MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 4339–4351 doi: 10.1111/j.1365-294X.2010.04815.x

# Multiple ice-age refugia in Pacific cod, *Gadus macrocephalus*

MICHAEL F. CANINO,* INGRID B. SPIES,* KATHRYN M. CUNNINGHAM,† LORENZ HAUSER† and W. STEWART GRANT‡§

*NOAA Fisheries, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, USA, †School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98105, USA, ‡Department of Biological Sciences, University of Alaska Anchorage, 3211 Providence Drive, Anchorage, AK 99508, USA, §Alaska Department of Fish and Game, 333 Raspberry Road, Anchorage, AK 99518, USA

## Abstract

**Pleistocene ice-ages greatly influenced the historical abundances of Pacific cod, *Gadus macrocephalus*, in the North Pacific and its marginal seas. We surveyed genetic variation at 11 microsatellite loci and mitochondrial (mt) DNA in samples from twelve locations from the Sea of Japan to Washington State. Both microsatellite (mean $H$ = 0.868) and mtDNA haplotype (mean $h$ = 0.958) diversities were large and did not show any geographical trends. Genetic differentiation between samples was significantly correlated with geographical distance between samples for both microsatellites ($F_{ST}$ = 0.028, $r^2$ = 0.33) and mtDNA ($F_{ST}$ = 0.027, $r^2$ = 0.18). Both marker classes showed a strong genetic discontinuity between northwestern and northeastern Pacific populations that likely represents groups previously isolated during glaciations that are now in secondary contact. Significant differences appeared between samples from the Sea of Japan and Okhotsk Sea that may reflect ice-age isolations in the northwest Pacific. In the northeast Pacific, a microsatellite and mtDNA partition was detected between coastal and Georgia Basin populations. The presence of two major coastal mtDNA lineages on either side of the Pacific Ocean basin implies at least two ice-age refugia and separate postglacial population expansions facilitated by different glacial histories. Northward expansions into the Gulf of Alaska were possible 14–15 kyr ago, but deglaciation and colonization of the Georgia Basin probably occurred somewhat later. Population expansions were evident in mtDNA mismatch distributions and in Bayesian skyline plots of the three major lineages, but the start of expansions appeared to pre-date the last glacial maximum.**

*Keywords*: *Gadus macrocephalus*, glacial refugia, Pacific cod, Pleistocene, postglacial colonization, vicariance

*Received 12 May 2010; revision received 15 July 2010; accepted 19 July 2010*

## Introduction

The Pleistocene ice-ages greatly influenced the abundances and distributions of marine species in northern seas. Stable isotopes of oxygen in ice and sediment cores reveal millennial-scale ocean-climate cycles in the north Pacific that varied from a few thousand years (Bond *et al.* 1993; Dansgaard *et al.* 1993) to 100 000 years ('Milankovitch' cycles; Imbrie *et al.* 1984). Ice-age conditions dominated the North Pacific about 80% of the time and have been punctuated with warm periods lasting only a few 1000 years (Lambeck *et al.* 2002). At glacial maxima, global sea level dropped 120 m and further reduced coastal habitats by decreasing extents of productive continental shelf areas (Mitrovica 2003). Massive terrestrial glaciers flowed onto continental shelves along most of the northern seas destroying shallow-water habitats, altering salinity and water temperatures (Clark *et al.* 1999; Tanaka & Takahashi 2005) and reducing the

Correspondence: Michael F. Canino, Fax: 1 206 526 6723; E-mail: mike.canino@noaa.gov

complexities of coastal food webs (McDonald *et al.* 1999). Coastal glaciers were prominent features in parts of the Okhotsk and Bering seas and along the northeast Pacific to Puget Sound (Clague 1989; Mann & Peteet 1994; Grosswald 1998). During these glaciations, populations of marine species survived in southern areas beyond the reach of the glaciers, or possibly offshore in northern ice-free areas (Takahashi 1998; Barrie & Conway 1999).

These glaciations influenced species in different ways depending on species' life history patterns and physiological tolerances. Several methods have been used to understand how marine species were affected by glacial cycles. One approach uses ecological niche models to hind-cast ice-age distributions and identify potential refuge areas (e.g., Bigg *et al.* 2008). This approach assumes that the same physical and biological variables that limit contemporary population abundances and distributions also limited historical populations. Another approach uses genetic markers to provide indirect evidence of historical isolations, population bottlenecks and colonizations (e.g., Bigg *et al.* 2008; Kenchington *et al.* 2009). Discontinuities in the geographical distributions of neutral genetic markers may indicate ice-age isolations of populations in refugia followed by secondary postglacial contact between previously isolated populations (Hewitt 1996). Species experiencing boom-bust cycles in abundance may show reductions in genetic diversity, reflecting historical population bottlenecks (Nei *et al.* 1975) or departures from mutation–drift equilibria (Fu 1997). Furthermore, geographical patterns in genetic diversity may give insights into modes of postglacial colonizations (Bernatchez & Wilson 1998). For example, latitudinal gradients in diversity may reflect stepwise colonizations by founders from small, newly established populations (Vucetich & Waite 2003). Alternatively, the absence of these gradients may be evidence for the correlated effects of high levels of gene flow during rapid colonizations (Hewitt 1996).

Here, we use the distributions of microsatellite and mitochondrial (mt) DNA markers to understand how ice-age re-organizations of the North Pacific influenced populations of Pacific cod *Gadus macrocephalus*. Pacific cod occur on continental shelves and slopes to 500 m from 34–63°N latitude on both sides of the North Pacific from the Yellow Sea, to the Okhotsk and Bering seas, and across the northeastern Pacific to Oregon (Allen & Smith 1988). Distributions of contemporary populations respond to changes in temperatures on decadal timescales. For example, cod actively avoid the summer 'cold' pool (0–2 °C) water mass in the Bering Sea (Ciannelli & Bailey 2005) but have followed its northward retreat during recent climate warming (Mueter & Litzow 2008). Pacific cod exhibit biological and ecological traits typical of cods, including high fecundity, fast growth rates and planktonic larval stages (Paul *et al.* 1990; Ormseth & Norcross 2009). These traits can facilitate high levels of gene flow between populations, swift colonizations and rapid population growth.

Despite the large potential for gene flow, several discrete populations have been identified on relatively small geographical scales (reviewed by Gustafson *et al.* 2000). A survey of microsatellite variation among northeastern Pacific populations showed strong isolation by distance that implies at least some restriction of gene flow (Cunningham *et al.* 2009). In addition, populations in the fjord environments of the Georgia Basin showed a greater degree of isolation than expected from the pattern of isolation by distance between coastal populations (Cunningham *et al.* 2009). It is uncertain, however, whether this divergence is because of contemporary or long-term isolation. A similar level of divergence occurs between populations in the northwestern Pacific. Although no mtDNA or random amplified polymorphic DNA frequency differences were detected among populations around Japan, these populations differed from populations in the Bering Sea (Saitoh 1998). Differences were also detected with allozyme markers between populations in the Okhotsk and Bering seas and populations in the Yellow Sea and Sea of Japan (Gong *et al.* 1991) and between the latter seas with microsatellites (Kim *et al.* 2010). On a larger scale, allozyme markers define two major northwest and northeast Pacific groups that meet in the western Bering Sea (Grant *et al.* 1987). This genetic discontinuity is mirrored in other fishes (Grant & Utter 1984; Grant *et al.* 1984) and is interpreted as a zone of sympatry following secondary contact between northwestern and northeastern Pacific groups.

If ice-age climate shifts produced glacial population extinctions and interglacial colonizations, genetic imprints may appear in contemporary populations as allele frequency differences and distinctive mtDNA phylogroups. The results of previous genetic studies indicate that Pacific cod survived ice-age maxima in refugia located on both sides of the North Pacific. However, the number and locations of these refugia are poorly understood. Populations may have survived at the southern margins of coastal glaciers or in northern areas offshore from unglaciated terrestrial areas that served as refugia for mammals (Byun *et al.* 1997), estuarine fishes (O'Reilly *et al.* 1993; Deagle *et al.* 1996) and plants (Warner *et al.* 1982). We posit that the numbers and locations of refugia should reflect the particular shoreline configurations, and the effects of historical isolations and demographic expansions in the Pacific Ocean basin and marginal seas.

