Effects of Pleistocene climatic fluctuations on the phylogeographic and demographic histories of Pacific herring (Clupea pallasii).

Permalink
https://escholarship.org/uc/item/86h9n3s2

Journal
Molecular ecology, 20(18)

ISSN
0962-1083

Authors
Liu, Jin-Xian
Tatarenkov, Andrey
Beacham, Terry D
et al.

Publication Date
2011-09-01

DOI
10.1111/j.1365-294x.2011.05213.x

License
https://creativecommons.org/licenses/by/4.0/
Effects of Pleistocene climatic fluctuations on the phylogeographic and demographic histories of Pacific herring (Clupea pallasii)

JIN-XIAN LIU,* ANDREY TATARENKOV,* TERRY D. BEACHAM,† VICTOR GORBACHEV,‡ SHARON WILDES§ and JOHN C. AVISE*

*Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA, †Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, BC V9T 6N7, Canada, ‡Laboratory of Population Genetics, Institute of Biological Problems of the North FEB RAS, 18 Portovaya St., Magadan, 685000, Russia, §Alaska Fisheries Science Center, 17109 Point Lena Loop Road, Juneau, AK 99801, USA

Abstract

We gathered mitochondrial DNA sequences (557 bp from the control region in 935 specimens and 668 bp of the cytochrome b gene in 139 specimens) of Pacific herring collected from 20 nearshore localities spanning the species’ extensive range along the North Pacific coastlines of Asia and North America. Haplotype diversity and nucleotide diversity were high, and three major phylogeographic lineages (sequence divergences ca. 1.5%) were detected. Using a variety of phylogenetic methods, coalescent reasoning, and molecular dating interpreted in conjunction with paleoclimatic and physiographic evidence, we infer that the genetic make-up of extant populations of C. pallasii was shaped by Pleistocene environmental impacts on the historical demography of this species. A deep genealogical split that cleanly distinguishes populations in the western vs. eastern North Pacific probably originated as a vicariant separation associated with a glacial cycle that drove the species southward and isolated two ancestral populations in Asia and North America. Another deep genealogical split may have involved either a vicariant isolation of a third herring lineage (perhaps originally in the Gulf of California) or it may have resulted simply from the long coalescent times that are possible in large populations. Coalescent analyses showed that all the three evolutionary lineages of C. pallasii experienced major expansions in their most recent histories after having remained more stable in the preceding periods. Independent of the molecular calibration chosen, populations of C. pallasii appear to have remained stable or grown throughout the periods that covered at least two major glaciations, and probably more.

Keywords: Clupea pallasii, mitochondrial DNA, phylogeography, Pleistocene glaciation, population demography

Received 14 March 2011; revision received 14 June 2011; accepted 17 June 2011

Introduction

The Pleistocene epoch was dominated by repeated episodes of global cooling and an expansion of continental ice sheets across much of North America and Eurasia (Lambeck et al. 2002). These climatic fluctuations strongly influenced the distribution and abundance of terrestrial and freshwater species in the Northern Hemisphere (Bernatchez & Wilson 1998; Avise 2000; Hewitt 2000, 2004). Although marine organisms were thought to be less vulnerable to such climatic oscillations (Hewitt 2000), recent studies indicate that Pleistocene glaciations had large impacts on coastal marine species also (Avise 2000; Liu et al. 2007; Wilson & Veraguth 2010).
The North Pacific Ocean that borders Asia and North America was greatly affected by Pleistocene glaciations and their associated impacts on sea temperatures, global sea levels and shoreline configurations (Morley & Dworetzky 1991). During each major cycle of glacial advance and retreat, extensive ice sheets episodically covered much of the continental shelf between the Alaskan Peninsula and British Columbia, and portions of the Kamchatka Peninsula (Mann & Hamilton 1995; Grosswald 1998; Kaufman & Manley 2004), while fluctuating sea levels periodically exposed and then flooded large portions of the continental shelf in the Bering Sea and north-eastern Asia (Mann & Hamilton 1995; Lambeck et al. 2002). The glacial-interglacial cycles with their alternating configurations of ice sheets and shorelines also caused pronounced fluctuations not only in sea surface temperature (CLIMAP 1981; Herbert et al. 2001; Lyle et al. 2001) but also in oceanic salinity, upwelling intensity and primary production (Sancetta & Silvestri 1984; Sancetta et al. 1992; Herbert et al. 2001; Marret et al. 2001). All of these and related environmental factors must have had huge impacts on the distribution and abundance of marine species in the North Pacific region (Fields et al. 1993; Roy et al. 1996; Graham et al. 2003). For example, Hubbs (1948) concluded that the presence of several relict fish species in the upper Gulf of California was the result of Pleistocene cooling (which drove marine species southward) and subsequent warming (which then trapped populations in the Gulf). However, the nature of biogeographic responses to glaciation probably has varied considerably among taxa (Arndt & Smith 1998; Marko 2004; Hickerson & Cunningham 2005), with some species likely ‘floating’ with their habitats as the latter shifted across latitudes but with other species remaining more stationary as they adapted to their altered local environments (Marko et al. 2010).

The Pacific herring, *Clupea pallasii*, is a pelagic-foraging marine fish that inhabits coastal waters and marginal seas on both sides of the North Pacific Ocean, at latitudes ranging from nearly subtropical to Arctic (Hay & McCarter 1997). In this extremely abundant species, mass spawning typically occurs from late winter to early summer in sheltered inlets, sounds, bays and estuaries (Haegle & Schweigert 1985; Hay 1985), with the timing of the spawn being earlier in the south and progressively later at higher latitudes (Hay 1985). Spawning aggregations are highly conspicuous: melt from millions of individuals sometimes turns the water milky-white across many kilometres (Hay et al. 2008). However, spawning sites are distributed in patchwork fashion with some spawning arenas separated by vast stretches of unsuitable reproductive habitat. Adhesive ova are deposited on marine vegetation and rocky substrates in intertidal and shallow subtidal zones (Hay 1985). Fertilized eggs hatch in 2–3 weeks, and the pelagic larvae typically metamorphose into juveniles about 2–3 months later (Hay 1985). After spending their first summer in inshore waters, juveniles gather in large schools and move offshore until maturation (Stocker et al. 1985), or some may remain in inshore waters until their first spawning event (Hay 1985). Age of sexual maturity is generally 2–5 years and increases with latitude (Hay 1985).

