Journal of Heredity 2006:97(6):571–580
doi:10.1093/jhered/esl033
Advance Access publication October 12, 2006

Biogeographic Evidence for Selection on Mitochondrial DNA in North Pacific Walleye Pollock *Theragra chalcogramma*

W. Stewart Grant, Ingrid B. Spies, and Michael F. Canino

From the Department of Biological Sciences, University of Alaska Anchorage, Anchorage, AK 99508 (Grant); and Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115 (Spies and Canino).

Address correspondence to W. Stewart Grant at the address above, or e-mail: phylogeo@yahoo.com.

Abstract

Three major mitochondrial DNA (mtDNA) haplogroups were identified in 5 data sets for North Pacific and Bering Sea walleye pollock. The common haplogroup A showed mirror-image clines on both sides of the North Pacific with high frequencies in southern areas ($P_A > 0.84$) and low frequencies in the Bering Sea ($P_A < 0.36$). Two additional haplogroups showed complimentary, but weaker, clines in the opposite direction. These clines are unlikely to have arisen by chance during postglacial colonizations of coastal waters in the North Pacific and Bering Sea, and they do not appear to reflect isolation by distance. Contrary to these trends, pollock at the western end of the Aleutian Island Archipelago were genetically more similar to Asian than to North American pollock, a pattern likely reflecting postglacial colonization. Haplogroup F_{ST} values for a given haplotype diversity were significantly larger than expected under the island model of migration and random drift, a result implicating natural selection. Frequencies of haplogroup A were highly correlated with sea surface temperature ($r > 0.91$), whereas frequencies of groups B and C showed negative correlations with temperature. Selection may be operating directly on mtDNA variability or may be mediated through cytonuclear interactions. This biogeographic evidence adds to a growing body of literature indicating that selection may play a greater role in sculpting mtDNA variability than previously thought.

In phylogeographic studies, mitochondrial DNA (mtDNA) variability among populations is usually interpreted to reflect the effects of gene flow and random drift. Gene flow between populations of marine fishes may be particularly important because of the large potential for movement between populations as larvae, juveniles, or adults. Although natural selection has been documented for some nuclear genes in marine fishes (Pogson 2001; Canino and Bentzen 2004), selection on mtDNA is generally assumed to be minimal (Avise 2000). However, several studies have demonstrated selection on mtDNA in other groups, including mammals (e.g., Ruiz-Pesini et al. 2004), insects (e.g., Rand and Kann 1996), and freshwater fishes (e.g., Moyer et al. 2005). A recent summary of genetic diversity in populations of vertebrates and invertebrates failed to confirm the overriding importance of demographic factors in sculpting mtDNA variability (Bazin et al. 2006).

Here, we examine mtDNA variability among populations of walleye pollock (hereafter pollock), *Theragra chalcogramma*, which extend around the rim of the North Pacific from the Sea of Japan to Puget Sound, Washington. Biological (e.g., Serobaba 1978; Hinckley 1987; FitzGerald et al. 2004) and genetic (e.g., Olsen et al. 2002) differences appear among populations despite the large potential for gene flow during planktonic egg and larval stages (Hinckley et al. 1991). Genetic differences among populations have been detected on scales of hundreds of kilometers with allozyme (Grant and Utter 1980; Olsen et al. 2002), mtDNA (Mulligan et al. 1992; Shields and Gust 1995; Kim et al. 2000; Olsen et al. 2002; Brykov et al. 2004), microsatellite DNA (Olsen et al. 2002; O'Reilly et al. 2004; Shubina et al. 2004), and nuclear (pantophysin [*Pan* I] Canino et al. 2005) gene markers. On a larger geographic scale, genetic markers reveal 2 parapatric groups across the North Pacific meeting in the Bering Sea (Iwata 1975; Olsen et al. 2002; O'Reilly et al. 2004). This subdivision has been variously attributed to ice-age vicariances (e.g., Grant and Utter 1984) or natural selection (e.g., Canino et al. 2005).

Whereas several studies have used mtDNA polymorphisms to survey regional populations of pollock, no range-wide studies have been conducted to define ocean-scale features of genetic variability. Here, we summarize

4 published mtDNA data sets (Mulligan et al. 1992; Kim et al. 2000; Olsen et al. 2002; Brykov et al. 2004) and, together with new mtDNA sequence data from Puget Sound, resolve ocean-wide patterns. The positions and abundances of haplotypes in genealogies from the various studies allow the identification of 3 major haplogroups in the data sets. Frequencies of these groups vary clinally along each side of the North Pacific. We evaluate 3 mechanisms potentially producing these clines: 1) random events during postglacial colonization, 2) isolation by distance, and 3) natural selection. Pollock inhabit coastal areas with strong latitudinal sea surface temperature (SST) gradients, and these mtDNA clines may reflect the effects of temperature-mediated selection.

Materials and Methods

Fifty juvenile walleye pollock were collected in June 2003 in Puget Sound near Port Townsend, Washington (USA), and tissue samples were preserved in 95% ethanol. Genomic DNA was extracted with a Qiagen DNeasy kit (Qiagen Inc., Valencia, CA). Polymerase chain reaction (PCR) primers, designed with Primer3 software (Rozen and Skaletsky 2000), amplified a 479-bp fragment of the cytochrome oxidase subunit I (COI) mitochondrial gene. Amplifications were conducted in 10 μl volumes containing ~100 ng template DNA, 10 mM Tris–HCl (pH 8.3), 50 μM KCl, 2.5 mM $MgCl_2$, 2.0 mM of each deoxyribonucleotide triphosphate, 5 pM of each primer (COI_gadid_F—5′ATCACCCGCTGATTTTTCTC3′ and COI_gadid_R—5′ATTGCCCCAAGAATTGATGA3′), and 0.5 U Biolase *Taq* polymerase (Biolase Technology, Inc., San Clemente, CA). A "touchdown" thermal cycle was performed in an MJ Research PTC-100 Thermalcycler (MJ Research, Inc., Waltham, MA). An initial denaturation step at 94 °C (2 min) was followed by 5 cycles of 94 °C (1 min), 56 °C (1 min [−1° per cycle]), and 72 °C (1 min), then by 25 cycles of 94 °C (30 s), 51 °C (30 s), and 72 °C (30 s), with a final extension at 72 °C (10 min). Amplicons were treated with ExoSAP-IT® (USB Corp., Cleveland, OH) to degrade unincorporated primers and deoxynucleotides.