## Materials and methods

### *Sample collection and DNA extraction*

Samples were collected from large spawning and prespawning aggregates of Pacific cod at eight general locations across the northeastern Pacific Ocean and at two locations in the northwestern Pacific from January to March, 2003–2006 (Table 1, Fig. 1). Replicate samples were taken at 2-year intervals at two locations, Unimak Pass (UP03 and UP05) and Kodiak Island (KI03 and KI05), Alaska. Two samples from commercial fishing operations in the central Aleutian Islands region (AD and AT) in 2006 were close to a sample collected during a trawl survey in 2005 (AI) (180 and 275 km, respectively). Pectoral fin clips were preserved in 95% ethanol and stored at room temperature prior to DNA extraction. DNA was extracted from pectoral fin tissue using Qiagen DNeasy kits (QIAGEN, Valencia, CA, USA) following the manufacturer's protocols.

### *Microsatellite and mtDNA amplification*

Eleven microsatellite markers isolated from Pacific cod (Canino *et al.* 2005) were screened using polymerase chain reaction (PCR) conditions as described in Cunningham *et al.* (2009). Two segments of mtDNA were sequenced for variation at a 1014 bp of the NADH subunit 2 (ND2) gene and a 495-bp segment of cytochrome *b* (cyt *b*). PCR primers t-Met and c-Trp and thermalcycling protocols following Orbacz & Gaffney (2000) were used to amplify ND2, except that the t-Met forward primer was modified to 5′-AAGCTCTTGGGCCCA-TACCC-3′ following alignment with an Atlantic cod sequence retrieved from GenBank (Accession number AM489716). Primers and PCR protocols for cyt *b* are described in Bakke & Johansen (2005). PCR amplicons were sequenced in both directions, and contigs were assembled using SEQUENCHER 4.8 (Gene Codes Corp., Ann Arbor, MI, USA). Aligned consensus sequences were created using BioEdit 7.0.9.0 (Hall 1999) and used to construct composite individual ND2/cyt *b* haplotypes. Sequences were deposited in GenBank (Accession numbers: ND2: EU729403–EU729423, FJ877032–FJ877135; cyt *b*: EU729368–EU729402, FJ876989–FJ877031).

### *Analysis of microsatellite DNA*

Summary statistics were calculated using FSTAT 2.9.3.2 (Goudet 2001) and GENETIX 4.05 (Belkhir 2000). Fit of genotype frequency to Hardy–Weinberg expectations (HWE) and genotypic linkage equilibrium were determined

**Table 1** Pacific cod sampling locations, collection years, abbreviations, latitude and longitude, and sample sizes for microsatellite (msat) and mitochondrial DNA (mtDNA) screening

| Location | Year | Abbr. | N Latitude | Longitude | Sample size | |
|---|---|---|---|---|---|---|
| | | | | | msat | mtDNA |
| East China Sea, Korea | 2005 | KO | 38°59′ | 128°42′E | 86 | 34 |
| Sea of Okhotsk, Japan | 2005 | JP | 44°20′ | 135°52′E | 90 | 34 |
| Near Islands, AK | 2006 | NI | 52°31′ | 173°52′E | | 36 |
| Central Aleutian Islands, AK | 2005 | AI | 51°50′ | 177°36′W | 92 | |
| Adak Island, AK | 2006 | AD | 51°40′ | 176°36′W | 45 | 36 |
| Atka Island, AK | 2006 | AT | 52°18′ | 173°38′W | 45 | |
| Unimak Pass, AK | 2003 | UP03 | 54°38′ | 168°10′W | 95 | 35 |
| | 2005 | UP05 | 54°38′ | 168°10′W | 87 | |
| Kodiak Island, AK | 2003 | KI03 | 57°48′ | 152°31′W | 94 | 33 |
| | 2005 | KI05 | 57°55′ | 152°18′W | 106 | |
| Hecate Strait, BC, Canada | 2004 | HS | 53°13′ | 130°57′W | 89 | 34 |
| Coastal Washington, WA | 2005 | WA | 47°55′ | 125°33′W | 69 | 37 |
| Strait of Georgia, WA | 2003 | SG | 48°54′ | 123°06′W | 94 | 35 |
| Puget Sound, WA | 2004, 2006 | PS | 47°35′ | 122°30′W | 18 | 32 |

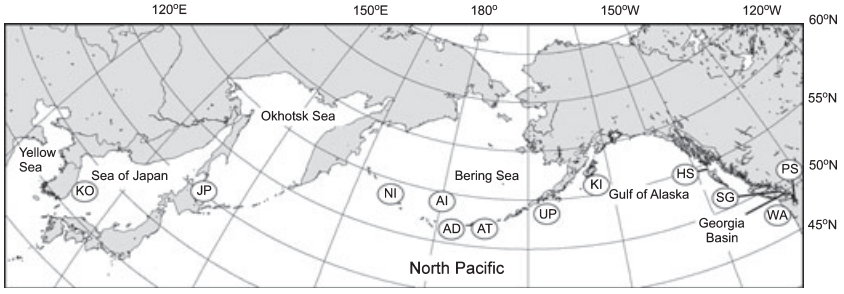


**Fig. 1** Sampling locations for Pacific cod. Sample abbreviations are as in Table 1.

using exact tests in GENEPOP 3.3 (Raymond & Rousset 1995) with Monte Carlo Markov Chain (MCMC) parameters of 5000 dememorization steps, 500 batches and 5000 iterations per batch. MICRO-CHECKER 2.2.3 (van Oosterhout *et al.* 2004) was used to check for the presence of null alleles, scoring errors, stuttering and large allele dropout.

The within-locus *k* test (Reich & Goldstein 1998) was used to compare microsatellite allelic distributions against those expected under mutation–drift equilibrium for evidence of a population expansion in Pacific cod. A recently expanded population is expected to have a smooth, unimodal distribution compared to a ragged, multimodal distribution expected for a deep genealogy in a population of constant size. The inter-locus *g* test (Reich & Goldstein 1998) was used to compare observed versus expected allele size variances across loci. This ratio is expected to be small in a recently expanded population in which allele genealogies show a recent coalescence, but large in a population of constant size because of longer histories of variable mutation rates among loci. Both tests were conducted with the EXCEL macro *kg*tests (Bilgin 2007). Significances of the *k* test values were determined by the proportion of loci giving a positive *k* value using a one-tailed binomial distribution with the probability of a positive *k* set at a conservative level ($P = 0.515$), derived from simulations of populations of constant size (Reich *et al.* 1999). The inter-locus *g* test statistics were compared with fifth-percentile values (Table 1 in Reich *et al.* 1999) for the number of loci and sample sizes.

$F_{ST}$ was estimated with GENEPOP from variances in allele frequencies using the unbiased estimator θ (Weir & Cockerham 1984). $R_{ST}$ (Slatkin 1995) was calculated over all samples and between sample pairs using SPAGEDI 2.1 (Hardy & Vekemans 2002). Significant differences between $F_{ST}$ and $R_{ST}$ were tested using 5000 permutations of allelic sizes among allelic states following Hardy *et al.* (2003). Significantly larger values of $R_{ST}$ indicate that mutation, in addition to drift and gene flow, has contributed to microsatellite frequency differentiation among samples.

The extent of geographical structure was assessed using STRUCTURE 2.3 (Pritchard *et al.* 2000). We used the LOCPRIOR model developed by Hubisz *et al.* (2009) and implemented in STRUCTURE, which allows sample locations to be used as priors in the clustering algorithm. This approach seemed reasonable as a previous analysis of microsatellite data indicated weak structure overall and that at least some of these locations represented distinct populations (Cunningham *et al.* 2009). From one to six population clusters were simulated assuming possible mixed ancestry and correlated allele frequencies by sampling 100 000 MCMC steps after discarding the first 50 000. Graphical output of estimated individual membership coefficients in each cluster was created using DISTRUCT 1.1 (Rosenberg 2004). Adjacent-sample pooling analyses (Buonaccorsi *et al.* 2004) were used to estimate $R_{ST}$ and $F_{ST}$ variance components to identify genetic discontinuities across sampling sites in the North Pacific using ARLEQUIN 3.5 (Excoffier *et al.* 2005) with 50 000 data permutations. The small sample from Puget Sound ($n = 18$) was pooled with the geographically proximate (150 km) sample from the Strait of Georgia, BC, for the analysis after exact tests for genic differentiation, and estimates of $F_{ST}$ between the two groups were not significant. Isolation by distance was tested by searching for correlations between genetic distances and the shortest geographical distances along continental margins (<200 m). Significance was assessed with Mantel's test and 1000 permutations of linearized $R_{ST}$ or $F_{ST}$ and distance matrices using ISOLDE in GENEPOP.