Based on its life history, we can speculate on how Pleistocene climatic variation might have impacted the demography and distribution of Pacific herring. This species prefers sheltered inlets and estuaries, with its current centre of abundance in the north-eastern Pacific stretching from British Columbia to southern Alaska (the coast south of British Columbia has fewer estuaries and bays). However, the species’ prime spawning area today was covered with an ice sheet during glacial maxima and thus presumably was much reduced or even unavailable for spawning during much of the Pleistocene. Another negative population effect of glaciations on herring populations may have been a reduction in the continental shelf area over which herring feed. The continental shelf is narrow along the Pacific coast of North America (especially south of British Columbia), and it would have been largely exposed during glacial maxima, thus reducing both feeding and spawning habitat. In the north-western Pacific, shelf areas along Kamchatka’s east coast likewise would have been greatly reduced. On the other hand, reduced sea surface temperature (SST) during glaciations may have opened new areas further to the south and may have compensated for the loss of suitable habitats in the north. [For example, SST was nearly 10 °C lower off northern California and 4 °C lower off Baja California, such that SST near Santa Barbara during glacial maxima would correspond to modern SST in Oregon (Herbert et al. 2001).] Another positive effect of cooling seas on herring populations might have been increased upwelling (Lyle et al. 2010), which increases the production of plankton that herring eat. Due in part to such opposing environmental influences, the overall effect of Pleistocene climatic cooling on herring populations thus has been uncertain.

An allozyme comparison with Atlantic herring, *Clupea harengus*, suggests that ancestors of *C. pallasii* entered the Pacific Ocean from the Atlantic soon after the opening of the Bering Strait about 3.5 million years ago (Ma) (Grant 1986). The pan-Pacific distribution and the life-history characteristics of *C. pallasii* would seem to make this high-latitude species especially susceptible to Pleistocene climatic fluctuations. Based on allozyme evidence, Grant & Utter (1984) distinguished two
genetic races of \textit{C. pallasii} (in the Asian–Bering Sea vs. the eastern North Pacific). Additional genetic subdivisions may exist within each of these genetic races that may be difficult to identify using approaches that utilize only information on allele frequencies without knowledge about genetic relationships between alleles. Nonrecombining rapidly evolving mtDNA can offer additional insights into past isolations (e.g. in ice-age refugia) and secondary contacts between previously isolated populations (Avise 2000; Hewitt 2000). Furthermore, variation in nucleotide level is highly amenable for inferences about past demographic history of the species (Drummond \textit{et al.} 2005; Ho & Shapiro 2011). Demographic studies on shorter (decadal) time scales have shown that recruitment in \textit{C. pallasii} can vary sharply in response to SST fluctuations (Schweigert 1995; Nagasawa 2001). Such short-term oscillations in abundance suggest that more dramatic paleoclimatic fluctuations of the Pleistocene may have affected \textit{C. pallasii} populations in a profound way, so that genomic signatures are strong enough to detect the extent and direction of those population size changes.

Here, we analyse mitochondrial (mt) DNA sequences from nearly 1000 specimens of \textit{C. pallasii} sampled across the species’ vast range. Our dual goals are to obtain a comprehensive picture of the demographic and phylogeographic structure of this species, and thereby better understand how Pleistocene climate oscillations may have affected that population structure.

\textbf{Materials and methods}

\textit{Sampling and sequencing}

\textit{Clupea pallasii} specimens were collected from spawning and prespawning aggregations at 20 localities across the species nearshore range in test fisheries or by research vessels during 1999–2010 (Table 1, Fig. 1). Tissue samples were preserved in either 95% ethanol or a saline solution (20% dimethyl sulfoxide, 0.25 M EDTA, saturated with NaCl, pH 8.0) (Seutin \textit{et al.} 1991). Genomic DNA was isolated from tissue samples. The first hypervariable segment, which was identical in length to the fragment amplified in the present study, was extracted and used for further analysis. The first hypervariable segment, which was identical in length to the fragment amplified in the present study, may be difficult to identify using approaches that utilize only information on allele frequencies without knowledge about genetic relationships between alleles. Nonrecombining rapidly evolving mtDNA can offer additional insights into past isolations (e.g. in ice-age refugia) and secondary contacts between previously isolated populations (Avise 2000; Hewitt 2000). Furthermore, variation in nucleotide level is highly amenable for inferences about past demographic history of the species (Drummond \textit{et al.} 2005; Ho & Shapiro 2011). Demographic studies on shorter (decadal) time scales have shown that recruitment in \textit{C. pallasii} can vary sharply in response to SST fluctuations (Schweigert 1995; Nagasawa 2001). Such short-term oscillations in abundance suggest that more dramatic paleoclimatic fluctuations of the Pleistocene may have affected \textit{C. pallasii} populations in a profound way, so that genomic signatures are strong enough to detect the extent and direction of those population size changes.

Here, we analyse mitochondrial (mt) DNA sequences from nearly 1000 specimens of \textit{C. pallasii} sampled across the species’ vast range. Our dual goals are to obtain a comprehensive picture of the demographic and phylogeographic structure of this species, and thereby better understand how Pleistocene climate oscillations may have affected that population structure.