Cycle sequencing was performed directly from PCR products on a LI-COR 4300S DNA analyzer (LI-COR Biosciences, Lincoln, NE) by combining 100 mM purified PCR product with labeled forward (IRD800, LI-COR Biosciences) and reverse (IRD700, LI-COR Biosciences) primers (ThermoSequenase II, Amersham Biosciences, Piscataway, NJ). Forward and reverse sequences were analyzed using e-SEQ 3.0.9 (LI-COR Biosciences), and sequence contigs were aligned with SEQUENCHER 4.0.5 (Gene Codes Corp., Ann Arbor, MI). Haplotype sequences were deposited in GenBank (DQ385429–DQ385443).

The data sets of Brykov et al. (2004), Kim et al. (2000), and Mulligan et al. (1992) consist of restriction fragment length polymorphism (RFLP) profiles of whole mtDNA or PCR amplicons (Table 1). Mulligan et al. (1992) produced whole mtDNA restriction profiles for 4 samples from the West Gulf of Alaska and Bering Sea (Figure 1) using 9 enzymes. Two errors appear in this study: restriction haplotype "F" for *Apa*I appeared in a table of haplotype frequencies but not in the table of restriction enzymes fragment profiles. Also restriction fragment profile "F" appears for *Bgl*I in the table of restriction profiles but not in the table of haplotype frequencies. These 2 enzymes were dropped from the analysis without a loss of phylogeographic resolution. Kim et al. (2000) used 8 restriction enzymes to survey whole mtDNA in 5 samples from the Northeastern Pacific and Bering Sea, and Brykov et al. (2004) used 6 restriction enzymes to survey 3 PCR fragments (ND3/4, ND5/6, Cyt *b*/CR) in 5 samples from the West Bering Sea. Olsen et al. (2002) surveyed 3 PCR fragments (ND5/6, Cyt *b*, COI) with 8 restriction enzymes in 6 spawning-area samples, extending from the Kamchatka Peninsula to the central Gulf of Alaska. A fifth data set from Shields and Gust (1995) of 2 noncoding regions (76 bp spacer region, 170 bp left side of the control region) did not provide sufficient phylogeographic resolution to be useful. For each RFLP data set, restriction profiles were concatenated to create a composite haplotype.

Estimates of haplotype (h) and nucleotide (Θ_π and Θ_W) diversity for individual samples followed Nei (1987), as implemented in ARLEQUIN 2.0 (Schneider et al. 2000). Watterson's (1978) homozygosity test, based on F, the sum of squared haplotype frequencies, tests the fit of the observed to the expected haplotype frequency spectrum under the infinite-alleles model for a population in mutation-drift stasis (Ewens 1972). Chakraborty's (1990) test, also based on Ewens' frequency spectrum, estimates the expected number of haplotypes under neutrality. Significances for both tests were evaluated with 1000 permutations. F_S (Fu 1997), based on the infinite-sites model, was also used to test for neutrality. Diversity statistics on both sides of the North Pacific were regressed with distance from the southernmost sample using Numerical Taxonomy and Multivariate Analysis System, (NTSYS, Exeter Software, Setauke, NY).

Gene genealogies for each data set were inferred from minimum spanning networks (ARLEQUIN). Restriction site polymorphisms on different basal branches in the genealogy were assumed to reflect different nucleotide substitutions in the restriction recognition site. Major haplogroups in each data set were identified with procedures outlined in Templeton et al. (1987). Three groups in common among the studies were identified by their abundances and positions (tip or interior) in the genealogies of pooled samples for each study. A weighted average of group frequencies was estimated for 14 areas. These frequencies were used to estimate F_{ST}, which was tested for significance with 1000 permutations of sample frequencies (Analysis of Molecular Variance, ARLEQUIN).

The distribution of F_{ST} and heterozygosity values among 24 allozymes and 3 microsatellite DNA loci (Olsen et al. 2002) was used to test the hypothesis that the F_{ST}'s and gene diversities of the 3 haplogroups fit the infinite-alleles and island-migration models of population structure (FDIST2, Beaumont and Nichols 1996). The 95% limits were generated with 50 000 iterations of 100 hypothetical populations with samples sizes of 100 individuals.

Table 1. Summary statistics of mtDNA variability and tests of neutrality in walleye pollock

Locality	*N*	No. of haplotypes		Homozygosity		F_S
		Observed	Expected	Observed	Expected	
Kim et al. (2000)						
1. Sea of Japan	30	7*	3.8	0.471	0.287	−3.46**
2. Pacific Hokkaido	20	9	6.3	0.260*	0.179	−5.34***
3. Kurile Islands	10	2	2.7	0.520	0.681	1.03
7. Central Bering Sea	20	9	7.7	0.205	0.179	−3.85**
9. Bogoslof Island	20	8	7.0	0.230	0.209	−3.38*
Pooled	100	20***	8.0	0.297***	0.125	−17.05***
Brykov et al. (2004)						
4. Ozernoi Bay	34	15	11.9	0.147	0.115	−5.11*
4. Olutovskyi Bay	40	19	19.9	0.080	0.088	−8.98***
5. Koryak Shelf	40	18	18.1	0.091	0.097	−7.28**
5. Navarin	40	21	16.6	0.103	0.076	−10.17***
6. Anadyr	38	16	13.7	0.129	0.111	−6.59**
Pooled	192	48***	32.4	0.080***	0.051	−25.88***
Shields and Gust (1995)						
4. W. Bering Sea	18	7	4.1	0.401*	0.234	−3.54**
8. W. Aleutians	10	2	1.5	0.820***	0.678	−0.34
9. E. Aleutians	38	12***	4.2	0.446***	0.167	−8.42***
10. S.E. Bering	14	9	7.6	0.184*	0.149	−6.08***
Pooled	80	21***	5.5	0.401***	0.109	−21.99***
Mulligan et al. (1992)						
7. Central Bering Sea	50	16**	9.1	0.224***	0.126	−6.49***
8. Adak Island	40	24	19.9	0.080**	0.062	−14.09***
9. Bogoslof Island	42	9	8.8	0.222	0.236	−1.99
11. Shelikof Strait	36	10	8.6	0.218	0.200	−1.48
Pooled	168	43***	17.8	0.150***	0.056	−26.22***
Olsen et al. (2002)						
3. E. Kamchatka	70	19*	11.9	0.184	0.116	−11.67***
9. Bogoslof Island	174	36	21.2	0.125	0.072	−26.97***
10. Unimak Pass	30	12	7.9	0.227*	0.146	−6.20***
11. Shelikof Strait	171	37***	15.8	0.171***	0.069	−26.88***
12. Prince Wm Snd	164	27***	12.4	0.215***	0.103	−15.13***
13. Middleton Island	75	19***	8.6	0.262*	0.120	−11.13***
Pooled	684	66***	21.0	0.170***	0.054	−26.23***
This study						
14. Puget Sound	50	15***	5.5	0.318***	0.137	−28.27***

Sample numbers correspond to locations in Figure 1.