### *Analysis of mitochondrial DNA*

Haplotype (*h*) and nucleotide (π) diversities were estimated with DnaSP 4.50.3 (Rozas *et al.* 2003). Departures from mutation–drift equilibrium under the infinite-sites model were tested with Fu's (1997) $F_S$ with ARLEQUIN and 10 000 data permutations. Analysis of molecular variation in ARLEQUIN with 50 000 permutations and adjacent-sample pooling analysis of $\Phi_{ST}$ and $F_{ST}$ were used to identify genetic discontinuities and to assess the importance of mutation in contributing to population structure. Correlations between linearized $\Phi_{ST}$ and $F_{ST}$ and the shortest geographical distance along continental margins (<200 m) between samples were estimated with Mantel's test and 1000 permutations of distances using ISOLDE in GENEPOP.

We tested for demographic and spatial range expansions with an analysis of the mtDNA nucleotide mismatch distributions using ARLEQUIN, as both processes may have influenced cod populations over successive glacial cycles. The timing of a population expansion was estimated with $\tau = 2ut$ (units of mutation time), where $u$ is the haplotype mutation rate per generation and $t$ is the number of generations. In addition, estimates of the effective female population size, $N_f$, were calculated from mean pairwise differences in the mismatch analysis in ARLEQUIN as $\theta = 2N_f u$ prior to and after the expansion event. Bootstrapped distributions (1000 replicates) were used to assess the fit of the mismatch distribution to an exponential growth model using Harpending's (1994) raggedness statistic, *r*, and to calculate 95% confidence intervals for estimated expansion dates and $2N_f u$ values estimated for times before expansion ($\theta_0$) and for current populations ($\theta_1$). A mean generation time of 6.0 years was used, based on 50% maturity at age 4.4 and 4.9 years in the Gulf of Alaska and eastern Bering Sea, respectively (Stark 2007).

Mutation rates were derived from an estimate of Pliocene divergence from Atlantic cod using 32 homologous Atlantic cod sequences (GenBank Accession numbers: EU877710–EU877741). The HKY+I model of substitution, indicated by the Akaike information criterion in MODELTEST 3.7 (Posada & Crandall 1998), was used to make a relaxed clock estimate of divergence with BEAST 1.4 (Drummond & Rambaut 2007), with a time since divergence between Atlantic cod and Pacific cod lineages at 3.6 Ma (Coulson *et al.* 2006; Carr & Marshall 2008). An MCMC run of 10 million steps produced an effective sample size (ESS) of 5390 and a mutation rate $\mu = 8.29 \times 10^{-9}$ site$^{-1}$ yr$^{-1}$. This calibration yielded a per haplotype rate of $\mu = 1.25 \times 10^{-5}$ yr$^{-1}$.

Historical demographics in three groups northwest Pacific (KO, JP), northeast Pacific (NI, AD, UP03, KI03, HS, WA) and Georgia Basin (SG, PS) were further estimated with coalescence theory and Bayesian skyline plots using BEAST. MCMC chains ranged from 30–168 million steps and were combined to achieve ESSs >320. Pre-expansion $N_e$ was estimated by the average of the first five median values of $\Theta = 2N_e\mu g$, and contemporary $N_e$ was estimated with the final median value of $\Theta g/2\mu$, where $g$ is generation time. Mutational units were converted to calendar time with the estimate of divergence from Atlantic cod.

## Results

### *Microsatellite variation*

Microsatellite heterozygosities ranged from $H_O = 0.523$ (Gma108) to $H_O = 0.955$ (Gma101) over loci and averaged $H_O = 0.868$ (Table S1, Supporting Information). Eight departures from HWE were detected, but only one heterozygote deficit and one excess were significant after correction for multiple tests. Linkage disequilibrium was detected between Gma103 and Gma108 in a single sample (JP), but no significant linkage disequilibrium was detected over all samples. Within-locus $k$ tests produced a significant number of negative $k$ values in nine of the 13 samples ($P \leq 0.05$), for pooled samples from Asia and North America ($P = 0.003$ and 0.004, respectively), and for samples over all ($P = 0.004$). However, no inter-locus $g$-tests were significant in any of the individual or the pooled regional samples.

A large genetic discontinuity between the northwest and northeast Pacific Ocean was detected in the STRUCTURE analysis (Fig. 2). Simulation of three population clusters produced the largest penalized log-likelihood probabilities (average log-likelihood $P$ minus half of the variance) and the smallest estimates for the parameter $r$, indicating that sample information was an informative model prior. $F_{ST}$ ranged from zero to 0.100 over all sample pairs with a global average of 0.028. All pairwise estimates of $F_{ST}$ were significant between samples from Asia (KO and JP) and North America after controlling the false discovery rate to $\alpha = 0.05$ following Benjamini & Hochberg (1995). Samples obtained from Puget Sound in 2004 and 2006 were pooled because of small sample sizes ($n = 9$ each) after a log-likelihood test showed no significant heterogeneity between years. In North America, samples south of Alaska, excluding Puget Sound, were differentiated from those in Alaska and from each other. The smaller Puget Sound sample was marginally differentiated from the other Georgia Basin sample (SG; $P = 0.057$) and from coastal Washington State (WA; $P = 0.067$). $R_{ST}$ over all samples (0.057) was more than twice as large as overall $F_{ST}$. Permutation tests verified that multilocus $R_{ST}$ values were significantly greater than $F_{ST}$ values between all sample pairs comparing Asia and North American populations. Significant correlations between genetic and geographical distance across the entire geographical range were observed for $R_{ST}$ ($r^2 = 0.283$, $P < 0.001$) and $F_{ST}$ ($r^2 = 0.325$; $P < 0.001$) (Fig. 3a). Pairwise estimates for both showed a distinct break between northwest and northeast Pacific populations that were much more pronounced for $R_{ST}$. The slope of the isolation-by-distance (IBD) relationship was about three times greater for $R_{ST}$ (0.042) than for $F_{ST}$ (0.015). Adjacent-sample pooling analysis of 12 samples across the North Pacific showed a strong break for $R_{ST}$ between the pooled Asian samples and the remaining samples, and the least amount of within-group heterogeneity, although $F_{ST}$ did not show this geographical structure (Table 2). Regional patterns of between-group and within-group variance components for $R_{ST}$ and $F_{ST}$ were not observed in northeastern Pacific samples (KI, HS, WA, SG, PS) when the Asian samples were excluded from the analysis, and most estimates for within-group variances were significantly greater than zero.

### *Mitochondrial DNA variation*

A total of 135 substitutions (121 transitions and 14 transversions) at 134 polymorphic sites defined 170 haplotypes (126 singletons) in 346 composite ND2/cyt *b*


**Fig. 2** Results of cluster analysis from STRUCTURE. Vertical lines show inferred membership proportions of individual genotypes in each of K = 3 simulated clusters. Sample abbreviations are as in Table 1.

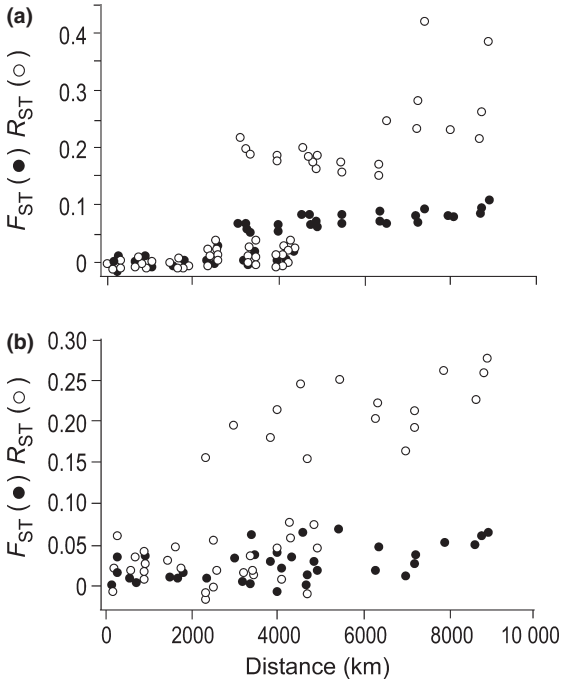


**Fig. 3** Pairwise values of linearized estimator of genetic differentiation versus geographic distance in Pacific cod. (a) $F_{ST}$ (closed circles) and $R_{ST}$ (open circles) values for microsatellites. (b) $F_{ST}$ (closed circles) and $\Phi_{ST}$ (open circles) for mtDNA variation.

mtDNA sequences. Haplotype frequencies for samples are given in Table S2, Supplemental Information. The number of haplotypes per sampled location ranged from 20 to 28. Haplotype diversity ($h$) ranged from 0.923 to 0.984 (mean $h$ = 0.958) but did not show any geographical trends (Table 3). Nucleotide diversity ($\pi$) ranged from 0.003 to 0.004 (mean $\pi$ = 0.0036) and also did not show a trend among samples. In seven of the 10 samples, the number of observed haplotypes was larger than the expected number, but this difference was significant only in UP03. Tests for departures from neutrality or population expansion, as indicated by $F_S$, were significant in all samples ($P < 0.01$). A parsimony haplotype network for these three groups appears in Fig. 4. Only four haplotypes (2.4%) were present in all three groups, six additional haplotypes were shared between the northeast Pacific and Georgia Basin samples, and two additional haplotypes were shared between the northeast Pacific and northwest Pacific samples. The remaining 158 haplotypes were unique to one of the three groups.