\textit{Materials and methods}

\textit{Sampling and sequencing}

\textit{Clupea pallasii} specimens were collected from spawning and prespawning aggregations at 20 localities across the species nearshore range in test fisheries or by research vessels during 1999–2010 (Table 1, Fig. 1). Tissue samples were preserved in either 95% ethanol or a saline solution (20% dimethyl sulfoxide, 0.25 M EDTA, saturated with NaCl, pH 8.0) (Seutin \textit{et al.} 1991). Genomic DNA was isolated from tissue samples. The first hypervariable segment, which was identical in length to the fragment amplified in the present study, may be difficult to identify using approaches that utilize only information on allele frequencies without knowledge about genetic relationships between alleles. Nonrecombining rapidly evolving mtDNA can offer additional insights into past isolations (e.g. in ice-age refugia) and secondary contacts between previously isolated populations (Avise 2000; Hewitt 2000). Furthermore, variation in nucleotide level is highly amenable for inferences about past demographic history of the species (Drummond \textit{et al.} 2005; Ho & Shapiro 2011). Demographic studies on shorter (decadal) time scales have shown that recruitment in \textit{C. pallasii} can vary sharply in response to SST fluctuations (Schweigert 1995; Nagasawa 2001). Such short-term oscillations in abundance suggest that more dramatic paleoclimatic fluctuations of the Pleistocene may have affected \textit{C. pallasii} populations in a profound way, so that genomic signatures are strong enough to detect the extent and direction of those population size changes.

Here, we analyse mitochondrial (mt) DNA sequences from nearly 1000 specimens of \textit{C. pallasii} sampled across the species’ vast range. Our dual goals are to obtain a comprehensive picture of the demographic and phylogeographic structure of this species, and thereby better understand how Pleistocene climate oscillations may have affected that population structure.

\textit{Materials and methods}

\textit{Sampling and sequencing}

\textit{Clupea pallasii} specimens were collected from spawning and prespawning aggregations at 20 localities across the species nearshore range in test fisheries or by research vessels during 1999–2010 (Table 1, Fig. 1). Tissue samples were preserved in either 95% ethanol or a saline solution (20% dimethyl sulfoxide, 0.25 M EDTA, saturated with NaCl, pH 8.0) (Seutin \textit{et al.} 1991). Genomic DNA was isolated from tissue samples. The first hypervariable segment, which was identical in length to the fragment amplified in the present study, may be difficult to identify using approaches that utilize only information on allele frequencies without knowledge about genetic relationships between alleles. Nonrecombining rapidly evolving mtDNA can offer additional insights into past isolations (e.g. in ice-age refugia) and secondary contacts between previously isolated populations (Avise 2000; Hewitt 2000). Furthermore, variation in nucleotide level is highly amenable for inferences about past demographic history of the species (Drummond \textit{et al.} 2005; Ho & Shapiro 2011). Demographic studies on shorter (decadal) time scales have shown that recruitment in \textit{C. pallasii} can vary sharply in response to SST fluctuations (Schweigert 1995; Nagasawa 2001). Such short-term oscillations in abundance suggest that more dramatic paleoclimatic fluctuations of the Pleistocene may have affected \textit{C. pallasii} populations in a profound way, so that genomic signatures are strong enough to detect the extent and direction of those population size changes.

Here, we analyse mitochondrial (mt) DNA sequences from nearly 1000 specimens of \textit{C. pallasii} sampled across the species’ vast range. Our dual goals are to obtain a comprehensive picture of the demographic and phylogeographic structure of this species, and thereby better understand how Pleistocene climate oscillations may have affected that population structure.
visualized by FigTree (ver.1.3.1; http://tree.bio.ed.ac.uk/software/figtree/). Genealogical relationships among haplotypes for both the control region and the cytochrome \( b \) gene were further assessed using a minimum spanning tree constructed by ARLEQUIN.

**Estimate of substitution rate**

A lineage-specific nucleotide substitution rate for herring was calculated with the equation \( \mu = dA / 2T \), where \( T \) is an estimate of 3.5 Myr divergence between \( C. pallasi \) and \( C. harengus \) (Grant 1986), and \( dA \) was the net average genetic distance between species estimated using 26 homologous \( C. harengus \) sequences. The net average genetic distance between lineages (\( dA = dXY - (dX + dY) / 2 \), where \( dXY \) is the mean distance between lineages \( X \) and \( Y \), and \( dX \) and \( dY \) are mean within lineage distances) was calculated using the Tamura and Nei model of nucleotide substitution for multiple hits with MEGA 4.1 (Tamura et al. 2007). Using this