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Regression analysis NTSYS, was used to test for correlations between group frequencies and 1) distances between samples and 2) historical SSTs (Reynolds 1988). SST for each sampling site was interpolated from isotherms on maps produced by the COADS database (1800–2002, Comprehensive Ocean Atmosphere Data Set; Woodruff et al. 1993).

Results

Thirteen transitions and 2 transversions along a 479-bp fragment of COI produced 15 haplotypes in a sample of 50 pollock from Puget Sound (Table 2). No indels were observed, and 11 haplotypes occurred as singletons. The number of haplotypes and the level of haplotype "homozygosity" significantly exceeded the expectations of the drift-mutation neutral model (Table 1).

Weighted averages of haplotype diversities for 14 areas appear in Figure 2. Among Asian samples, *h* increased from 0.55 in the Sea of Japan to 0.93 in the Northwestern Bering Sea and was highly correlated with distance from the Sea of Japan ($r = 0.926$, $P = 0.008$). Θ_π rose from 0.3% to 2.4% and was highly correlated with distance ($r = 0.927$, $P = 0.008$). Θ_W increased from 0.6% to 3.7% and was also highly correlated with distance ($r = 0.941$, $P = 0.005$). Diversity gradients were less steep, but also significant, in the Northeastern Pacific. Haplotype diversity varied from 0.70 in Puget Sound to 0.95 in the Southeastern Bering Sea and was correlated with distance from Puget Sound

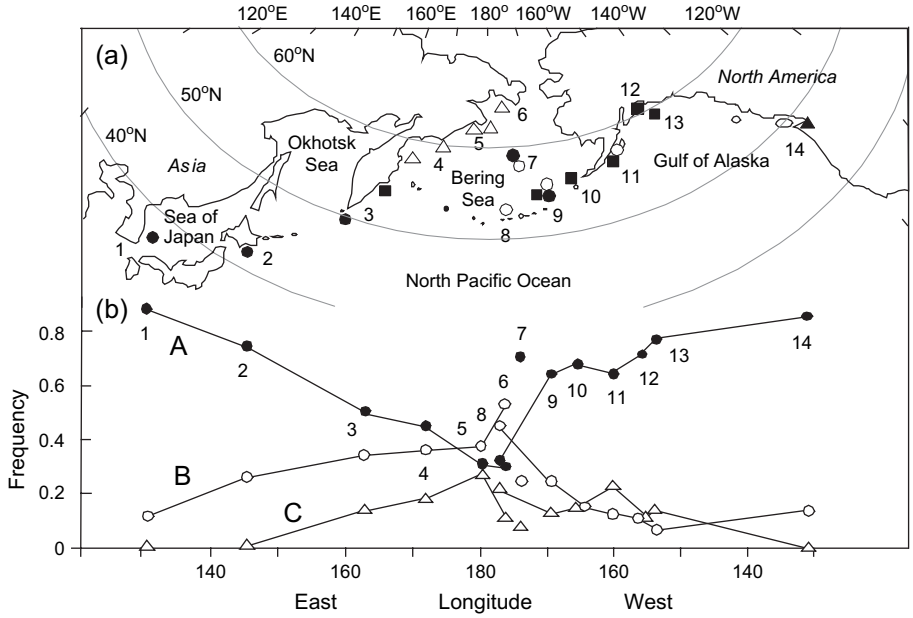

Figure 1. (a) Map showing locations of pollock samples in Mulligan et al. (1992) (open circles), Kim et al. (2000) (closed circles), Olsen et al. (2002) (closed squares), and Brykov et al. (2004) (open triangles). (b) Frequencies of mtDNA clades **A** (closed circles), **B** (open circles), and **C** (open triangles). Collection localities: 1, West Sea of Japan; 2, Pacific Hokkaido; 3, Kurile Islands; 4, Ozernoi and Olutovskyi bays; 5, Cape Navarin; 6, Gulf of Anadyr; 7, Central Bering Sea ("Donut" hole); 8, West Aleutian Islands; 9, East Aleutian Islands; 10, Southeastern Bering Sea; 11, Shelikof Strait; 12, Prince William Sound; 13, Middleton Island; and 14, Puget Sound.

($r = 0.855$, $P = 0.014$). Θ_π varied from 0.3% to 1.2% and was also correlated with distance ($r = 0.828$, $P = 0.022$). Θ_W was also correlated with distance ($r = 0.791$, $P = 0.034$).

Tests of drift-mutation neutrality were made on the original samples (Table 1). Significant values of F_S appeared in all but 4 samples. These departures from neutrality generally reflect greater numbers of haplotypes and higher levels of haplotype homozygosity than expected from Ewens' distribution. The number of significant departures from neutrality for both variables tended to be larger in North American than in Asian samples.

Table 2. Haplotype frequencies of COI in pollock in Puget Sound, Washington

Haplotype	GenBank accession no.	Frequency
1	DQ385429	3
2	DQ385430	3
3	DQ385431	1
4	DQ385432	1
2	DQ385433	6
6	DQ385434	1
7	DQ385435	1
8	DQ385436	27
9	DQ385437	1
10	DQ385438	1
11	DQ385439	1
12	DQ385440	1
13	DQ385441	1
14	DQ385442	1
15	DQ385443	1

Although summary statistics for individual samples reveal distinctive geographic patterns, trends in haplotype frequencies among areas across the North Pacific remain obscure. In an attempt to combine the various data sets, homologous haplogroups among studies were identified by their positions in haplotype networks and by frequencies in geographically overlapping samples among individual data sets (Figure 3). Three major groups, **A**, **B**, and **C**, appeared in the data sets of Mulligan et al. (1992), Olsen et al. (2002), and Brykov et al. (2004). Two major groups appeared in the data sets of Kim et al. (2000) and in the sequences for Puget Sound. Group **C** appeared to be absent in the southernmost samples on both sides of the North Pacific.