As with the microsatellite data, significant genetic differentiation based upon haplotype frequencies, $F_{ST}$, or $\Phi_{ST}$, was found in all pairwise comparisons between Asian and North American samples, although the two samples from the northwest Pacific (KO and JP) did not differ ($P$ = 0.177). One Georgia Basin sample (SG) was differentiated from all other northeast Pacific coastal samples, excluding UP03, but consistent patterns of differentiation among the latter group were not found. Significant isolation by distance was evident in mtDNA variation across the entire sample range. Correlations between geographical distance and $F_{ST}$ ($r^2$ = 0.183;

**Table 2** Adjacent-sample pooling analysis of microsatellite variation among 11 samples of Pacific cod extending from the Sea of Japan, Korea to Puget Sound, Washington. Sample abbreviations are as in Table 1

| Sample pooling scheme | | | | | | | | | | $R_{ST}$ | | | | $F_{ST}$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| KO | JP | AK | AD | AT | UP03-05 | KI03-05 | HS | WA | SG+PS | $R_{CT}$ | $P$ | $R_{SC}$ | $P$ | $F_{CT}$ | $P$ | $F_{SC}$ | $P$ |
| **(*)** | (* | * | * | * | * | * | * | * | *) | 0.076 | 0.100 | 0.027 | <0.001 | −0.0004 | 0.504 | <0.001 | 0.0017 |
| (* | *) | (* | * | * | * | * | * | * | *) | 0.098 | 0.022 | 0.007 | <0.001 | 0.0000 | 0.312 | <0.001 | 0.0023 |
| (* | * | *) | (* | * | * | * | * | * | *) | 0.039 | 0.042 | 0.024 | <0.001 | 0.0000 | 0.479 | <0.001 | 0.0011 |
| (* | * | * | *) | (* | * | * | * | * | *) | 0.030 | 0.014 | 0.027 | <0.001 | 0.0000 | 0.798 | <0.001 | 0.0009 |
| (* | * | * | * | *) | (* | * | * | * | *) | 0.021 | 0.008 | 0.030 | <0.001 | 0.0000 | 0.936 | <0.001 | 0.0010 |
| (* | * | * | * | * | *) | (* | * | * | *) | 0.008 | 0.144 | 0.037 | <0.001 | 0.0000 | 0.359 | <0.001 | 0.0065 |
| (* | * | * | * | * | * | *) | (* | * | *) | 0.005 | 0.091 | 0.039 | <0.001 | 0.0000 | 0.108 | <0.001 | 0.0522 |
| (* | * | * | * | * | * | * | *) | (* | *) | 0.002 | 0.267 | 0.040 | <0.001 | 0.0001 | 0.067 | <0.001 | 0.0011 |
| (* | * | * | * | * | * | * | * | *) | (*) | 0.003 | 0.300 | 0.042 | <0.001 | 0.0003 | 0.101 | <0.001 | 1.0000 |
| (* | * | * | * | * | * | * | * | * | *) | 0.041 | <0.001 | | | 0.0001 | 0.001 | | |
| | | (*) | (* | * | * | * | * | * | *) | 0.0003 | 0.250 | 0.008 | <0.001 | −0.0001 | 0.753 | <0.001 | 0.002 |
| | | (* | *) | (* | * | * | * | * | *) | 0.0024 | 0.212 | 0.007 | <0.001 | 0.0000 | 0.787 | <0.001 | 0.003 |
| | | (* | * | *) | (* | * | * | * | *) | 0.0039 | 0.108 | 0.006 | <0.001 | 0.0000 | 0.893 | <0.001 | 0.002 |
| | | (* | * | * | *) | (* | * | * | *) | 0.0038 | 0.085 | 0.006 | <0.001 | 0.0000 | 0.774 | <0.001 | 0.007 |
| | | (* | * | * | * | *) | (* | * | *) | 0.0039 | 0.108 | 0.006 | <0.001 | 0.0000 | 0.893 | <0.001 | 0.002 |
| | | (* | * | * | * | * | *) | (* | *) | 0.0103 | 0.036 | 0.004 | <0.001 | 0.0001 | 0.107 | <0.001 | 0.180 |
| | | (* | * | * | * | * | * | *) | (*) | 0.0170 | 0.125 | 0.003 | <0.001 | 0.0003 | 0.111 | <0.001 | 1.000 |
| | | (* | * | * | * | * | * | * | *) | 0.0079 | <0.001 | | | 0.0001 | 0.019 | | |

| Sample | *N* | $A_O$ | $A_E$ | *h* | *k* | π | $F_S$ |
|---|---|---|---|---|---|---|---|
| KO | 34 | 20 | 15.6 | 0.9234 | 4.45 | 0.003 | −9.319*** |
| JP | 34 | 27 | 24.9 | 0.9768 | 4.74 | 0.003 | −22.224*** |
| NI | 36 | 28 | 28.4 | 0.9841 | 5.54 | 0.004 | −20.436*** |
| AD | 36 | 23 | 18.2 | 0.9397 | 5.53 | 0.004 | −10.849*** |
| UP03 | 35 | 24** | 17.0 | 0.9328 | 4.83 | 0.003 | −14.758*** |
| KI03 | 33 | 27 | 24.4 | 0.9773 | 5.44 | 0.004 | −20.857*** |
| HS | 34 | 22 | 22.6 | 0.9679 | 5.82 | 0.004 | −9.554** |
| WA | 37 | 26 | 25.8 | 0.9745 | 5.08 | 0.003 | −16.871*** |
| SG | 35 | 23 | 20.0 | 0.9529 | 5.73 | 0.004 | −10.853*** |
| PS | 32 | 21 | 17.8 | 0.9456 | 6.39 | 0.004 | −7.951** |

**Table 3** Sample size (*N*), observed and expected number of mtDNA haplotypes (*A*), haplotype diversity (*h*), average number of pairwise nucleotide differences (k), nucleotide diversity (π) and Fu's ($F_S$) test of neutrality for Pacific cod. Sample abbreviations are as in Table 1

*$P \leq 0.05$; **$P \leq 0.001$; ***$P \leq 0.001$.

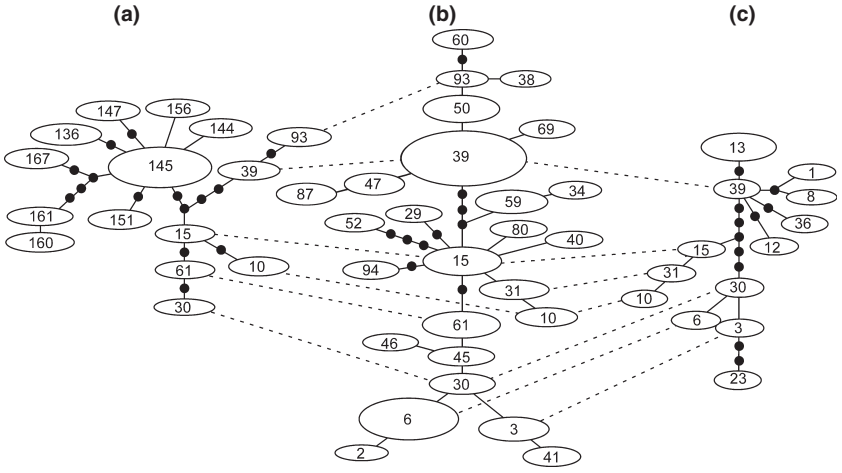


**Fig. 4** Parsimony-based mitochondrial DNA haplotype networks (singletons excluded) for three majors groups of Pacific cod. (a) Northwest Pacific; (b) northeast Pacific; (c) Georgia Basin. Dashed lines connect haplotypes shared between groups. Closed circles represents inferred but unsampled intermediate haplotypes.