### Table 1

<table>
<thead>
<tr>
<th>Sample</th>
<th>Abb.</th>
<th>Date of collection</th>
<th>All data</th>
<th>Lineage A</th>
<th>Lineage B</th>
<th>Lineage C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>( N )</td>
<td>( \mu )</td>
<td>( \pi )</td>
<td>( n (%) )</td>
</tr>
<tr>
<td>Aleksandrovsk*</td>
<td>AL</td>
<td>2008</td>
<td>50</td>
<td>0.974</td>
<td>0.011</td>
<td>50 (100.0)</td>
</tr>
<tr>
<td>Bering Sea</td>
<td>BS</td>
<td>2007</td>
<td>45</td>
<td>0.904</td>
<td>0.010</td>
<td>45 (100.0)</td>
</tr>
<tr>
<td>Gizhigin Bay*</td>
<td>GI</td>
<td>2005</td>
<td>52</td>
<td>0.952</td>
<td>0.012</td>
<td>52 (100.0)</td>
</tr>
<tr>
<td>Ilpir*</td>
<td>KK</td>
<td>2007</td>
<td>16</td>
<td>1.000</td>
<td>0.012</td>
<td>16 (100.0)</td>
</tr>
<tr>
<td>Krasnogorsk*</td>
<td>OZ</td>
<td>2008</td>
<td>48</td>
<td>0.935</td>
<td>0.010</td>
<td>48 (100.0)</td>
</tr>
<tr>
<td>Peter the Great Bay*</td>
<td>PV</td>
<td>2007</td>
<td>68</td>
<td>0.954</td>
<td>0.010</td>
<td>68 (100.0)</td>
</tr>
<tr>
<td>Port Moller</td>
<td>PM</td>
<td>2010</td>
<td>47</td>
<td>0.930</td>
<td>0.010</td>
<td>47 (100.0)</td>
</tr>
<tr>
<td>Taiu Bay*</td>
<td>AR</td>
<td>2006</td>
<td>52</td>
<td>0.936</td>
<td>0.012</td>
<td>52 (100.0)</td>
</tr>
<tr>
<td>Vznmorey*</td>
<td>VZ</td>
<td>2008</td>
<td>49</td>
<td>0.967</td>
<td>0.011</td>
<td>49 (100.0)</td>
</tr>
<tr>
<td>Columbia River</td>
<td>CR</td>
<td>2010</td>
<td>53</td>
<td>0.995</td>
<td>0.021</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Knight Inlet</td>
<td>KI</td>
<td>2003</td>
<td>46</td>
<td>0.989</td>
<td>0.020</td>
<td>—</td>
</tr>
<tr>
<td>Nootka Sound</td>
<td>NO</td>
<td>2001</td>
<td>46</td>
<td>0.973</td>
<td>0.020</td>
<td>—</td>
</tr>
<tr>
<td>North Sitka</td>
<td>NS</td>
<td>2000</td>
<td>42</td>
<td>0.988</td>
<td>0.019</td>
<td>2 (4.8)</td>
</tr>
<tr>
<td>Portage Inlet</td>
<td>PI</td>
<td>1999</td>
<td>44</td>
<td>0.814</td>
<td>0.016</td>
<td>—</td>
</tr>
<tr>
<td>Selwyn Inlet</td>
<td>SI</td>
<td>2001</td>
<td>43</td>
<td>0.987</td>
<td>0.020</td>
<td>4 (9.3)</td>
</tr>
<tr>
<td>San Francisco Bay</td>
<td>SF</td>
<td>2001</td>
<td>44</td>
<td>0.975</td>
<td>0.019</td>
<td>—</td>
</tr>
<tr>
<td>Simpson Sound</td>
<td>SP</td>
<td>2007</td>
<td>48</td>
<td>0.982</td>
<td>0.020</td>
<td>5 (10.4)</td>
</tr>
<tr>
<td>South Sitka</td>
<td>SS</td>
<td>2000</td>
<td>47</td>
<td>0.993</td>
<td>0.021</td>
<td>—</td>
</tr>
<tr>
<td>Tomales Bay</td>
<td>TB</td>
<td>2002</td>
<td>47</td>
<td>0.986</td>
<td>0.020</td>
<td>—</td>
</tr>
<tr>
<td>Yakutat Bay</td>
<td>YB</td>
<td>2008</td>
<td>48</td>
<td>0.990</td>
<td>0.020</td>
<td>3 (6.3)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>935</td>
<td>0.986</td>
<td>0.020</td>
<td>442</td>
<td>169</td>
</tr>
</tbody>
</table>

Abbreviation of populations (Abb.), date of collection, number of specimens (\( N \)), haplotype diversity (\( \mu \)) and nucleotide diversity (\( \pi \)) are shown for each population and for the entire data set. Number (\( n \)) and proportion of individuals (\( \% \)), and number of haplotypes (nh) and private haplotypes (np) for the phylogenetic lineages A, B and C in different populations are also shown.

*Sequences of these populations were retrieved from GenBank.

Fig. 1 Sample locations for Pacific herring. Populations are marked by abbreviations that correspond to Table 1. Pie charts show the frequencies of lineages A, B and C.
method, the lineage-specific divergence rate for control region was estimated at 1.59%/Myr. To address the issue of time dependency of molecular rate estimates (sensu Ho et al. (2005); see Discussion for more details), we additionally considered an alternative rate that is 10× faster. The 10-fold difference was chosen for two reasons: first, the corresponding higher divergence rate (15.9%/Myr) is near the high end of estimates reported for fishes (e.g. Bowen et al. 2006); second, Ho et al. (2008) reported up to 10× elevated rates on short time scales, so an order-of-magnitude range of rate estimates should bracket the most extreme plausible cases.

Population genetics

Population structure was evaluated with an analysis of molecular variance (AMOVA) that incorporates haplotype frequency and sequence divergence among haplotypes (Excoffier et al. 1992). We conducted AMOVA analyses with two groups: the Asian–Bering Sea group and the eastern North Pacific assemblage. The significance of covariance components associated with the different possible levels of genetic structure was tested using 5000 permutations. In addition, pairwise genetic divergence values between populations were estimated using the fixation index \( F_{ST} \) (Excoffier et al. 1992). The significance of the \( F_{ST} \) was evaluated by 10 000 permutations for each pairwise comparison. All of the genetic structure calculations were performed in ARLEQUIN. Because the GTR + I + G model is not implemented in ARLEQUIN, genetic distances between haplotypes were corrected for multiple hits using the Tamura and Nei model of nucleotide substitution with a gamma shape parameter (\( \alpha = 0.56 \)) correction for heterogeneity of mutation rates.

Demographic history

The \( F_S \) test of Fu (1997) was calculated for the control region data using ARLEQUIN. Fu’s \( F_S \) is very sensitive to population demographic expansions, which generally lead to large negative \( F_S \) values. Changes in effective population size (\( N_e \)) across time were inferred using Bayesian skyline analyses, which enable past demographic changes to be inferred from the current patterns of genetic diversity within a population (Drummond et al. 2005). BEAST v1.6 (Drummond & Rambaut 2007) was used to create the Bayesian skyline plots with 20 groups. Separate analyses were performed for each major mitochondrial lineage. To test for convergence, we performed multiple analyses that were run for \( 10^8 \) iterations with a burn-in of \( 10^7 \) under the GTR + I + G model, a strict molecular clock and a stepwise skyline model. Genealogies and model parameters were sampled every 10 000 iterations. All operators were optimized automatically. Results of replicate runs for each lineage were pooled using LogCombiner, and the effective sample size (ESS) for each parameter exceeded 200. Skyline plots were generated by Tracer1.4 (Rambaut & Drummond 2007). Historical demographic expansions were further investigated using the ‘mismatch distribution’ of all pairwise differences between mitochondrial sequences using ARLEQUIN. The time since population expansion was estimated using the equation \( \tau = 2ut \), where \( u \) is the mutation rate for the whole DNA sequence under study and \( t \) is the time since expansion.

