling site</th>
<th>% rare haplo. (SD)</th>
<th>$h$ (SD)</th>
<th>$π$ (SD)</th>
<th>$θ$ (SD)</th>
<th>$D$</th>
<th>$P$</th>
<th>$τ$ (age)</th>
<th>$θ_1 (N_{f1})$</th>
<th>$θ_1 (N_{f1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Northern anchovy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grays Harbor</td>
<td>5.9</td>
<td>0.743</td>
<td>0.004</td>
<td>0.003</td>
<td>−0.328</td>
<td>0.393</td>
<td>3.932</td>
<td>0.001</td>
<td>3.775</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(365 000)(48)</td>
<td></td>
<td>(0.181 × 10⁶)</td>
</tr>
<tr>
<td>Eureka</td>
<td>25.0</td>
<td>0.779</td>
<td>0.004</td>
<td>0.004</td>
<td>−0.755</td>
<td>0.245</td>
<td>3.684</td>
<td>0.005</td>
<td>3.568</td>
</tr>
<tr>
<td></td>
<td>(0.083)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(342 000)(242)</td>
<td></td>
<td>(0.17 × 10⁶)</td>
</tr>
<tr>
<td>San Diego 94</td>
<td>35.7</td>
<td>0.857</td>
<td>0.004</td>
<td>0.004</td>
<td>−0.842</td>
<td>0.221</td>
<td>2.912</td>
<td>0.001</td>
<td>9.824</td>
</tr>
<tr>
<td></td>
<td>(0.077)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(270 000)(48)</td>
<td></td>
<td>(0.47 × 10⁶)</td>
</tr>
<tr>
<td>San Diego 01*</td>
<td>36.7</td>
<td>0.915</td>
<td>0.005</td>
<td>0.007</td>
<td>−1.326</td>
<td>0.092</td>
<td>3.89</td>
<td>0.008</td>
<td>4.949</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(361 000)(387)</td>
<td></td>
<td>(0.46 × 10⁶)</td>
</tr>
<tr>
<td>San Diego Total</td>
<td>36.4</td>
<td>0.899</td>
<td>0.005</td>
<td>0.007</td>
<td>−1.444</td>
<td>0.070</td>
<td>3.255</td>
<td>0.005</td>
<td>9.441</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(302 000)(242)</td>
<td></td>
<td>(0.46 × 10⁶)</td>
</tr>
<tr>
<td>Ensenada*</td>
<td>21.7</td>
<td>0.810</td>
<td>0.004</td>
<td>0.006</td>
<td>−1.594</td>
<td>0.049</td>
<td>1.267</td>
<td>1.938</td>
<td>3.151</td>
</tr>
<tr>
<td></td>
<td>(0.080)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(118 000)(93 683)</td>
<td></td>
<td>(0.15 × 10⁶)</td>
</tr>
<tr>
<td>South of Cedros Is.</td>
<td>22.6</td>
<td>0.803</td>
<td>0.004</td>
<td>0.006</td>
<td>−1.525</td>
<td>0.057</td>
<td>0.875</td>
<td>1.601</td>
<td>12.554</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(81 000)(77 392)</td>
<td></td>
<td>(0.61 × 10⁶)</td>
</tr>
<tr>
<td>Bahia Magdalena</td>
<td>51.4</td>
<td>0.949</td>
<td>0.006</td>
<td>0.010</td>
<td>−1.856</td>
<td>0.021</td>
<td>3.036</td>
<td>0.288</td>
<td>46.533</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(282 000)(13 922)</td>
<td></td>
<td>(2.25 × 10⁶)</td>
</tr>
<tr>
<td>Guaymas</td>
<td>25.0</td>
<td>0.890</td>
<td>0.005</td>
<td>0.005</td>
<td>−1.116</td>
<td>0.141</td>
<td>3.255</td>
<td>0.005</td>
<td>9.441</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(302 000)(242)</td>
<td></td>
<td>(0.46 × 10⁶)</td>
</tr>
<tr>
<td><strong>Global</strong></td>
<td>31.3</td>
<td>0.855</td>
<td>0.005</td>
<td>0.014</td>
<td>−2.166</td>
<td>0.004</td>
<td>3.352</td>
<td>0.043</td>
<td>6.167</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(311 000)(2078)</td>
<td></td>
<td>(0.29 × 10⁶)</td>
</tr>
<tr>
<td><strong>Pacific sardine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vancouver Is.</td>
<td>45.5</td>
<td>0.897</td>
<td>0.005</td>
<td>0.008</td>
<td>−1.721</td>
<td>0.033</td>
<td>2.343</td>
<td>0.000</td>
<td>116.4</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(276 000)(0)</td>
<td></td>
<td>(3.1 × 10⁶)</td>
</tr>
<tr>
<td>San Diego</td>
<td>52.4</td>
<td>0.929</td>
<td>0.005</td>
<td>0.008</td>
<td>−1.747</td>
<td>0.030</td>
<td>2.053</td>
<td>0.000</td>
<td>4350.0</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(241 000)(0)</td>
<td></td>
<td>(115.3 × 10⁶)</td>
</tr>
<tr>
<td>Bahia de Todos Santos</td>
<td>31.6</td>
<td>0.880</td>
<td>0.004</td>
<td>0.006</td>
<td>−1.683</td>
<td>0.037</td>
<td>1.854</td>
<td>0.000</td>
<td>3388.7</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(218 000)(0)</td>
<td></td>
<td>(89.8 × 10⁶)</td>
</tr>
<tr>
<td>Guaymas</td>
<td>21.4</td>
<td>0.826</td>
<td>0.003</td>
<td>0.005</td>
<td>−1.825</td>
<td>0.023</td>
<td>1.499</td>
<td>0.000</td>
<td>3488.75</td>
</tr>
<tr>
<td></td>
<td>(0.073)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(176 000)(0)</td>
<td></td>
<td>(92.4 × 10⁶)</td>
</tr>
<tr>
<td><strong>Global</strong></td>
<td>39.5</td>
<td>0.885</td>
<td>0.005</td>
<td>0.012</td>
<td>−2.063</td>
<td>0.008</td>
<td>2.004</td>
<td>0.000</td>
<td>4330.00</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(229 000)(0)</td>
<td></td>
<td>(83.1 × 10⁶)</td>
</tr>
</tbody>
</table>