Frequencies of these haplogroups varied considerably across the North Pacific (Figure 1). The frequency of **A** along the Asian coast decreased clinally from 0.87 in the Sea of Japan to 0.34 in the northern Bering Sea. Along North America, frequencies of **A** varied clinally from 0.84 in Puget Sound to 0.64 around Bogoslof Island. Frequencies of groups **B** and **C** also varied clinally, but in the opposite direction from **A**. Samples 7 and 8 in the center of the species' range, however, did not fit into these clinal patterns. In a three-dimensional ordination of samples by group frequencies, the sample from the western end of the Aleutians (No. 8) clustered with Russian samples (Nos 4–6), and samples from the central Bering (No. 7) in international waters (the "donut hole") clustered with Alaskan samples (Nos 9–13) (Figure 4). Excluding these 2 samples,

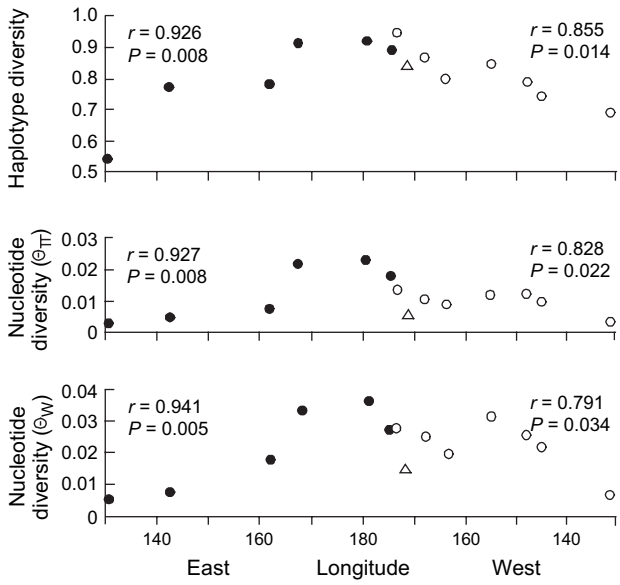

Figure 2. Haplotype diversity (h) and nucleotide diversity (Θ_π and Θ_W) averaged among data sets for 14 areas. Closed circles represent Asian samples and open circles represent North America samples. Open triangles represent samples near Adak Island in the western Aleutian Island Archipelago.

clines were steeper among Asian than among North American populations. Levels of differentiation among Asian populations (F_{ST} = 0.053–0.059) were two to three times higher than among North American populations (F_{ST} = 0.015–0.022) (Table 3).

Unexpectedly high values of F_{ST} can potentially identify genetic markers under natural selection (Baer 1999). Most allozyme and microsatellite DNA loci from Olsen et al. (2002) show low levels of differentiation among populations ($F_{ST} < 0.009$) for a given level of heterozygosity, except for two allozyme loci, *Sod-2* and *Mpi* (Figure 5). Both these enzymes are suspected to be under selection (Iwata 1975; Olsen et al. 2002). F_{ST} values for the 3 mtDNA haplogroups on both sides of the North Pacific exceeded the 95% limit of F_{ST} expected under the island-migration and infinite-alleles models of population structure.

Regression analysis revealed a significant decrease in the frequencies of group **A** with distance from the southernmost sample on each side of the North Pacific ($r > -0.90$, Table 3). Frequencies of **A** were significantly correlated with SST ($r > 0.90$). On the Asian side, significant correlations appeared between frequencies of **B** and distance ($r = 0.946$) and between **B** and SST ($r = -0.920$), but not on the North American side of the North Pacific. Significant positive correlations ($r > 0.83$) appeared between the frequency of **C** and distance, and significant negative correlations ($r > -0.81$) appeared between **C** and SST on both sides of the North Pacific.

Discussion

Three mechanisms, possibly in combination, may account for these patterns of mtDNA variability in pollock. First, pollock populations may bear imprints of postglacial colonizations (Hewitt 1996). Under this "historical" hypothesis, populations have not reached migration–drift equilibrium across the North Pacific and may still reflect demographic effects of historical colonizations. Under a second hypothesis, restricted gene flow and random drift have produced isolation by distance among populations, resulting in haplogroup frequency clines. Although these historical and equilibrium hypotheses may explain some aspects of the results, the clinal patterns in haplotype frequencies are more likely due to natural selection.

Before discussing these hypotheses in more detail, we offer 2 prominent caveats:

1. The apparent geographic patterns observed in this survey of mtDNA data depend on correctly identifying homologous haplogroups among studies. Identifications are possible because the lack of recombination along mtDNA molecule preserves linkage groups. Although different combinations of restriction enzymes were used in the different studies, each study detected the same underlying molecules and, hence, should detect the same mitogenomic haplotypes. The large data set of Olsen et al. (2002) provides the backbone for these identifications because of wide geographic coverage, large sample sizes, and geographic overlap with samples from other studies.
2. Geographic population structure may not be fully resolved by surveys of genetic markers in some studies of pollock because samples were collected out of the spawning season, except for samples in Olsen et al. (2002) and the sample from Puget Sound. However, significant haplotype frequency shifts between adjacent samples of nonspawning fish indicate that movement may be largely limited to in- and offshore seasonal migrations. The lack of significance of Chakraborty's amalgamation test and the general lack of departures from Hardy–Weinberg proportions for allozyme loci (Grant and Utter 1980) in most areas also indicate the lack of population mixing out of the spawning season.

Postglacial Colonization

During glacial maxima, climate cooling and drops in sea level modified shorelines, ocean currents, and plankton productivity in the central North Pacific (Hopkins 1972; Sancetta 1983; Oba et al. 1991). During these maxima, perennial sea ice covered much reduced areas of the Okhotsk and Bering seas (Mann and Peteet 1994). Inshore areas were also covered by tidewater glaciers, which extended along Kamchatka Peninsula and from the Aleutian Islands to Puget Sound (Molnia 1986; Mann and Peteet 1994; Barrie and Conway 1999). These sea-level glaciers disrupted inshore pollock nursery areas through combinations of physical displacements, temperature effects, and shifts in current patterns. More importantly, the collapse of plankton productivity around the rim of the North Pacific (Sancetta 1983) likely led to the collapse of many pollock populations, which are regulated by