Results

Summary statistics

A 557-bp sequence of the 5′ end of the control region was obtained for 935 specimens. Sequence comparison revealed 408 distinct haplotypes defined by 143 polymorphic sites with 129 transitions, 35 transversions and nine indels. Most of the haplotypes (309 or 76%) were singletons, being represented by a single sequence in the sample. Of the remaining 99 haplotypes, 85 were shared among populations and 14 were found in more than one individual but only in one population. A moderate heterogeneity in mutation rates among sites was indicated by an estimate of the gamma distribution shape parameter (\( \alpha = 0.56 \)). Overall values for haplotype (\( h \)) and nucleotide diversities (\( \pi \)) were 0.99 and 0.02, respectively.

A 668-bp fragment of the cytochrome \( b \) gene was obtained for a subset of 139 individuals from eight populations. A total of 62 substitutions (57 transitions and five transversions) at 59 polymorphic sites defined 66 haplotypes, 51 of which were singletons. Overall values for \( h \) and \( \pi \) were 0.96 and 0.008, respectively.

Phylogenetic relationships and three lineages

The phylogenetic tree obtained by Bayesian analysis under the GTR + I + G model using the complete data set of control region haplotypes revealed three distinct lineages with high posterior probabilities (labelled A, B and C in Fig. 2). This result was further supported by the minimum spanning trees for both the control region haplotypes and the cytochrome \( b \) haplotypes. At the control region, the three lineages each were separated from one another by six mutational steps (Figs 3–5), and the net genetic distances between lineages (±S.E) were as follows: A/B, 0.015 (±0.005); A/C, 0.014 (±0.005); and B/C, 0.014 (±0.004). Applying the lineage-specific control region divergence rate (1.59%/Myr), the evolutionary separations among these lineages occurred about 1 Ma, in the lower Pleistocene. Assuming the
A ‘faster’ rate of molecular divergence (15.9%/Myr), the lineages may have been separated for at least 100 000 years. At the cytochrome b gene, a minimum spanning tree of haplotypes also revealed the three lineages, which in this case were separated from one another by two or three mutational steps each (Fig. 6).

Overall, lineage A (169 haplotypes, 442 individuals) included all haplotypes that were found in the Asian and Bering Sea group. However, representatives of lineage A also were observed at low frequency in several populations in the eastern North Pacific (Table 1; Fig. 1). By contrast, lineages B (109 haplotypes, 210 individuals) and C (130 haplotypes, 283 individuals) were common and broadly sympatric in all populations of the eastern North Pacific. Although the frequencies of lineage C tended to be higher than those of lineage B in most of these populations, there were no obvious shifts in the frequencies of either of these two lineages along the eastern North Pacific coast (Table 1, Fig. 1). Within lineages A, B and C, mean pairwise genetic distances (±SE) between individuals were respectively as follows: 0.010 (±0.005), 0.014 (±0.007) and 0.013 (±0.007).

For each of the three major lineages, the topology of the minimum spanning tree was quite deep and characterized by multiple star-like polytomies with relatively shallow genetic divergences (Figs 3–5). Prevalent or common haplotypes usually were found in the centre of a given star-like polytomy. Within each of these three lineages, no distinct phylogeographic apportioning of localities was observed (Figs 3–5). The minimum spanning tree for cytochrome b haplotypes also was star shaped for each lineage (Fig. 6).

**Population structure**

The global *amova* analysis showed that 38.4% of the total genetic variation was apportioned between the Asian–Bering Sea group vs. the group in the eastern North Pacific (*P* < 0.01). A smaller (1.2%) but still significant (*P* < 0.01) amount of genetic diversity was present among populations within groups, and 60.4% of the total variation occurred within populations. Genetic subdivision was highly significant between groups (*Φ*<sub>CT</sub> = 0.38, *P* < 0.01) and among populations within groups (*Φ*<sub>SC</sub> = 0.02, *P* < 0.01).

The pairwise *Φ*<sub>ST</sub> values further revealed Krasnogorsk as being the most divergent population within the Asian–Bering Sea group (*Φ*<sub>ST</sub> = 0.03–0.09; *P* < 0.01) and Portage Inlet as being the most divergent population within the eastern North Pacific group (*Φ*<sub>ST</sub> = 0.08–0.16; *P* < 0.01). Indeed, an *amova* analysis that excluded the Krasnogorsk population revealed no significant structure in the remainder of the Asian–Bering Sea assemblage (*Φ*<sub>ST</sub> = 0.00; *P* = 0.5), and an *amova* analysis that excluded the Portage Inlet population revealed no significant genetic structure within the remainder of the eastern North Pacific assemblage (*Φ*<sub>ST</sub> = 0.00; *P* = 0.08).

**Population demography**

The *F*<sub>S</sub> tests for all the three lineages were negative and highly significant (Table 2), which might indicate population expansion. The mismatch distributions for all three major lineages (A, B and C) were generally unimodal and closely matched the expected distributions under a sudden expansion model (Table 2, Fig. 7).
Based on $\tau$ values and a divergence rate of 1.6% per Myr, the timing of expansion of all the three lineages dated to the middle Pleistocene (Table 2). If a 10$\times$ faster rate is applied, the timing of expansion would be in upper Pleistocene. Bayesian skyline plots revealed two episodes of Pleistocene demographic expansion in each
of the three lineages (Fig. 8). According to slower divergence rate of 1.59%/Myr, the first older demographic expansion began about 1.2–1.4 Ma, followed by a more recent abrupt demographic expansion that started 140–190 thousand years (kyr) ago. Furthermore, Bayesian skyline plot suggested one episode of population decline in lineage C that started 400 kyr ago, but this change was within 95% HPD intervals and thus not strongly supported. Use of the 10-fold faster mutation rate shortens the timing of these events by a factor of 10.
Challenges of molecular clock calibration

MtDNA substitution rates are known to vary considerably across animal lineages (Galtier et al. 2009), so lineage-specific mutation rates are necessary to obtain reliable timeframes for evaluating evolutionary hypotheses. Using the separation of Pacific from Atlantic herring populations that presumably attended the opening of the Bering Straight approximately 3.5 Ma (as evidenced by allozyme data; Grant 1986), the lineage-specific divergence rate for herring mitochondrial sequence was estimated at 1.59 %/Myr. Another study that used a fish-specific calibration for ITS1 arrived at similar time interval (Domanico et al. 1996) for this separation and thus confirmed that the ancestors of Pacific herring entered the Pacific Basin from Arctic regions during the Upper Pliocene rather than during a Pleistocene interglacial period.