$P = 0.003$) or $\Phi_{ST}$ ($r^2 = 0.368$; $P < 0.013$) were both significant across the entire range (Fig. 3b), but not between samples taken only from North America (~4850 km). An adjacent-sample pooling analysis using two groups indicated that $\Phi_{ST}$ was consistently larger than $F_{ST}$ in the western North Pacific. This gap

**Table 4** Adjacent-sample pooling analysis of mitochondrial DNA variation among 10 samples of Pacific cod extending from the Sea of Japan, Korea, to Puget Sound, Washington. Sample abbreviations are as in Table 1

| Sample pooling scheme | | | | | | | | | | $\Phi_{ST}$ | | | | $F_{ST}$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| KO | JP | NI | AD | UP03 | KI03 | HS | WA | SG | PS | $\Phi_{CT}$ | *P* | $\Phi_{SC}$ | *P* | $F_{CT}$ | *P* | $F_{SC}$ | *P* |
| (*) | (* | * | * | * | * | * | * | * | *) | 0.157 | 0.096 | 0.067 | <0.001 | 0.031 | 0.097 | 0.021 | <0.001 |
| (* | *) | (* | * | * | * | * | * | * | *) | 0.188 | 0.023 | 0.027 | 0.003 | 0.026 | 0.021 | 0.018 | <0.001 |
| (* | * | *) | (* | * | * | * | * | * | *) | 0.093 | 0.010 | 0.057 | <0.001 | 0.010 | 0.061 | 0.022 | <0.001 |
| (* | * | * | *) | (* | * | * | * | * | *) | 0.053 | 0.018 | 0.073 | <0.001 | 0.003 | 0.213 | 0.025 | <0.001 |
| (* | * | * | * | *) | (* | * | * | * | *) | 0.028 | 0.075 | 0.086 | <0.001 | 0.003 | 0.250 | 0.025 | <0.001 |
| (* | * | * | * | * | *) | (* | * | * | *) | 0.025 | 0.155 | 0.088 | <0.001 | 0.003 | 0.265 | 0.025 | <0.001 |
| (* | * | * | * | * | * | *) | (* | * | *) | 0.009 | 0.132 | 0.096 | <0.001 | 0.007 | 0.111 | 0.023 | <0.001 |
| (* | * | * | * | * | * | * | *) | (* | *) | 0.025 | 0.064 | 0.092 | <0.001 | 0.019 | 0.044 | 0.020 | <0.001 |
| (* | * | * | * | * | * | * | * | *) | (*) | 0.001 | 0.310 | 0.100 | <0.001 | 0.011 | 0.200 | 0.025 | <0.001 |
| | | (*) | (* | * | * | * | * | * | *) | 0.173 | 0.023 | 0.030 | 0.0002 | 0.019 | 0.065 | 0.019 | <0.001 |
| | | (* | *) | (* | * | * | * | * | *) | 0.138 | 0.014 | 0.026 | 0.0008 | 0.013 | 0.002 | 0.020 | <0.001 |
| | | (* | * | *) | (* | * | * | * | *) | 0.123 | 0.007 | 0.021 | 0.0360 | 0.014 | 0.036 | 0.018 | <0.001 |
| | | (* | * | * | *) | (* | * | * | *) | 0.149 | 0.002 | 0.022 | 0.0030 | 0.025 | 0.005 | 0.014 | <0.001 |
| | | (* | * | * | * | *) | (* | * | *) | 0.132 | 0.002 | 0.016 | 0.0170 | 0.029 | 0.002 | 0.009 | <0.001 |
| | | (* | * | * | * | * | *) | (* | *) | 0.072 | 0.024 | 0.055 | <0.001 | 0.018 | 0.004 | 0.015 | <0.001 |
| | | (* | * | * | * | * | * | *) | (*) | 0.039 | 0.033 | 0.074 | <0.001 | 0.012 | 0.018 | 0.018 | <0.001 |

diminished across the North Pacific, and the two statistics were similar in the eastern North Pacific (Table 4). In this analysis, a separation between Asian samples (KO, JP) and the remaining samples from North America showed the largest divergence ($\Phi_{CT}$ = 0.188, $P$ = 0.023; $F_{CT}$ = 0.026, $P$ = 0.021) and the smallest within-group divergences between samples ($\Phi_{SC}$ = 0.027, $P$ = 0.003; $F_{SC}$ = 0.018, $P$ = 0.000). However, all values of $\Phi_{SC}$ and $F_{SC}$ were significant ($P < 0.010$), reflecting a pattern of isolation by distance among populations. An adjacent-sample pooling analysis of eight samples, excluding the Asian samples, indicated a significant break between samples from the northeast Pacific and the two samples from the Georgia Basin ($\Phi_{CT}$ = 0.029, $P$ = 0.037; $F_{CT}$ = 0.023, $P$ = 0.036) (Table 4). This partition also produced the smallest divergences among populations within the two groups ($\Phi_{SC}$ = 0.016; $F_{SC}$ = 0.010), even though both were significant ($P \leq 0.026$).

Nucleotide mismatch distributions were generally unimodal for each of the phylogroups (Fig. 5), although none conformed strictly to a sudden demographic or spatial expansion model ($P$ values for Harpending's $r$ = 0.090–0.980 for six tests). Expansion events for the three phylogroups dated to the mid-Pleistocene, ranging from *ca.* 297 to 218 kyr BP (Table 5). Pre-expansion estimates of $N_{f0}$ were below 10 000 for all groups and postexpansion estimates ranged from 241 to 478 million females. Net sequence divergences ($d$) among the three clades, assuming an estimated mutation rate of $\mu = 8.29 \times 10^{-9}$ site$^{-1}$ yr$^{-1}$, yielded an estimated separation between the northeast Pacific and Georgia Basin phylogroups about 16 000 years ago.

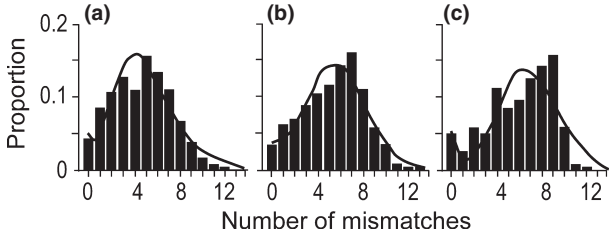


**Fig. 5** Mitochondrial DNA haplotype nucleotide mismatch distributions for Pacific cod phylogroups. (a) Northwest Pacific; (b) northeast Pacific; (c) Georgia Basin. curves are fitted to bootstrapped mismatch distributions.

Bayesian skyline plots of northwest Pacific, northeast Pacific and Georgia Basin population groups all indicated recent population expansions, with the northeast Pacific group showing the largest growth from about 1.5 million to over 43 million effective females (Fig. 6). The expansion of the northwest Pacific populations was less abrupt than expansions in the other two groups and began about 500 kyr BP. Onsets of expansion occurred at about 145 kyr BP for Georgia Basin populations and at about 181 kyr BP for northeast Pacific populations.

## Discussion

The biogeographical conclusions of this study are based on surveys of microsatellite and mtDNA population markers across most of the contemporary geographical distribution of Pacific cod. The range of mutation rates in these markers allowed the resolution of both deep divergences resulting from ice-age isolations and contemporary population structure shaped by postglacial colonizations and gene flow. In particular, maternally inherited mtDNA resolved evolutionary lineages that were not apparent in tests of nuclear marker allele frequencies. The STRUCTURE analysis of microsatellites and the adjacent-sample pooling analysis for both marker classes confirmed the presence of two major evolutionary lineages in the northwestern and northeastern Pacific Ocean previously reported by Grant *et al.* (1987), plus another now confined to the fjord environments of the Georgia Basin. These markers also revealed additional genetic structure within these regions that likely arose from ice-age perturbations. Considerable separation between the northwestern and northeastern Pacific groups is indicated by large microsatellite ($R_{ST}$ = 0.098,

**Table 5** Estimates of historical and contemporary demographic parameters based on mtDNA mismatch distributions for three regional Pacific cod mtDNA clades fit to an exponential population growth model. Θ and effective female population sizes (in thousands) are estimated before ($\Theta_0$, $N_{f0}$) and after ($\Theta_1$, $N_{f1}$) population expansions, respectively. Estimates of times to population expansions were calculated from ($\tau = 2ut$), where $u$ is the haplotype mutation rate per generation and $t$ is the number of generations. Sample abbreviations are as in Table 1

| | | $N_{f0}$ (thousands) | | | | $N_{f1}$ (thousands) | | | Expansion event (kyr BP) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Phylogroup | Samples | $n$ | $\Theta_0$ | mean | 95% c.i. | $\Theta_1$ | mean | 95% CI | $\tau$ | date | 95% CI |
| NW Pacific | JP, KO | 68 | 0.24 | 1.6 | 0–17.6 | 18041.1 | 477621 | 210810–2647337 | 5.49 | 218 | 102–339 |
| NE Pacific | NI, AD, UP03, KI03, HS, WA | 211 | 0.38 | 10.0 | 0–74.8 | 14139.2 | 374317 | 360–2647337 | 6.34 | 252 | 120–340 |
| Georgia Basin | PS, SG | 67 | 0.28 | 7.3 | 0–7.6 | 9116.6 | 241349 | 400–2647337 | 7.49 | 297 | 151–407 |