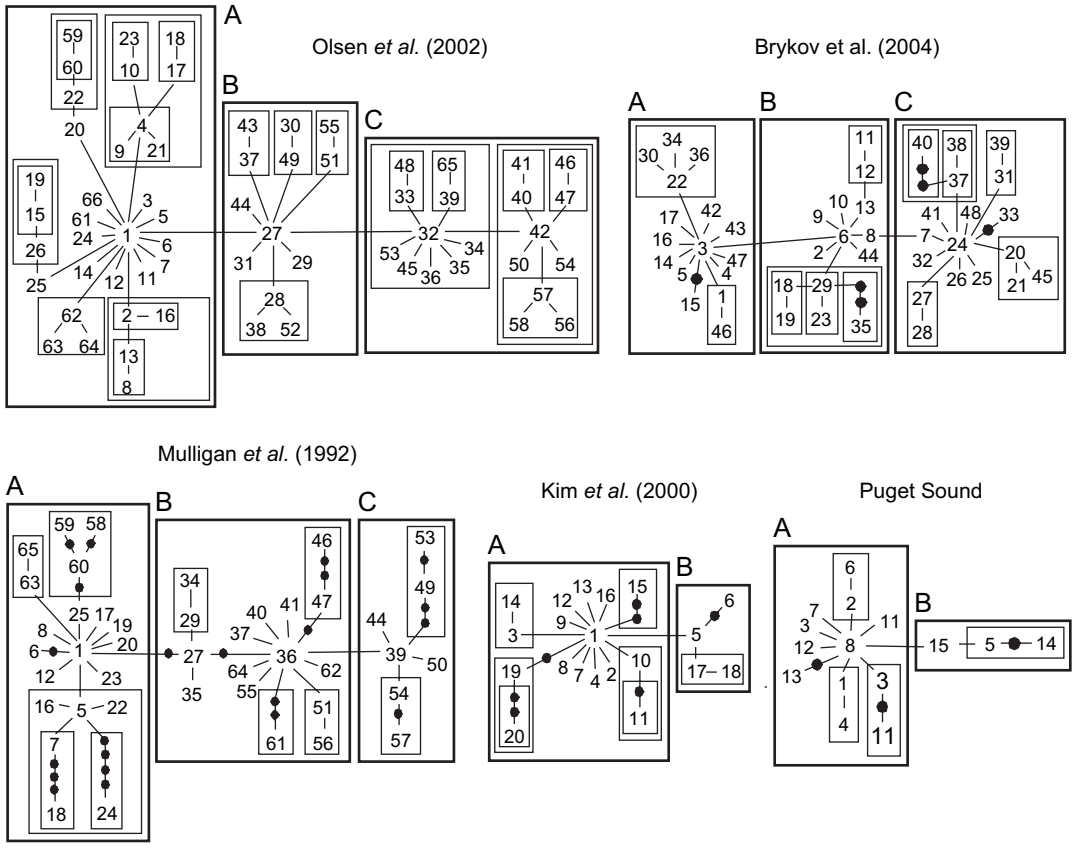

Figure 3. Minimum spanning trees for mtDNA data sets of Mulligan et al. (1992), Kim et al. (2000), Olsen et al. (2002), Brykov et al. (2004), and the present study. Major haplogroups (bold boxes) were identified with the procedures of Templeton et al. (1987).

bottom–up trophic dynamics (Ware and Thomson 2005). Climatic upheavals apparently led to population extinctions in the central North Pacific, producing 2 large refuge populations in the Northwestern and Northeastern Pacific. These populations subsequently expanded into the central North Pacific and marginal seas after glaciers receded 12 000–14 000 years ago.

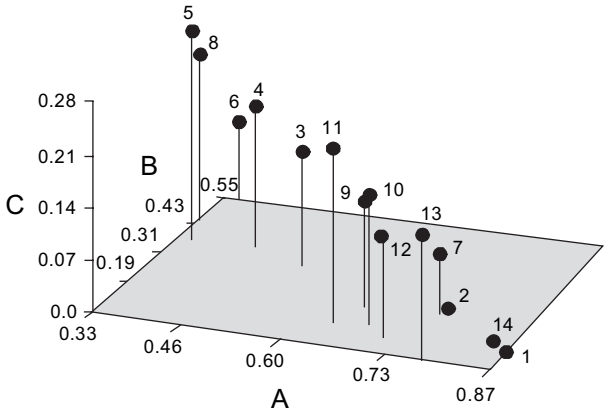

Figure 4. Three-dimensional ordination of samples of pollock based on mtDNA haplogroup frequencies.

Genetic imprints of these colonizations may still be present in contemporary populations. Under an historical scenario, the geographic patterns in haplotype frequencies across the Bering Sea are a legacy of ice-age separation and secondary contact between refuge populations. Genetic imprints of ice-age vicariances and re-associations appear in other North Pacific fishes, including Pacific herring (Grant and Utter 1984), yellowfin sole (Grant et al. 1983), and Pacific cod (Grant et al. 1987). In pollock, the pervasive departures from neutrality among samples may reflect these ice-age population disturbances.

Features of mtDNA variability in pollock, however, diminish the suitability of this model for explaining parallel clines on both sides of the North Pacific. The highest haplotype and nucleotide diversities appear in northern populations, an unlikely pattern, if colonization occurred in a stepwise fashion by "pioneer" migrants (Hewitt 1996). If pioneer colonization were important, reduced levels of diversity in northern populations should appear from founder effects, and a chaotic haplotype frequency distribution would be imprinted on pollock populations. These 2 features characterize populations of freshwater fishes in previously glaciated areas of North America (Bernatchez and Wilson 1998).

Table 3. Haplotype diversity, divergence among populations (F_{ST}), and correlations between haplogroup frequency and distance from southernmost location, and between frequency and SST

				Distance		SST	
Haplogroup	*h*	F_{ST}	*P*	*r*	*P*	*r*	*P*
Northwestern Pacific–Western Bering Sea (Nos 1–6)							
A	0.500	0.054	0.0016	−0.992	0.0001	0.942	0.0049
B	0.448	0.059	0.0013	0.946	0.0042	−0.920	0.0095
C	0.258	0.053	0.0028	0.868	0.0251	−0.815	0.0483
Northeastern Pacific–Eastern Bering Sea (Nos 9–14)							
A	0.498	0.015	0.0166	−0.902	0.0139	0.912	0.0113
B	0.231	0.015	0.0179	0.318	0.5384	−0.345	0.5025
C	0.198	0.022	0.0037	0.835	0.0384	−0.823	0.0444

However, in marine waters, the potential for movement is greater because of the lack of land barriers to dispersal. If colonization occurred by "tidal waves" of migrants, no latitudinal gradient would be expected. Additionally, the lack of intermediate haplotype frequencies in the central part of the range does not support a hypothesis of mixing between 2 divergent populations.

While the parallel clines on both sides of the North Pacific likely reflect the effects of natural selection (see below), genetic profiles of some populations may be better explained by patterns of postglacial colonization. In particular, the population near the western Aleutians (location 8) is more similar in haplogroup frequencies and genetic diversity to Russian populations than to neighboring North American populations. This population does not have haplogroup frequencies predicted by the observed correlation with temperature among North American populations. A similar genetic affinity between Asian and western Aleutian Island populations has also been observed for sockeye salmon (Habicht C, personal communication), and both examples likely reflect similar colonization histories of the Aleutians.

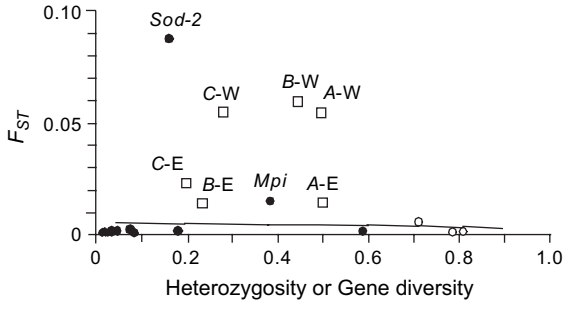

Figure 5. F_{ST} and gene diversity in pollock. Closed circles represent 24 allozyme loci (Olsen et al. 2002), and open circles represent 3 microsatellite DNA loci (Olsen et al. 2002). Open squares represent mtDNA haplogroups from western (W) and eastern (E) areas of the study area or range. Line represents upper 95% limit of expected values under the infinite-alleles and island model of population structure.