Ho et al. (2005) recently argued that ‘internal’ molecular substitution rates estimated over short evolutionary timescales within a species can greatly exceed ‘external’ rates estimated over longer periods of time, e.g. between distant species. [The extent of such acceleration remains hotly contested and was shown by some authors to be either nonexistent (Emerson 2007) or greatly overstated (Burridge et al. 2008).] If the time-frame-dependent nature of rate calibrations is indeed as serious as suggested by Ho and colleagues, its repercussions on estimates of population-level events would be enormous. To account to some extent for such concerns, we have included two very different alternative rate calibrations in our current analyses. Thus, in addition to the external calibration based on the presumed split between Pacific and Atlantic herring (which we hereafter call the ‘slow’ or ‘phylogenetic’ rate), we also considered an internal rate that is 10-fold faster to introduce an

### Table 2 Summary of molecular diversity for Clupea pallasi control region and cytochrome b data

<table>
<thead>
<tr>
<th>Groups</th>
<th>( N )</th>
<th>No. of haplotype</th>
<th>( S )</th>
<th>( h )</th>
<th>( \pi )</th>
<th>( t )</th>
<th>( P )</th>
<th>( \tau )</th>
<th>( P(SDD) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>935</td>
<td>408</td>
<td>143</td>
<td>0.986</td>
<td>0.020</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Lineage A</td>
<td>442</td>
<td>169</td>
<td>93</td>
<td>0.955</td>
<td>0.011</td>
<td>—24.59</td>
<td>0.00</td>
<td>—6.52</td>
<td>0.01</td>
</tr>
<tr>
<td>Lineage B</td>
<td>210</td>
<td>109</td>
<td>76</td>
<td>0.968</td>
<td>0.014</td>
<td>—24.72</td>
<td>0.00</td>
<td>—6.52</td>
<td>0.01</td>
</tr>
<tr>
<td>Lineage C</td>
<td>283</td>
<td>130</td>
<td>84</td>
<td>0.970</td>
<td>0.014</td>
<td>—24.46</td>
<td>0.00</td>
<td>—6.52</td>
<td>0.01</td>
</tr>
<tr>
<td>Cyt b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>139</td>
<td>66</td>
<td>59</td>
<td>0.963</td>
<td>0.008</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Lineage A</td>
<td>50</td>
<td>15</td>
<td>18</td>
<td>0.811</td>
<td>0.003</td>
<td>—8.32</td>
<td>0.00</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Lineage B</td>
<td>35</td>
<td>18</td>
<td>27</td>
<td>0.889</td>
<td>0.007</td>
<td>—6.52</td>
<td>0.01</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Lineage C</td>
<td>54</td>
<td>33</td>
<td>35</td>
<td>0.961</td>
<td>0.006</td>
<td>—26.01</td>
<td>0.00</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Number of individuals (\( N \)), number of haplotypes, number of segregating sites (\( S \)), haplotype diversity (\( h \)), nucleotide diversity (\( \pi \)) and Fu’s \( F_S \) with corresponding \( P \) values for each grouping of samples are shown. Time since expansion measured in mutational time unit (\( \tau \)), absolute time (Ma) since expansion (\( t \)) and \( P \)-value for the sum of the squared deviations under the hypothesis of sudden expansion [\( P(SDD) \)] for control region data of each lineage are also shown.

**Discussion**

**Challenges of molecular clock calibration**

MtDNA substitution rates are known to vary considerably across animal lineages (Galtier et al. 2009), so lineage-specific mutation rates are necessary to obtain reliable timeframes for evaluating evolutionary hypotheses.
appropriate element of caution into our phylogeographic scenarios. Importantly, irrespective of the particular mutation rate chosen, the Pacific herring displays a general population stability or even growth throughout much of its history, suggesting that Pleistocene climatic changes did not have an overall negative impact on the historical demography of this species.

**Pleistocene lineage diversification**

Three major genealogical lineages and a distinct phylogeographic pattern were detected in our mtDNA analyses of *Clupea pallasii*. The geographical distribution of lineage A strongly implies an origin in the western North Pacific, whereas the spatial arrangement and broad sympatry of lineages B and C suggest an origin for the latter groups somewhere in the eastern North Pacific region. Based on the herring-specific divergence rates utilized in the current study, lineages A, B and C in the Pacific herring separated from one another approximately 1.0 Ma.

Divergence among the major herring lineages coincides with the onset of rapid climatic changes and oceanographic shifts that typified the mid-Pleistocene. By analysing siliceous microfauna assemblages in sediments, Morley & Dworetzky (1991) concluded that the first major cooling event in the North Pacific (comparable in severity to the late-Pleistocene glacial events) occurred about 1.3 Ma. A pronounced lowering of SST also was registered by changes in the diatom assemblage during this same general period of time (Morley & Dworetzky 1991). Thus, at that time, the distribution of *C. pallasii* probably shifted southward in response to the extreme glacial conditions in the subarctic Pacific, eventuating in a vicariant divergence between coastal populations in the western North Pacific (lineage A) vs. those in the eastern North Pacific region (lineages B and C). The glacial history of the North Pacific appears to have produced similar genetic subdivisions in several other marine and anadromous fish in this region (Grant et al. 1987, 2010; Seeb & Crane 1999; Beacham et al. 2009; Canino et al. 2010; Ravago-Gotanco & Juinio-Meñez 2010), although the magnitude of genetic subdivision varies among species. Under this phylogeographic scenario for Pacific herring, spatial overlap in the current distribution of eastern and western herring lineages in the North Pacific (such as we observed at sites SI and SP) would be a result of secondary contact sometime after the range of *C. pallasii* shifted northward during interglacial periods of global warming. Considering that there is no genetic divergence between the lineage A haplotypes in the eastern North Pacific group and those from the Asian–Bering Sea group, secondary contact in these cases probably occurred after the last glacial maximum (LGM).