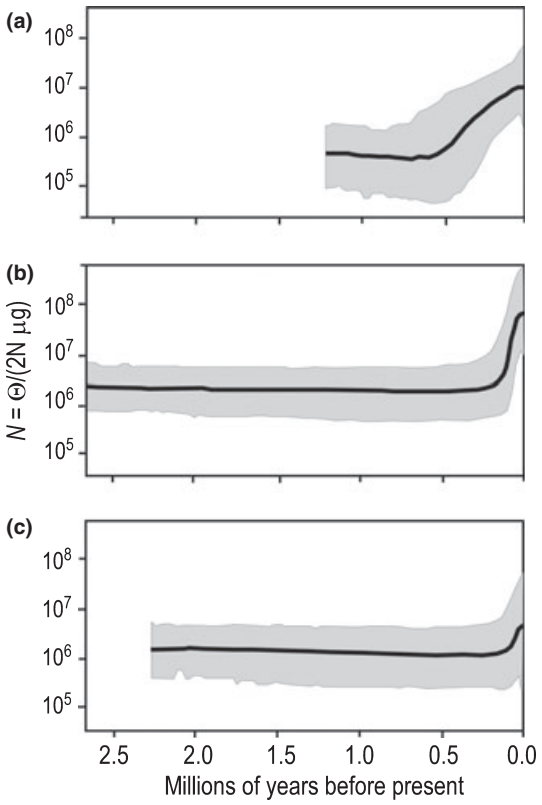


**Fig. 6** Bayesian skyline plots of historical effective female population sizes for three Pacific cod phylogroups. (a) Northwest Pacific; (b) northeast Pacific; (c) Georgia Basin. Shaded areas represent upper and lower bounds of 95% highest posterior density intervals.

$P = 0.020$) and mtDNA ($\Phi_{ST} = 0.188$, $P = 0.023$) divergences, by unique distributions of haplotypes in each group and by different demographic histories. The glacial history of the North Pacific has also produced similar genetic subdivisions in other marine species (*e.g.* Grant & Utter 1984), although the locations of the contact zones differ among species because of differences in life history patterns that influenced postglacial colonization patterns.

Alternatively, this genetic architecture could potentially reflect different selective pressures across the North Pacific. Selection has been invoked to explain ocean-wide frequency shifts in mtDNA haplotypes in Atlantic cod (Árnason 2004) and walleye pollock (Grant *et al.* 2006), but it is unlikely to influence microsatellite and mtDNA variants in the same way. The observed geographical frequency shifts appear to be better explained by historical isolations and limited gene flow between genetically diverged populations (Vasemägi 2006). Hence, we examine the results in the light of palaeoceanographical data for the northwest Pacific, Bering Sea and northeast Pacific for evidence of ice-age refugia.

### *Northwestern Pacific Ocean and marginal seas*

The existence of at least two Asian groups has been supported by significant allozyme (Gong *et al.* 1991) and microsatellite (Kim *et al.* 2010) frequency differences between samples from the Yellow Sea and Sea of Japan and samples from the Okhotsk Sea. In this study, microsatellite, but not mtDNA variants, showed significant frequency differences between the southern Sea of Japan (KO) and the Okhostk Sea (JP). Although these two groups differed in microsatellite allele frequencies, they belonged to the same mtDNA phylogroup. Both the mismatch distribution and the Bayesian skyline plots gave signals of population growth that began in the middle Pleistocene, well before the start of postglacial warming *ca.*15 kyr BP. These genetic signals appear to have been integrated over several ice-age cycles and, hence, do not correlate with a particular palaeoclimatic event. An apparent population expansion pre-dating the most recent postglacial warming has also been observed in the mtDNA architecture of Atlantic cod (Carr & Marshall 2008). Microsatellite $R_{ST}$ and mtDNA $\Phi_{ST}$ estimates exceeded corresponding values of $F_{ST}$ to a greater extent in northwest Pacific than in northeast Pacific samples, indicating longer isolations among populations so that mutation, in addition to random drift, has contributed to genetic population structure (Hardie *et al.* 2006).

The presence of two contemporary Asian groups of Pacific cod may be explained by more recent subdivision of the mid-Pliestocene phylogroup expansion from a single refugium. This could have arisen during sea level low stands associated with ice-age cooling, when access to the Sea of Japan through the northern Tatarskiy, Soya and Tsugaru straits was blocked and largely limited by a sill across the southern Tsushima Strait (Kitamura *et al.* 2001). The genetic structure of the redlip mullet, *Chelon haematocheilus*, was also apparently shaped by ice-age isolations in the South China Sea, East China Sea and the Sea of Japan (Liu *et al.* 2007). A northern group of Pacific cod was probably isolated along the Pacific Ocean shores of Japan, as the Okhotsk Sea was smaller during glacial peaks, covered with sea ice and had reduced levels of productivity (Okazaki *et al.* 2005), conditions that would have been inhospitable to Pacific cod.

### *Northeastern Pacific Ocean and Bering Sea*

Samples from the northeast Pacific showed significant departures from neutrality that are likely due to population expansion. Both the mismatch distribution and Bayesian skyline plot for the northeast Pacific samples indicate that population growth pre-dated the most

recent postglacial warming. As with the results for the Asian populations, signals of population growth appear to have been integrated over several ice-age cycles.

Present-day populations of Pacific cod in the northeast Pacific and Bering Sea most likely originated from a southern refuge population, as oceanic conditions in the Bering Sea and along the shores of the Gulf of Alaska would not have supported Pacific cod populations during ice-age maxima. During these periods, terrestrial glaciers flowed onto continental shelves along the northeast Pacific from the Aleutian Islands and along the shores of the Gulf of Alaska and British Columbia to Puget Sound (Mann & Peteet 1994; Mann & Hamilton 1995; Barrie & Conway 1999), likely forcing surviving populations southward. Spawning and nursery habitats were greatly altered by drops in sea level by as much as 120 m during glacial maxima that exposed large areas of the continental shelf in the Bering Sea (Bering Land Bridge) and along the northeast Pacific (Mann & Peteet 1994; Mann & Hamilton 1995; Ager 2003). The southern outflow of Beringian rivers produced a low-salinity lens that intensified sea-ice formation (Sancetta *et al.* 1985), and floating continental ice sheets covered the southeastern Bering Sea (Sancetta & Robinson 1983; Tanaka & Takahashi 2005). This smaller Bering Sea no longer supported a productive shallow-water ecosystem. Drops in sea level also reduced the sizes of suitable near-shore spawning and nursery habitats around the Gulf of Alaska.

### *Georgia Basin*

The genetic separation between northeast Pacific coastal and Georgia Basin populations observed in both marker classes indicates long-term allopatric isolation. Only ten of the 135 observed haplotypes in these regions were shared between the two phylogroups. Frequencies of microsatellite alleles differed to a greater extent between these groups than expected by the IBD relationships observed among coastal populations (Cunningham *et al.* 2009). Individual assignments by STRUCTURE produced a large membership of Georgia Basin fish in the third of three simulated clusters. The abrupt transition of those assignments from the Washington State coast to the Georgia Basin (Fig. 2) supports the presence of a spatial barrier isolating these two groups (Cunningham *et al.* 2009).

The origins of the northeast Pacific and Georgia Basin phylogroups have been determined by timings of glacial maxima and deglaciations. The most parsimonious explanation is a single refuge located below the southern limits of coastal ice during glacial maxima. The clear isolation-by-distance pattern and large microsatellite diversities in the coastal northeast Pacific groups suggest an unimpeded postglacial expansion of cod into the Gulf of Alaska, Bering Sea and Aleutian Islands. Our divergence date estimate of *ca.* 16 kyr BP between the northeast Pacific and Georgia Basin phylogroups is consistent with colonization of the Georgia Basin by fish from an outer coastal population following the last Pleistocene glaciation. The Puget lobe of the Cordilleran ice sheet extended across the northern portion of the Georgia Basin as late as 14.5–14.0 kyr BP (Porter & Swanson 1998; Clague & James 2002), although Puget Sound was largely ice-free by about 15 kyr BP (Menounos *et al.* 2008). Glaciers in more northern coastal areas receded a few thousand years earlier than glaciers in southern areas. Hence, colonization of the Gulf of Alaska was possible by 14–15 kyr BP, but colonization of the Georgia Basin may not have occurred until later (Mann & Hamilton 1995; Menounos *et al.* 2008).