Isolation by Distance

Under a second scenario, haplotype distributions arise from a balance between gene flow and random drift. If gene flow were restricted by larval retention mechanisms and spawning-site fidelity (Sinclair 1988), random drift would produce frequency differences among populations. Stepping-stone gene flow between populations along continental margins could then lead to isolation by distance and potentially produce the observed haplotype frequency clines.

Two factors, however, argue against an isolation-by-distance explanation for the parallel clines in the Northwestern and Northeastern Pacific. First, drift is expected to be minimal since the last glacial maximum 18 000 year ago, because population sizes have likely numbered in the hundreds of millions of fish (Ianelli et al. 1998). Second, the significantly larger than expected values of F_{ST} for clades **A**, **B**, and **C**, relative to values of F_{ST} for allozyme and microsatellite DNA markers, indicate that the observed geographic variability is unlikely to reflect drift and gene flow alone.

Selection

The rejection of historical and equilibrium neutral models implicates natural selection. One obvious selective agent is SST, which can influence enzyme kinetics, physiology, and, hence, ecological fitness (Powers et al. 1991). A steep temperature gradient occurs in the Northwestern Pacific, where mean SST varies from 22 °C in the southern Sea of Japan to about 4 °C in the Gulf of Anadyr in the northern Bering Sea (COADS). In the Northeastern Pacific, the temperature gradient is shallower, ranging from about 11 °C in Puget Sound to about 5 °C in the southeastern Bering Sea. Significant correlations between clade frequency and temperature support the argument that selection is operating on the mtDNA in pollock. Further, the slopes of the 2 haplotype frequency clines mirror the steepnesses of the corresponding temperature gradients.

In addition to these correlations, excesses of low-frequency haplotypes are also consistent with selection. Even though selection may be operating on a limited portion of mtDNA, departures from neutrality can be detected over the whole mtDNA molecule because of linkage among genes

and the lack of recombination (genetic hitchhiking, Maynard-Smith and Haigh 1974). Reduced diversities in the southern parts of the range may indicate that selection is more intense in warmer waters. The apparent absence of haplogroup **C** in warm waters on both sides of the North Pacific also supports this argument, but larger samples from these areas are needed to confirm this supposition.

Evidence for temperature-mediated selection on some nuclear genes has been found in pollock (Iwata 1975; Canino and Bentzen 2004) and Atlantic cod (Pogson 2001; Karlsson and Mork 2003). In both species, the frequencies of some pantophysin (*Pan* I) alleles are correlated with SST gradients (Canino et al. 2005; Case et al. 2005; Sarvas and Fevolden 2005). In pollock, the most common *Pan* I allele also appears in a mirror-image cline on both sides of the North Pacific (Canino et al. 2005). Frequencies of this allele decrease from 0.91 in the north central Bering Sea to about 0.60 around Japan and in Puget Sound. Frequencies average about 0.80 in a broad area extending across the southern Bering Sea and Gulf of Alaska, an east–west corridor with little change in SST. In that study, a partial Mantel's test showed a highly significant correlation ($r = 0.746$, $P = 0.002$) between SST differences and pairwise F_{ST} (Canino et al. 2005).

In Atlantic cod, the sister species of pollock (Pogson and Mesa 2004), frequencies of the most common *Pan* I allele (*Pan* I^A) vary from 0.51 to 0.76 in the northwestern Atlantic off Canada but drop to 0.38 in Iceland and 0.17 in the North Sea (Pogson et al. 2001). This frequency cline is associated with a west-to-east increase in SST, a pattern induced by the Gulf Stream Current. A similar cline associated with temperature also occurs with latitude from the North Sea, along the Norwegian coast to the Barents Sea (Case et al. 2005; Sarvas and Fevolden 2005).

Also in Atlantic cod, a frequency cline for mtDNA Cyt *b* haplotypes appears across the North Atlantic (Árnason 2004). In this instance, the common haplotype (A) decreases from about 0.82 off Canada to 0.16 in the Baltic Sea. A second haplotype (E) shows a complementary, but less precipitous, cline across the North Atlantic. The distribution of synonymous and replacement substitutions along the 250 bp of cytochrome *b* failed to indicate that selection was affecting this portion of the gene. If selection is responsible for these clines, the selected region lies elsewhere on the mtDNA molecule or on an associated nuclear gene.

Natural selection can influence mtDNA in several ways. First, mutations can alter the functions of encoded proteins and transfer RNAs directly. mtDNA encodes components of enzymes involved in oxidative phosphorylation (OXPHOS), which generates energy and heat, two adaptively important functions. In humans, haplotype frequency differences between northern and equatorial populations have been interpreted to reflect adaptive differences in OXPHOS coupling (Ruiz-Pesini et al. 2004). Haplotypes prevalent in equatorial populations produce more energy than northern haplotypes, which produce more heat. Although pollock are cold blooded, energy efficiency in different temperature environments is nonetheless a key metabolic feature that is subject to strong selection pressure in other vertebrates (Powers et al. 1991).

If selection is focused on a particular mtDNA gene, this gene cannot be identified from the data summary here. A preliminary analysis of 10 mitogenomic sequences (Yanagimoto et al. 2004) from Japan and the Southeastern Bering Sea with Phylogenetic Analysis by Maximum Likelihood (Yang 1998) failed to detect a selective signal for the 13 protein-coding mtDNA genes. However, the locations of these samples and their small sample sizes provide little power for these tests.

A second mode of selection may involve cytonuclear interactions. Proteins encoded by mtDNA form large complexes with proteins encoded by nuclear genes to enable the OXPHOS system. In fruit flies, changes in nuclear background influence the fitness of particular mtDNA haplotypes (Sackton et al. 2003). In primate evolution, a larger energy-demanding brain appears to have driven mtDNA haplotypic evolution (Grossman et al. 2001). A third form of selection may stem from cytonuclear interactions induced by parasitic bacteria (e.g., Jiggins and Tinsley 2005). Although similar interactions have not been documented in fishes, they, nonetheless, carry a heavy load of Rickettsia-like infections (Fryer and Mauel 1997), and future work may demonstrate this effect.