* Mismatch analyses with one individual removed randomly from the sample (Least square procedure did not converge after 1800 steps with the original sample set).
in coastal fisheries) depend on the relevant timescale. On an ecological timescale, populations of anchovies, sardines and many other marine species may be sustained primarily by local recruits, even though these species show range-wide connectivity in population genetic assays (Cowen et al. 2000; Hellberg et al. 2002; Swearer et al. 2002). On an evolutionary timescale, species with high population self-recruitment are expected to fragment into new taxa more readily than species with long-distance dispersal.

**Population genetic structure — northern anchovy**

Although northern anchovy show a large potential for gene flow through larval or adult dispersal, some biological and genetic information indicates population isolation. Previous studies of proteins (Vrooman et al. 1981), allozymes (Hedgecock et al. 1989, 1994; Hedgecock 1991, 1994; Díaz-Jaimes et al. 1999) and karyotypes (Uribe-Alcocer et al. 1996) have revealed shallow but significant subpopulation structure along the west coast of North America. These

---

*Fig. 4 Mismatch distribution for each sampling site for northern anchovy (*Engraulis mordax*) and Pacific sardine (*Sardinops sagax*). *n* refers to the number of pairwise comparisons, *P*-values compare the observed distribution (bars) to the distribution modelled for sudden exponential growth (curves). Values < 0.05 reject the model of sudden exponential growth. The y-axis represents the relative frequency of classes of pairwise comparisons (scaleless).*
data generally support the recognition of northern, central and southern stocks (Fig. 1). Furthermore, analysis of allozyme variability within the central stock indicates heterogeneity among schools, age classes and sexes (Hedgecock et al. 1994), the kind of chaotic genetic patchiness that characterizes many marine invertebrates (Johnson & Black 1984). This genetic patchiness is not geographically localized, so that it cannot be used to define stocks for fishery management.