The divergence between lineage B and lineage C also coincides with the first major cooling in the North

---

**Fig. 8** Bayesian skyline plots showing changes through time of $N_e T$ ($N_e$ = effective population size; $T$ = generation time) inferred with BEAST based on mtDNA control region sequences of three lineages shown with each pane. Shown at the bottom of the figure on a separate axis are times under an alternative (10x faster) calibration rate for mtDNA. Black lines represent median estimates of $N_e T$; light lines are the upper and lower 95% highest posterior density (HPD) limits of $N_e T$; and the vertical dotted line represents the median estimate of time to most recent common ancestor (TMRCA). The upper 95% HPD on TMRCA is at the right end of the plot, whereas the lower 95% HPD is the dashed line to the left of the median. The $y$-axis is plotted on a logarithmic scale.
Pacific, at which time the SST off Southern California and Baja California were more than 4 °C cooler during the winter than they are today (Moore et al. 1980). One likely consequence is that these temperature shifts probably drove the range of C. pallasii southward towards the Baja California coast. Although C. pallasii is currently rare in Southern California, fishery records indicate that it resided along the northern Baja California coast and spawned in San Diego Bay in recent historical times (McHugh 1954; Miller & Lea 1972; Spratt 1981). Another relevant point about the physical environment is that a seaway across the middle of the Baja Peninsula existed approximately 1.0–1.6 Ma, thus providing a direct connection between the middle of the Gulf of California and the Pacific Ocean (Upton & Murphy 1997; Riddle et al. 2000). Thus, C. pallasii might have invaded the Gulf of California through this mid-Peninsula channel and then become vicariously isolated from Pacific Ocean populations after the seaway closed approximately 1 Ma. Mid-Pleistocene allopatric divergences following the closure of the mid-Peninsula seaway have been detected in several other Baja disjunct fish species (Present 1987; Terry et al. 2000; Huang & Bernardi 2001; Bernardi et al. 2003).

Whether the third lineage in C. pallasii arose in the Gulf of California (as opposed to some other location) remains uncertain, but if it did, the Gulf lineage later could have migrated out of the Gulf (perhaps via the southern tip of the Baja Peninsula) during periods of oceanic cooling associated with glaciation events, thereby regaining contact and mixing with its Pacific Ocean counterpart, with the net result being the current sympatric distribution of lineages B and C in the north-eastern Pacific region. It is impressive that today there are no obvious shifts in the frequency of either herring lineage along the vast stretch of eastern Pacific coastline from central California to Alaska. This indicates, perhaps, that the ranges of the two lineages formerly were much narrower than they are now, ensuring thorough mixing prior to the range expansion. For marine species with high dispersal ability, an extensive mixture of diverged lineages is known to be possible. For example, Gaither et al. (2010) monitored the fate of two allopatric lineages of Bluestriped Snapper (Lutjanus kasmira) introduced to Hawaii a half-century ago and found that both lineages are now present in similar frequency (and at about the same proportion as in the original introduction) at all surveyed sites across a 2000-km stretch of the Hawaiian Archipelago. These observations confirm that the ranges of marine species can in some cases expand quickly across long distances.

Another (but in our opinion perhaps less likely) scenario is that origin of herring lineages B and C did not involve geographical isolation of any sort but instead registers a deep coalescent event in one large and historically well-mixed population. In principle, about 2N generations are required for neutral haplotypes in a panmictic population to coalesce, where N is the female effective population size (Avise 2004). Thus, if a population of Pacific herring remained stable or grew throughout much of its history, particular matrilines theoretically could coexist for millions of generations and thereby accumulate numerous sequence differences. Although the widespread primary sympatry of deeply separated mtDNA lineages has only rarely been reported in any species (Avise 2000), we certainly acknowledge its theoretical possibility in Pacific herring, especially given that our demographic analyses strongly imply that herring populations have either remained stable or have grown in size at particular times in their recent history.

Population structure

The AMOVA shows pronounced genetic structure between the Asian–Bering Sea matrilineal group and the north-eastern Pacific group of herring, a finding generally consistent with the allozyme survey by Grant & Utter (1984). Compared with the allozyme markers, however, the rapidly evolving and maternally inherited nonrecombining mtDNA molecule should provide a better opportunity to uncover interactions between these two genetic assemblages. The Alaska Coastal Current enters the south-eastern Bering Sea via Aleutian Island channels that should facilitate dispersal of C. pallasii from the Gulf of Alaska to the Bering Sea (Mann & Hamilton 1995). Thus, it may seem surprising that we did not detect herring lineages B or C in the three Bering Sea populations surveyed. However, this peculiarity may reflect past conditions more so than present conditions. Beringia was not covered by ice during the glacial periods, so the possibility of faunal exchange consistently existed between the Bering Sea and Asian coastal waters to the south. In contrast, during glacial periods, the Bering Sea was separated from the Gulf of Alaska by the Aleutian Islands that were covered with ice (and the Gulf of Alaska’s coastal itself was also covered by ice). Perhaps herring populated the Bering Sea during the last glacial maxima, but even if not the Asian race could have arrived there in the early stages of an interglacial while south-eastern Alaska and British Columbia were still under ice and thus inaccessible for herring. In any event, after the Asian race reached high population density in Bering Sea, it might have been difficult for other lineages to establish there. As the ice sheet blanketing Alaska and British Columbia started receding, the coast became available for colonization. Judging by current distribution of lineages A–C,
most herring colonists in this region came from the American continental coastline to the south, but some fish might also have come from the then-populated Bering Sea through newly opened passes in eastern Alaska or perhaps directly from Kamchatka along the Aleutian Islands. At present, herring abundance is greater in the Bering Sea along south-western Alaska than in the Gulf of Alaska (Hay et al. 2008); therefore, with similar migration rates, one might expect gene flow across the Aleutian Island channels to be greater southward than in the opposite direction. Similar patterns of dispersal are suspected for at least two other fish species in this region (Hawkins et al. 2005; Orr & Hawkins 2008). Regardless of the specific colonization route(s) utilized by herring, the main conclusion is that contemporary ocean currents alone are not necessarily a reliable indicator of gene flow among marine populations.