Conclusions

The results reported here prompt an ocean-wide survey of mtDNA sequences in pollock populations to confirm the geographic patterns inferred in this study. Whereas ice-age extinctions and colonizations may explain some components of genetic variability in pollock, natural selection most likely explains the mirror-image clines on both sides of the North Pacific. An additional question is whether a similar regimen of selection has produced the mtDNA clines in Atlantic cod. Although most nucleotide polymorphisms appear to be neutral, changes at only a few nucleotide sites can lead to adaptive shifts (Lande 1981). In the case of mtDNA, which lacks recombination, the signal of selection is imprinted on the entire molecule.

The collection of genetic data for many species of marine fish, and for pollock in particular, has been motivated largely by the need to resolve stock boundaries for fishery management. In this context, molecular genetic markers are most useful when their geographic distributions can be interpreted solely in terms of gene flow and random drift. Little is known about how the influence of selection on genetic markers alters inferences about stock structure. However, a growing body of evidence indicates that mtDNA (Bazin et al. 2006), microsatellite DNA (Kashi et al. 1997; Hammock and Young 2005), and nuclear DNA (Ford 2002) experience higher levels of selection than is assumed in many studies of natural populations.

Acknowledgments

V. Brykov and J. Olsen kindly supplied original data for this study. M. Grant, F. Lecomte and two anonymous reviewers provided insightful comments on the manuscript. Editorial assistance was aptly provided by B. Bowen.

References

Árnason E. 2004. Mitochondrial cytochrome *b* DNA variation in the high-fecundity Atlantic cod: trans-Atlantic lines and shallow gene genealogy. Genetics. 166:1871–1885.

Avise JC. 2000. Phylogeography: the history and formation of species. Cambridge (MA): Harvard University Press.

Baer CF. 1999. Among-locus variation in F_{ST}: fish, allozymes, and the Lewontin-Krakauer test revisted. Genetics. 152:653–659.

Barrie JV, Conway KW. 1999. Late Quaternary glaciation and postglacial stratigraphy of the northern Pacific margin of Canada. Quat Res. 51:113–123.

Bazin E, Glémin S, Galtier N. 2006. Population size does not influence mitochondrial genetic diversity in animals. Science. 28:570–572.

Beaumont MA, Nichols RA. 1996. Evaluating loci for use in the genetic analysis of population structure. Proc R Soc Lond B. 263:1619–1626.

Bernatchez L, Wilson CC. 1998. Comparative phylogeography of nearctic and palearctic fishes. Mol Ecol. 7:431–452.

Brykov VA, Polyakova NE, Priima TF, Katugin ON. 2004. Mitochondrial DNA variation in northwestern Bering Sea walleye pollock, *Theragra chalcogramma* (Pallas). Environ Biol Fishes. 69:167–175.

Canino MF, Bentzen P. 2004. Evidence for positive selection at the pantophysin (*Pan* I) locus in walleye pollock, *Theragra chalcogramma*. Mol Biol Evol. 21:1391–1400.

Canino MF, O'Reilly PT, Hauser L, Bentzen P. 2005. Genetic differentiation in walleye pollock (*Theragra chalcogramma*) in response to selection at the pantophysin (*Pan* I) locus. Can J Fish Aquat Sci. 62:2519–2529.

Case RAJ, Hutchinson WF, Hauser L, Van Oosterhout C, Carvalho G. 2005. Macro- and micro-geographic variation in pantophysin (*Pan* I) allele frequencies in NE Atlantic cod *Gadus morhua*. Mar Ecol Prog Ser. 301:267–278.

Chakraborty R. 1990. Mitochondrial DNA polymorphism reveals hidden heterogeneity within some Asian populations. Am J Hum Genet. 47:87–94.

Ewens WJ. 1972. The sampling theory of selectively neutral alleles. Theor Popul Biol. 3:87–112.

FitzGerald JL, Thorrold SR, Bailey KM, Brown AL, Severin KP. 2004. Elemental signatures in otoliths of larval walleye pollock (*Theragra chalcogramma*) from the northeast Pacific Ocean. Fish Bull. 102:604–616.

Ford MJ. 2002. Applications of selective neutrality tests to molecular ecology. Mol Ecol. 11:1245–1262.

Fryer JL, Mauel MJ. 1997. The rickettsia: an emerging group of pathogens in fish. Emerg Infect Dis. 3:137–144.

Fu Y-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 147:915–925.

Grant WS, Bakkala R, Utter FM, Teel DJ, Kobayashi T. 1983. Biochemical genetic population structure of yellowfin sole, *Limanda aspera*, of the North Pacific Ocean and Bering Sea. Fish Bull. 81:667–677.

Grant WS, Utter FM. 1980. Biochemical genetic variation in walleye pollock, *Theragra chalcogramma*: population structure in the southeastern Bering Sea and the Gulf of Alaska. Can J Fish Aquat Sci. 37:1093–1100.

Grant WS, Utter FM. 1984. Biochemical population genetics of Pacific herring (*Clupea pallasi*). Can J Fish Aquat Sci. 41:856–864.

Grant WS, Zhang CI, Kobayashi T, Ståhl G. 1987. Lack of genetic stock discretion in Pacific cod (*Gadus macrocephalus*). Can J Fish Aquat Sci. 44: 490–498.

Grossman LI, Schmidt TR, Wildman DE, Goodman M. 2001. Molecular evolution of aerobic energy metabolism in primates. Mol Phylogenet Evol. 18:26–36.

Hammock EAD, Young LJ. 2005. Microsatellite instability generates diversity in brain and sociobehavioral traits. Science. 308:1630–1634.

Hewitt GM. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol J Linn Soc. 58:247–276.

Hinckley S. 1987. The reproductive biology of walleye pollock, *Theragra chalcogramma*, in the Bering Sea, with reference to spawning stock structure. Fish Bull. 85:481–498.

Hinckley S, Bailey KM, Picquelle SJ, Schumacher JD, Stabeno PJ. 1991. Transport, distribution, and abundance of larval and juvenile walleye pollock (*Theragra chalcogramma*) in the western Gulf of Alaska. Can J Fish Aquat Sci. 48:91–98.

Hopkins DM. 1972. The paleogeography and climatic history of Beringia during Late Cenozoic time. J Arctic Nord Stud. 12:121–150.

Ianelli J, Fritz L, Honkalehto T, Williamson N, Waters G. 1998. Bering Sea–Aleutian Islands walleye pollock assessment for 1999. In: Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands regions. Section 1. Anchorage (AK): North Pacific Fishery Management Council. p. 1–79.

Iwata M. 1975. Genetic identification of walleye pollock (*Theragra chalcogramma*) populations on the basis of tetrazolium oxidase polymorphism. Comp Biochem Physiol B. 50:197–201.

Jiggins FM, Tinsley MC. 2005. An ancient mitochondrial polymorphism in *Adalia bipuntata* linked to a sex-ratio-distorting bacterium. Genetics. 171:1115–1124.