In this study, conventional population statistics, $F_{ST}$ and $\Phi_{ST}$, for anchovy were not significant, except for a single comparison between opposite ends of the species’ range. The dearth of subpopulation structure is also reflected in the AMOVA, which failed to detect significant levels of mtDNA differentiation among putative stocks. This outcome may be due in part to low statistical power under conditions of high haplotype diversity. However, the mtDNA and allozyme data are not highly discordant, as the former indicates homogeneity, and the latter indicates near homogeneity (very shallow structure and patchiness). As noted by Hedgecock et al. (1989), the structure observed in allozyme studies may be “resulting from ecological, not historical processes”. The lack of mtDNA haplotype heterogeneity in our study is consistent with these results. No haplotype frequency differences were detected and an AMOVA failed to detect significant levels of differentiation among putative stocks.

**Diversity gradients**

In contrast to the subpopulation homogeneity described above, mtDNA diversity varied significantly among populations of both species. We observed a diversity peak in the south-central portion of the range and a decline in both directions along the coast (Fig. 2). This pattern is not apparent in previous allozyme surveys (Hedgecock et al. 1989; for example), but diploid loci are expected to be less sensitive in this regard: a four-fold decrease in population size would be needed to show a reduction comparable with that observed in haploid mtDNA.

The mtDNA diversity gradients most likely arose from extirpations and recolonizations in marginal areas. Two recent cases of population recovery illustrate this process. The anchovy was rare in the Gulf of California prior to 1985, but abundances suddenly increased so that by 1986 this species supported a commercial fishery of 2086 metric
tons (Cisneros-Mata et al. 1995). In a similar fashion, the sardine north of San Francisco disappeared in the late 1940s (Murphy 1966; Schweigert 1988). Their reappearance in Canadian and Alaskan waters in the early 1990s was followed by a rapid expansion to a stock size of ~89 000 metric tons in 1997 (McFarlane & Beamish 2001). In both cases, mtDNA diversities in the newly abundant stock are expected to be reduced compared with those in neighbouring regions.

The basin model is supported by mtDNA data, but this model is not sufficient to describe the long-term dynamics of anchovy and sardine populations. For example, the basin model assumes a single stationary refuge during adverse conditions. Baumgartner et al. (1992) used this assumption to infer historical population fluctuations from the abundance of scales in sediment cores from the Santa Barbara Basin. If the location of the basin population moves in response to climate and oceanographic conditions, scale abundance in sediment cores may reflect changes in geographical range rather than historical abundance. In support of a more complex basin model, Rodríguez-Sánchez et al. (2001, 2002) demonstrated that during periods of ocean cooling between the early 1940s and 1975, the centre of abundance for sardine shifted from California to Baja California. If basin populations shift along the coast in response to climate variability, historical inferences based on sediment cores may have to be re-evaluated.

On decadal timescales, population expansions at the periphery of the distribution may be the result of growth by depleted, but locally recruiting subpopulations. Alternatively, marginal subpopulations may be extirpated during climate shifts, and then reseeded by fish from neighbouring populations. Recent explosions of anchovy and sardine in the Gulf of California may help to distinguish these alternatives. Rodríguez-Sánchez et al. (2001) postulated that a southerly shift in the centre of abundance of sardine during the early 1940s to 1960s led to the invasion of sardine into the Gulf of California. However, Holmgren-Urba & Baumgartner (1993) demonstrated that anchovy have been present in the Gulf (at low densities) for at least 250 years. If this small population recently expanded in isolation, we would expect the signatures of strong genetic drift and corresponding haplotype frequency shifts, neither of which are apparent in our data sets. Allozyme studies by Hedgcock et al. (1989) likewise did not detect an isolated population. Therefore, we discount the self-recruitment model for Gulf populations of sardine and anchovy, and support the hypothesis that the boom in Gulf populations is coupled with connectivity to the Pacific coast subpopulations. The monotonic decline in mtDNA diversity from the central part of the ranges of both species offers further support for colonizations, rather than local recruitment.