The complex genetic structure of Pacific cod populations reflects both limits on gene flow that allows divergence among contemporary populations, and divergence between populations isolated during late Pleistocene glaciations. Isolated populations may have survived in fjord-like environments at the edge of the continental shelf around the NE Pacific in areas that provided suitable spawning and nursery habitats. In addition to a refuge at the southern margin of the coastal ice sheet, refugia for Pacific cod may also have existed offshore from the Queen Charlotte Islands (O'Reilly *et al.* 1993) and offshore from Beringia (Bickham *et al.* 1995; Seeb & Crane 1999). Limited amounts of dispersal or incomplete lineage sorting since postglacial colonization can account for shared haplotypes between coastal and Georgia Basin populations.

In conclusion, the results of this study confirm the presence of a genetic discontinuity across the Bering Sea that represents a secondary contact zone between two major population groups isolated by mid-Pleistocene glaciation. The relatively large amounts of genetic diversity in present-day populations of Pacific cod indicate that they did not experience severe population bottlenecks during the Pleistocene. At least two glacial refugia existed in coastal areas on both sides of the North Pacific. Additional ice-age isolations in marginal seas in the northwest Pacific and staggered deglaciations in the northeast Pacific likely produced additional genetic subdivision and distinct evolutionary lineages in contemporary populations.

## Acknowledgements

We thank S. McFarlane, W. Palsson, V. Okimura, A. Abookire, R. Wilmot, D. Kubiak, O. Ormseth, K. Spalinger, D. Urban, D. Pengilly, S. Neidetcher, P. Munro, L. Conners, L. Logerwell, J. Shaklee, S. Young, G. Bargmann, S. Hoffman, K. Masse, M.

Rust, Q. Jo, D. Jang, T. Yanagimoto and AFSC Staff biologists for sample collections. M. Grant, B. Bowen and two anonymous referees provided helpful comments on earlier manuscript drafts. This work was funded, in part, by a grant from the Washington Sea Grant Program, University of Washington, pursuant to National Oceanic and Atmospheric Administration Award No. NA04OAR4170032, Project No. R/F-147. The views expressed herein are those of the author(s) and do not necessarily reflect the views of NOAA or any of its sub-agencies. This research is contribution EcoFOCI-0750 to NOAA's Fisheries-Oceanography Coordinated Investigations.

## References

Ager TA (2003) Late Quaternary vegetation and climate history of the central Bering land bridge from St. Michael Island, western Alaska. *Quaternary Research*, **60**, 19–32.

Allen MJ, Smith GB (1988) Atlas and zoogeography of the common fishes in the Bering Sea and northeastern Pacific. *NOAA Technical Report NMFS*, **66**, 151.

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.

Bakke I, Johansen SD (2005) Molecular phylogenetics of Gadidae and related Gadiformes based on mitochondrial DNA sequences. *Marine Biotechnology*, **7**, 61–69.

Barrie JV, Conway KW (1999) Late Quaternary glaciation and postglacial stratigraphy of the northern Pacific margin of Canada. *Quaternary Research*, **51**, 113–123.

Belkhir K (2000) *GENETIX, logiciel sous Windows pour la génétique des populations*. Laboratoire Génomes et Populations, CNRS UPR 9060, Univeristé de Montpellier II, Montpelier, France.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate—a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological*, **57**, 289–300.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of nearctic and palearctic fishes. *Molecular Ecology*, **7**, 431–452.

Bickham JW, Wood CC, Patton JC (1995) Biogeographic implications of cytochrome *b* sequence and allozymes in sockeye (*Oncorhynchus nerka*). *Journal of Heredity*, **86**, 140–144.

Bigg GR, Cunningham CW, Ottersen G, Pogson GH, Wadley MR, Williamson P (2008) Ice-age survival of Atlantic cod: agreement between palaeoecology models and genetics. *Proceedings of the Royal Society of London. Series B*, **275**, 163–173.

Bilgin R (2007) Kgtests: a simple Excel Macro program to detect signatures of population expansion using microsatellites. *Molecular Ecology Notes*, **7**, 416–417.

Bond G, Broecker W, Johnsen S *et al.* (1993) Correlations between climate records from North Atlantic sediments and Greenland ice. *Nature*, **365**, 143–147.

Buonaccorsi VP, Weserman M, Stannard J, Kimbrell C, Lynn E, Vetter RD (2004) Molecular genetic structure suggests limited larval dispersal in grass rockfish, *Sebastes rastrelliger*. *Marine Biology*, **145**, 779–788.

Byun SA, Koop BF, Reimchen TE (1997) North American black bear mtDNA phylogeography: implications for morphology and the Haida Gwaii glacial refugium controversy. *Evolution*, **51**, 1647–1653.

Canino MF, Spies IB, Hauser L (2005) Development and characterization of novel di- and tetra-nucleotide microsatellite markers in Pacific cod (*Gadus macrocephalus*). *Molecular Ecology Notes*, **5**, 469–471.

Carr SM, Marshall HD (2008) Intraspecific phylogeographic genomics from multiple complete mtDNA Genomes in Atlantic cod (*Gadus morhua*): origins of the ''Codmother,'' transatlantic vicariance and midglacial population expansion. *Genetics*, **180**, 381–389.

Ciannelli L, Bailey KM (2005) Landscape dynamics and resulting species interactions: the cod-capelin system in the southeastern Bering Sea. *Marine Ecology Progress Series*, **291**, 227–236.

Clague JJ (1989) Quaternary geology of the Canadian Cordillera. In: *Quaternary Geology of Canada and Greenland*. (ed. Fulton RJ), pp. 15–96. Geology of Canada 1, Geological Survey of Canada, Ottawa. (Volume 1 also published as *Geological Society of America, Geology of North America*, **K-1**)

Clague JJ, James TS (2002) History and isostatic effects of the last ice sheet in southern British Columbia. *Quaternary Science Reviews*, **21**, 71–78.

Clark PU, Alley RB, Pollard D (1999) Northern Hemisphere ice-sheet influences on global climate change. *Science*, **286**, 1104–1111.

Coulson MW, Marshall HD, Pepin P, Carr SM (2006) Mitochondrial genomics of gadine fishes: implications for taxonomy and biogeographic origins from whole-genome data sets. *Genome*, **49**, 1115–1130.

Cunningham KM, Canino MF, Spies IB, Hauser L (2009) Genetic isolation by distance and localized fjord population structure in Pacific cod (*Gadus macrocephalus*): limited effective dispersal in the northeastern Pacific Ocean. *Canadian Journal of Fisheries and Aquatic Sciences*, **66**, 153–166.

Dansgaard W, Johnsen SJ, Clausen HB *et al.* (1993) Evidence for general instability of past climate from 250-kyr ice-core record. *Nature*, **325**, 318–321.

Deagle BE, Reimchen TE, Levin DB (1996) Origins of endemic stickleback from the Queen Charlotte Islands: mitochondrial and morphological evidence. *Canadian Journal of Zoology*, **74**, 1045–1056.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214. doi:10.1186/147-2148-7-214.

Excoffier LG, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchiking, and background selection. *Genetics*, **147**, 915–925.

Genbank http://www.ncbi.nlm.nih.gov/Genbank/ for haplotype 001 NADH—haplotype NADH 021 [Accession numbers EU729403–EU729423]; for isolate PS297 NADH-KO64 NADH [Accession numbers FJ877032–FJ877135]; for haplotype 001 cytochrome B—haplotype 035 cytochrome B [Accession numbers EU729368–EU729402]; for isolate PS297 cytochrome B—isolate KO56 cytochrome B [Accession numbers FJ876989–FJ877031]; for isolate 3K_NC_14_02—isolate Tr11 [Accession numbers: EU877710–EU877741]; for

*Gadus morhua* complete mitochondrial genome [Accession number AM489716]; accessed July 14, 2010.

Gong Y, Park YC, Kim SS (1991) Study of the management unit of fisheries resources by genetic method. 1. Genetic similarity of Pacific cod in the North Pacific. *Bulletin of National Fisheries Research and Development Institute*, **45**, 47–61.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html.

Grant WS, Utter FM (1984) Biochemical population genetics of Pacific herring (*Clupea pallasi*). *Canadian Journal of Fisheries and Aquatic Sciences*, **41**, 856–864.

Grant WS, Teel DJ, Kobayashi T, Schmitt C (1984) Biochemical population genetics of Pacific halibut (*Hippoglossus stenolepis)* and comparison with Atlantic halibut (*H. hippoglossus*). *Canadian Journal of Fisheries and Aquatic Sciences*, **41**, 1083–1088.