Within the major herring lineages, the pairwise \( \Phi_{ST} \) values identified the Krasnogorsk and Portage Inlet sites as home to the most divergent populations within the Asian–Bering Sea and the eastern North Pacific assemblages, respectively. The genetic distinctiveness of these two populations also is supported by the haplotype compositions within each. Private haplotypes in high frequency and the overrepresented and underrepresented common haplotypes observed in both populations suggest that their histories may have included some rare stochastic population events such as reproductive sweepstakes (Hedgcock 1994). Indeed, strong interyear variation in reproductive recruitment is known for some \( C. pallasii \) populations, and such variation could promote genetic heterogeneity across space and time (Schweigert 1995; Nagasawa 2001), as also has been evidenced at regional scales by microsatellite markers (O’Connell et al. 1998; Small et al. 2005; Beacham et al. 2008; Sugaya et al. 2008). The distinctive nature of the Portage Inlet population as registered in mtDNA is also supported by an extensive analysis of microsatellite markers (Beacham et al. 2008). The particular nuclear and cytoplasmic genotypes observed in this and in the Krasnogorsk population provide special opportunities to study the fidelity (or lack thereof) of individuals to specific spawning sites and to assess the rates of dispersal and gene flow in these particular herring populations.

**Consequence of Pleistocene glaciations for demography**

Several of our statistical analyses of mtDNA haplotypes in \( C. pallasii \) yielded signatures of major demographic expansions likely associated with large-scale ecological shifts during the Pleistocene. During glacial periods, temperate coastlines in the North Pacific housed extensive cold, rocky shore ecosystems (Lyle et al. 2001; Graham et al. 2003) with extensive stands of macroalgae (kelp) that provided productive and complex environments for many marine species (Paine 2002). Increased upwelling (Lyle et al. 2010) would lead to higher plankton production and more food for \( C. pallasii \). Considering the environmental conditions favoured by Pacific herring (Alderdice & Velsen 1971; Haegele & Schweigert 1985), the cold and productive kelp-based ecosystems of the Pleistocene glacial periods likely provided ideal habitats for \( C. pallasii \) and therefore may have contributed to the demographic expansions that we detected in our genetic analyses. In response to less favourable interglacial warming, \( C. pallasii \) then may have shifted its range northward and found optimal spawning habitats in areas that were affected by coastal glaciers. Studies indicate that recruitment of \( C. pallasii \) is strongly influenced by annual mean SST even over decadal timescales, with higher spawning outputs at lower temperatures (Schweigert 1995; Nagasawa 2001). Indeed, a northerly shift in the distribution of \( C. pallasii \) in California was recorded in response to a warm El Nino off California (Spratt 1984), thus showing that even mild changes in oceanographic conditions can affect herring distributions. Such range shifts in the past might have counteracted any otherwise negative effects of altered environmental conditions on \( C. pallasii \).

Any assumption of faster evolutionary rates for mtDNA would of course shorten the timescales of all of our scenarios commensurately. For example, if we apply a 10× faster clock calibration (15.9%/Myr), then the initial population expansion of Pacific herring occurred 125–140 kya and the most recent expansion would have begun about 14–19 kya (a date that coincides with the onset of the current interglacial period). Such applications of alternative molecular rates illustrate how different yet plausible clock calibrations can drastically change conclusions about the timings of particular historical demographic events.

**Concluding remarks**

Our results indicate that climatic fluctuations of the Pleistocene left decipherable genetic footprints on several phylogeographic and demographic facets of herring history. With respect to phylogeography, a deep division between mtDNA lineages in the western vs. eastern North Pacific probably registers vicariant population separations and subsequent range shifts associated with Pleistocene glacial cooling. With respect to demography, our coalescent analyses revealed that populations of \( C. pallasii \) experienced dramatic population expansions at particular times during the Pleistocene and also that the Pleistocene glacial–interglacial cycles were not detrimental to herring population sizes.
overall. Therefore, we conclude that climatic oscillations in Pleistocene glacial–interglacial cycles had no major adverse effects in the demographic history of C. pallasii. This demographic outcome contrasts with a common situation in the terrestrial realm where Pleistocene oscillations often had negative impacts on population size in many animal and plant species (Hewitt 2000; Shapiro et al. 2004; Stiller et al. 2010). Because the demographic consequences of Pleistocene perturbations can in some cases be quite different for marine and terrestrial species, it would be interesting to deduce demographic histories for many other vagile marine species that had the potential to shift their ranges in response to severe climatic shifts.

Acknowledgements

We thank the following who kindly provided samples: T. Hartill, R. Larson, M. Pritchett, and K. Hebert from the Alaska Department of Fish and Game, J. Schweigert from Fisheries and Oceans Canada, Y. Carolsfeld from the World Fisheries Trust, and M. Litz, R.L. Emmett, and J.J. Vollenweider from the NOAA Fisheries. Our genetic work was funded by University of California at Irvine.

References


CLIMAP (1981) Seasonal Reconstruction of the Earth’s Surface at the Last Glacial Maximum. The Geological Society of America Map and Chart Series No. 36, Boulder, CO.