On longer timescales of the late Pleistocene, palaeoclimatic data indicate progressively larger climatic swings with massive continental glaciations (Crowley & North 1991), which lowered sea levels to the edge of the continental shelf (Lambeck et al. 2002) and influenced current patterns and upwelling intensities (Rahmstorf 2002). Our analyses of mtDNA haplotypes indicate contraction and population growth on the same timescale. Values of Tajima’s D were negative for all samples, but were significant (i.e. more negative) only in southern anchovy subpopulations. This finding may be due to lower levels of haplotype diversity, which contributed to lower values of \( \theta_0 \), compared to \( \pi \), and thus produced D-values close to 0 (Table 2). Because haplotypes were only a few mutations apart, \( \pi \) is necessarily small, and D largely depends on the magnitude of \( \theta_0 \) (see Table 2). In view of the fluctuations recorded in anchovy and sardine fisheries, the negative D-values most likely reflect population crashes and recoveries, rather than selective sweeps (Tajima 1989; Rand 1996).

Population histories — crashes and colonizations

Conclusions about effective population size and exponential growth depend on estimates of molecular clock rates and generation times. How robust are these estimates? The divergence rate of 2%/Myr for cytochrome b is calibrated by a vicariant separation of anchovies Ctenognathus dentatus and C. mysticus across the Isthmus of Panama (Lecomte et al. unpublished data) and for sardines by the separation of genera Sardinops and Sardina after the closure of Tethys Sea (Grant & Bowen 1998). Estimates of generation time are difficult to calculate for short-lived fish that spawn repetitively throughout the year (every 6.5–10.6 days for northern anchovy; Bindman 1985 from Butler et al. 1993). Anchovies and sardines both mature in less than 2 years, but anchovies seldom reach 4 years of age, whereas sardines can exceed 8 years of age. The estimated generation time of 4.4 years for Pacific sardine is probably appropriate for unexploited populations (Murphy 1966), but may be an upper bound for populations under stress from fishing pressure or environmental change. Reducing the sardine generation time from 4.4 to 2.2 years (for example) would not alter the conclusions about the age of the demographic expansion, but would diminish \( N_f \) estimates (based on \( \theta_0 \) and \( \theta_1 \)) by half. With this caution, we feel that the estimates of clock rate and generation time are sufficient to support our qualitative conclusions.

Northern anchovy have persisted in the California Current for at least five million years, based on palaeontological data (Fitch 1969) and genetic comparisons to sister taxa (Grant & Bowen 1998). This provides an evolutionary backdrop to interpret the coalescence analysis for anchovies. Estimates from the mtDNA data indicate a period of rapid population growth 290 000 years BP, from an effective population of \( N_f = 14 \ 000 \). Although this indicates an extreme reduction in census size relative to periods of peak abundance, the reduction was not sufficient to deplete
The anchovy and sardine, living with uncertainty in the California Current, offer new perspectives at the crossroads of evolution and wildlife management. The collapse and recovery of regional populations, with corresponding genetic homogenization, indicates that macro-evolutionary partitions are unlikely to arise within this region. The arena for speciation is clearly among the isolated upwelling regions of Chile, Australia, South Africa, Japan and (Baja) California, all of which are occupied by members of the genera Engraulis and Sardinops (Bowen & Grant 1997; Grant & Bowen 1998). Contemporary taxonomy of sardines and anchovies is consistent with this conclusion (Whitehead 1985; Whitehead et al. 1988).