Grant WS, Zhang CI, Kobayashi T, Ståhl G (1987) Lack of genetic stock discretion in Pacific cod (*Gadus macrocephalus*). *Canadian Journal of Fisheries and Aquatic Sciences*, **44**, 490–498.

Grant WS, Spies IB, Canino MF (2006) Biogeographic evidence for selection on mitochondrial DNA in North Pacific walleye pollock *Theragra chalcogramma*. *Journal of Heredity*, **97**, 571–580.

Grosswald MG (1998) Late-Weichselian ice sheets in Arctic and Pacific Siberia. *Quaternary International*, **45–46**, 3–18.

Gustafson RG, Lenarz WH, McCain BB *et al.* (2000) Status review of Pacific Hake, Pacific cod, and walleye pollock from Puget Sound, Washington. *NOAA Technical Memorandum* NMFS-NWFSC-44, 275p.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Hardie DC, Gillett RM, Hutchings JA (2006) The effects of isolation and colonization history on the genetic structure of marine-relict populations of Atlantic cod (*Gadus morhua*) in the Canadian Arctic. *Canadian Journal of Fisheries and Aquatic Sciences*, **63**, 1830–1839.

Hardy OJ, Vekemans X (2002) SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Hardy OJ, Charbonnel N, Freville H, Heuertz M (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics*, **163**, 1467–1482.

Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Imbrie J, Hays JD, Martinson DG *et al.* (1984) The orbital theory of Pleistocene climate: support from a revised chronology of the marine $\delta^{18}O$ record. In: *Milankovitsh and Climate, Part 1* (eds Berger AL, Imbrie J, Hays JD, Kukla GJ, Saltzman B), pp. 269–305. Reidel Publishing, Boston, MA.

Kenchington EL, Harding GC, Jones MW, Prodöhl PA (2009) Pleistocene glaciation events shape genetic structure across the range of the American lobster, *Homarus americanus*. *Molecular Ecology*, **18**, 1654–1667.

Kim M, An HS, Choi KH (2010) Genetic characteristics of Pacific cod populations in Korea based on microsatellite markers. *Fisheries Science*, **74**, 595–603.

Kitamura A, Takano O, Takata H, Omote H (2001) Late Pliocene-early Pleistocene paleoceanographic evolution of the Sea of Japan. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **172**, 81–98.

Lambeck K, Esat TM, Potter EK (2002) Links between climate and sea levels for the past three million years. *Nature*, **419**, 199–206.

Liu JX, Gao TX, Wu SF, Zhang YP (2007) Pleistocene isolation in the Northwestern Pacific marginal seas and limited dispersal in a marine fish, Chelon haematocheilus (Temminck & Schlegel, 1845). *Molecular Ecology*, **16**, 275–288.

Mann DH, Hamilton TD (1995) Late Pleistocene and Holocene Paleoenvironments of the North Pacific Coast. *Quaternary Science Reviews*, **14**, 449–471.

Mann DH, Peteet DM (1994) Extent and timing of the last glacial maximum in Southwestern Alaska. *Quaternary Research*, **42**, 136–148.

McDonald D, Pedersen TF, Crusius J (1999) Multiple late Quaternary episodes of exceptional diatom production in the Gulf of Alaska. *Deep-Sea Research II*, **46**, 2993–3017.

Menounos B, Osborn G, Clague JJ, Luckman BH (2008) Latest Pleistocene and Holocene glacier fluctuations in western Canada. *Quaternary Science Reviews*, **28**, 2049–2074.

Mitrovica JX (2003) Recent controversies in predicting post-glacial sea-level change. *Quaternary Science Reviews*, **22**, 127–133.

Mueter FJ, Litzow MA (2008) Sea ice retreat alters the biogeography of the Bering Sea continental shelf. *Ecological Applications*, **18**, 309–320.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1–10.

Okazaki Y, Takahashi K, Katsuki K *et al.* (2005) Late Quaternary paleoceanographic changes in the southwestern Okhotsk Sea: evidence from geochemical, radiolarian, and diatom records. *Deep Sea Research II*, **52**, 2332–2350.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Orbacz EA, Gaffney PM (2000) Genetic structure of tautog (*Tautoga onitis*) population assayed by RFLP and DGGE analysis of mitochondrial and nuclear genes. *Fishery Bulletin, US*, **98**, 336–344.

O'Reilly P, Reimchen TE, Beech TE, Strobeck C (1993) Mitochondrial DNA in *Gasterosteus* and Pleistocene glacial refugium on the Queen Charlotte islands, British Columbia. *Evolution*, **47**, 678–684.

Ormseth OA, Norcross BL (2009) Causes and consequences of life-history variation in North American stocks of Pacific cod. *ICES Journal of Marine Science*, **66**, 349–357.

Paul AJ, Paul JM, Smith RL (1990) Consumption, growth and evacuation in the Pacific cod, *Gadus macrocephalus*. *Journal of Fish Biology*, **37**, 117–124.

Porter SC, Swanson TW (1998) Radiocarbon age constraints on rates of advance and retreat of the Puget lobe of the Cordilleran ice sheet during the last glaciation. *Quaternary Research*, **50**, 205–213.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Pritchard JK, Stephens M, Donnely P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Raymond M, Rousset F (1995) GENEPOP (Version 3); population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Reich DE, Goldstein DB (1998) Genetic evidence for a Paleolithic human population expansion in Africa. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 8119–8123.

Reich DE, Feldman MW, Goldstein DB (1999) Statistical properties of two tests that use multilocus data sets to detect population expansions. *Molecular Biology and Evolution*, **16**, 453–466.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Saitoh K (1998) Genetic variation and local differentiation in the Pacific cod *Gadus macrocephalus* around Japan revealed by mtDNA and RAPD markers. *Fisheries Science*, **64**, 673–679.

Sancetta C, Robinson SW (1983) Diatom evidence on Wisconsin and Holocene events in the Bering Sea. *Quaternary Research*, **20**, 232–245.

Sancetta C, Heusser L, Labeyrie L, Naidu AS, Robinsion SW (1985) Wisconsin-Holocene paleoenvironments of the Bering Sea: evidence from diatoms, pollen, oxygen isotopes, and clay minerals. *Marine Geology*, **62**, 55–68.

Seeb LW, Crane PA (1999) High genetic heterogeneity in chum salmon in western Alaska, the contact zone between northern and southern lineages. *Transactions of the American Fisheries Society*, **128**, 58–87.

Slatkin M (1995) A measure of population subdivision based upon microsatellite allele frequencies. *Genetics*, **139**, 457–462.

Stark JW (2007) Geographic and seasonal variations in maturation and growth of female Pacific cod (*Gadus macrocephalus*) in the Gulf of Alaska and Bering Sea. *Fishery Bulletin, US*, **105**, 396–407.

Takahashi K (1998) The Bering and Okhotsk seas: modern and past paleoceanographic changes and gateways impact. *Journal of Asian Earth Sciences*, **16**, 49–58.

Tanaka S, Takahashi K (2005) Late Quaternary paleoceanographic changes in the Bering Sea and the western subarctic Pacific based on radiolarian assemblages. *Deep-Sea Research II*, **52**, 2131–2149.

Vasemägi A (2006) The adaptive hypothesis of clinal variation revisited: single-locus clines as a result of spatially restricted gene flow. *Genetics*, **173**, 2411–2414.

Vucetich JA, Waite TA (2003) Spatial patterns of demography and genetic processes across the species' range: null hypotheses for landscape conservation genetics. *Conservation Genetics*, **4**, 639–645.

Warner BG, Mathewes RW, Clague JJ (1982) Ice-free conditions on the Queen Charlotte Islands, British Columbia, at the height of the Late Wisconsin glaciation. *Science*, **218**, 675–677.

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

---

Michael Canino is interested in using genetic markers in the management and conservation of marine species in the North Pacific. Ingrid Spies is a graduate student at the University of Washington studying quantitative models of genetic and population dynamics processes influencing the abundances of North Pacific fishes. Kathryn Cunningham recently graduated from the University of Washington and has pursued a career in education. Lorenz Hauser is a professor, who leads a research group that is focused on the use of molecular genetic markers to understand the dynamics of wild populations. Stewart Grant has a career-long interest in the understanding of historical and contemporary sources of population structure in North Pacific species.

---

## Supporting information

Additional supporting information may be found in the online version of this article.

**Table S1** Sample size ($N$), number of alleles ($N_A$), allelic richness ($N_S$) corrected to the smallest sample size ($N = 18$), expected ($H_e$) and observed ($H_o$) heterozygosities, and estimates of $F_{IS}$ for 11 microsatellite loci in Pacific cod

**Table S2** Frequency of Pacific cod mtDNA haplotypes. Sample abbreviations are as in Table 1

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.