Karlsson S, Mork J. 2003. Selection-induced variation at the pantophysin locus (*Pan* I) in a Norwegian fjord population of cod (*Gadus morhua* L.). Mol Ecol. 12:3265–3274.

Kashi Y, King D, Soller M. 1997. Simple sequence repeats as a source of quantitative genetic variation. Trends Genet. 13:74–78.

Kim S-S, Moon D-Y, Yang W-S. 2000. Mitochondrial DNA analysis for stock identification of walleye pollock, *Theragra chalcogramma*, from the North Pacific Ocean. Bull Nat Fish Res Dev Inst Korea. 58:22–27.

Lande R. 1981. The minimum number of genes contributing to quantitative variaiton between and within populations. Genetics. 99:541–553.

Mann DH, Peteet DM. 1994. Extent and time of the last glacial maximum in Southwest Alaska. Quat Res. 42:136–148.

Maynard-Smith J, Haigh J. 1974. The hitchhiking effect of a favourable gene. Genet Res. 23:23–35.

Molnia BF. 1986. Glacial history of the northeastern Gulf of Alaska—a synthesis. In: Hamilton TD, Reed KM, Thorson RM, editors. Glaciation in Alaska-the geologic record. Anchorage (AK): Alaska Geological Society. p. 219–235.

Moyer GR, Winemiller KO, McPhee MV, Turner TF. 2005. Historical demography, selection, and coalescence of mitochondrial and nuclear genes in *Prochilodus* species of northern South America. Evolution. 59:599–610.

Mulligan TJ, Chapman RW, Brown BL. 1992. Mitochondrial DNA analysis of walleye pollock, *Theragra chalcogramma*, from the eastern Bering Sea and Shelikof Strait, Gulf of Alaska. Can J Fish Aquat Sci. 49:319–326.

Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.

Oba T, Kato M, Kitrazato H, Koizumi I, Omura A, Sakai T, Takayama T. 1991. Paleoenvironmental changes in the Japan Sea during the last 85,000 years. Paleoceanography. 6:499–518.

Olsen JB, Merkouris SE, Seeb JE. 2002. An examination of spatial and temporal genetic variation in walleye pollock (*Theragra chalcogramma*) using allozyme, mitochondrial DNA, and microsatellite data. Fish Bull. 100:752–764.

O'Reilly PT, Canino MF, Bailey KM, Bentzen P. 2004. Inverse relationship between F_{ST} and microsatellite polymorphism in the marine fish, walleye pollock (*Theragra chalcogramma*): implications for resolving weak population structure. Mol Ecol. 13:1799–1814.

Pogson GH. 2001. Nucleotide polymorphism and natural selection at the pantophysin *Pan* I locus in the Atlantic cod, *Gadus morhua* L. Genetics. 157:317–330.

Pogson GH, Mesa KA. 2004. Positive Darwinian selection at the pantophysin (*Pan* I) locus in marine gadid fishes. Mol Biol Evol. 21:65–75.

Pogson GH, Taggart CT, Mesa KA, Boutilier RG. 2001. Isolation by distance in the Atlantic cod, *Gadus morhua*, at large and small geographic scales. Evolution. 55:131–146.

Powers DA, Lauerman T, Crawford D, DiMichele L. 1991. Genetic mechanisms for adapting to a changing environment. Annu Rev Genet. 25: 629–659.

Rand DM, Kann LM. 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. Mol Biol Evol. 13:735–748.

Reynolds RW. 1988. A real-time global sea surface temperature analysis. J Clim. 1:75–86.

Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, editors. Bioinformatics methods and protocols: methods in molecular biology. Totowa (NJ): Humana Press. p. 365–386. Available from: http://frodo.wi.mit.edu/cgi-bin/primer3/primer3.cgi.

Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC. 2004. Effects of purifying and adaptive selection on regional variation in human mtDNA. Science. 303:223–226.

Sackton TB, Haney RA, Rand DM. 2003. Cytonuclear coadaptation in *Drosophila*: disruption of cytochrome *C* oxidase activity in backcross genotypes. Evolution. 57:2315–2325.

Sancetta C. 1983. Effect of Pleistocene glaciation upon oceanographic characteristics of the North Pacific Ocean and Bering Sea. Deep Sea Res. 30: 851–869.

Sarvas TH, Fevolden SE. 2005. Pantophysin (*Pan* I) locus divergence between inshore *v.* offshore and northern *v.* southern populations of Atlantic cod in the north-east Atlantic. J Fish Biol. 67:444–469.

Schneider S, Roessli D, Excoffier L. 2000. Arlequin ver. 2.00: a software for population genetics data analysis. Geneva, Switzerland: University of Geneva, Genetics and Biometry Laboratory. Available from: http://anthro.unige.ch/arlequin/.

Serobaba II. 1978. Data on the population structure of the walleye pollock, *Theragra chalcogramma* (Pallas), from the Bering Sea. J Ichthyol. 17:219–231.

Shields GF, Gust JR. 1995. Lack of geographic structure in mitochondrial DNA sequences of Bering Sea walleye pollock, *Theragra chalcogramma*. Mol Mar Biol Biotechnol. 4:69–82.

Shubina EA, Mel'nikova MN, Glubokov AI, Mednikov BM. 2004. Analysis of the genetic structure of northwestern Bering Sea walleye pollock, *Theragra chalcogramma*. Environ Biol Fish. 69:177–185.

Sinclair M. 1988. Marine populations. an essay on population regulation and speciation. Seattle (WA): University of Washington Press.

Templeton AR, Boerwinkle E, Sing CF. 1987. A cladistic analysis of phenotypic associations with haplotypes inferred from restrictions endonuclease mapping. I. Basic theory and analysis analysis of alcohol dehydrogenase activity in *Drosophila*. Genetics. 117:343–351.

Ware DM, Thomson RE. 2005. Bottom-up ecosystem trophic dynamics determine fish production in the northeastern Pacific. Science. 308:1280–1284.

Watterson G. 1978. The homozygosity test of neutrality. Genetics. 88: 405–417.

Woodruff SD, Lobker SJ, Wolter K, Worley SJ, Elms JD. 1993. Comprehensive Ocean-Atmosphere Data Set (COADS) Release 1a: 1980–1992. Earth Syst Monit. 4:1–8.

Yanagimoto T, Kitamura T, Kobayashi T. 2004. Complete nucleotide sequence and variation of mitochondrial DNA from 10 individuals of walleye pollock, *Theragra chalcogramma*. Fish Sci. 70:885–895.

Yang Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol Biol Evol. 15:568–573.

Received March 1, 2006
Accepted July 31, 2006

Corresponding Editor: Brian Bowen