How do these pelagic fishes respond to climate variability? On decadal timescales, fishery data indicate contractions of species ranges to a central refuge ‘basin’ (Rodríguez-Sánchez et al. 2002), while genetic data bear the imprints of subsequent geographical expansions. This process yields the shallow genealogies observed with mtDNA sequences. On millennial timescales, global coolings and extreme climatic deteriorations have led to the cessation of upwelling and to the collapse of primary and secondary productivity. These disturbances precipitate major population crashes that are recorded in the genetic architectures of both anchovies and sardines. The sardines may be more prone to extinction, as indicated by their shallow history in the northeast Pacific (quarter million years) relative to the anchovy (five million years).

These fishes do not show phylogeographical partitions in the northeastern Pacific, as observed in several inshore species (Burton 1998; Bernardi 2000; Dawson et al. 2001), yet significant differences exist in the distribution of their genetic diversity. Allozyme studies of anchovy reveal shallow and transient genetic heterogeneity (Hedgecock et al. 1994), while the mtDNA studies reveal diversity gradients in both species. Neither observation can be explained in the framework of a single randomly mating population. This raises a subtle point: northern anchovy and Pacific sardine are not divided into geographically structured subpopulations, but neither do they form single panmictic populations. While the genetic data support the basin model over a decadal scale of range contraction and expansion, local regimes may be largely self-seeding on a generation-to-generation timescale.

Wildlife management must consider the prosperity of individual species, as well as the ecosystems that support these species. While there is little chance of benignly altering the physical processes that create healthy upwelling ecosystems, the implementation of harvest limits can allow local subpopulations to reseed themselves over the short-term. On longer timescales, the conservation of these pelagic species depends on an understanding of micro-evolutionary processes, specifically the role of the refuge basin that allows persistence of a core population. Disruption of the core population could be a recipe for collapse and extinction.

Acknowledgements

This study was funded by a National Science Foundation grant (DEB-9727048 to BWB) jointly awarded by the Population Biology...
and Biological Oceanography Divisions, and by a sabbatical grant (to WSG) by the Foundation for Research Development, Pretoria, South Africa. FL was supported by a PhD Internship grant from Fonds pour la Formation des Chercheurs et l’Aide à la Recherche, Québec, Canada, and by a research grant from the Natural Sciences and Engineering Research Council of Canada to JJD. RR holds a fellowship from COFAA-IPN, Mexico. Special thanks to those who provided samples, including G. A. MacFarlane, B. Andrew, J. Graves, P. E. Smith, R. Hensley, J. Shaklee, and those F. Allendorf who provided logistic support and helpful discussion, including A.L. Bass, D.E. Campton, M. Courtney, G. Daigle, M. Grant, S.A. Karl, S. Ortega, C. Rocha, L.A. Rocha, P. Taylor, R.J. Toonen, R. Waples, Graham Wallis and two anonymous reviewers.

References


Clark FN, Janssen JFJ (1945) Movements and abundance of the sardine as measured by tag returns. California Department of Fish and Game, Fishery Bulletin, 61, 7–42.


MacCall AD (1990) *Dynamic Geography of Marine Fish Populations*. Washington Sea Grant Program, Seattle.


Rodríguez-Sánchez R, Lluch-Belda D, Villalobos H, Ortega-Garcia S (2002) Dynamic geography of small pelagic fish populations in...
Schramm CT (1985) Implications of radiolarian assemblages for the Late Quaternary paleoceanography of the eastern equatorial Pacific. Quaternary Research, 24, 204–218.


This study is part of a PhD internship project (by FL, supervised by BWB) on the phylogeography and molecular ecology of clupeoid fishes. Frederic Lecomte’s major interests are related to the molecular ecology and evolution of small pelagic fishes. Stewart Grant has a career-long interest in the evolution and zoogeography of marine organisms. Julian Dodson’s interests include the evolutionary ecology of fishes, life history evolution, early life-history ecology and the conservation of aquatic resources. Rubén Rodríguez is interested in the ecology and management of coastal marine fishes. Brian Bowen studies the phylogeography and conservation genetics of marine organisms at Hawaii Institute of Marine Biology.