

U.S. Fish and Wildlife Service

Genetic Population Structure of Alaska Eulachon

Alaska Fisheries Technical Report Number 106



U.S. Forest Service



Conservation Genetics Laboratory
November 2009



The Alaska Region Fisheries Program of the U.S. Fish and Wildlife Service conducts fisheries monitoring and population assessment studies throughout many areas of Alaska. Dedicated professional staff located in Anchorage, Juneau, Fairbanks, and Kenai Fish and Wildlife Offices and the Anchorage Conservation Genetics Laboratory serve as the core of the Program's fisheries management study efforts. Administrative and technical support is provided by staff in the Anchorage Regional Office. Our program works closely with the Alaska Department of Fish and Game and other partners to conserve and restore Alaska's fish populations and aquatic habitats. Additional information about the Fisheries Program and work conducted by our field offices can be obtained at:

<http://alaska.fws.gov/fisheries/index.htm>

The Alaska Region Fisheries Program reports its study findings through two regional publication series. The **Alaska Fisheries Data Series** was established to provide timely dissemination of data to local managers and for inclusion in agency databases. The **Alaska Fisheries Technical Reports** publishes scientific findings from single and multi-year studies that have undergone more extensive peer review and statistical testing. Additionally, some study results are published in a variety of professional fisheries journals.

Disclaimer: The use of trade names of commercial products in this report does not constitute endorsement or recommendation for use by the federal government.

Genetic Population Structure of Alaska Eulachon

Blair G. Flannery¹, John K. Wenburg¹, Cara J. Lewis¹,
Brenda L. Norcross², and Robert E. Spangler³

Abstract

Genetic variation was assayed at 14 microsatellite loci to investigate the genetic population structure of Alaska eulachon. Alaska eulachon have high levels of genetic diversity and relatively large effective population sizes (N_e), although a large variance in reproductive success is likely responsible for a low N_e/N ratio. Eulachon exhibit a low degree of genetic divergence ($G_{ST} = 0.005$) that is structured by broad-scale geographic regions. Overall, there is a significant correlation between genetic and geographic distance, suggesting that gene flow is geographically restricted and follows an isolation-by-distance (IBD) model, with geographic distance explaining 28% of the genetic variation. However, closer analysis reveals an absence of IBD within regions and that gene flow is primarily restricted by geographic distance between regions. The demographic independence of the regions warrants separate management regimes. However, it may be prudent to take a precautionary approach and conserve potential spawning habitat throughout the geographical range of eulachon because of high gene flow and variable use of rivers for spawning.

Introduction

Estimation of genetic population structure and diversity is important for defining meaningful conservation units and assessing genetic health. Physical or behavioral barriers can result in intraspecific reproductive isolation, allowing for the accumulation of genetic differences through genetic drift and the formation of genetic population structure. The diverging effects of genetic drift are constrained by limited time of separation, gene flow, and large effective population sizes (N_e). These factors may be significant for marine species because of potential recent colonization following Pleistocene deglaciation, the absence of obvious barriers to dispersal, and large census sizes, making detection of genetic divergence difficult even when suggested by differences in life history, morphology, and parasites (Hauser and Carvalho 2008). Whereas morphology and life history differences may represent phenotypic plasticity and not local adaptation, the observation of genetic population structure indicates demographic independence, which must be accounted for in fishery management (Hauser and Carvalho 2008). Long term sustained yield, ultimately the goal of fishery management, can only be accomplished through conservation of genetic resources to maintain diversity and a

1 U.S. Fish and Wildlife Service, Conservation Genetics Laboratory, 1011 E. Tudor Rd. Anchorage, AK 99503.

2 Institute of Marine Science, School of Fisheries and Ocean Sciences, University of Alaska Fairbanks, P.O. Box 75722, Fairbanks, AK 99775.

3 U.S. Forest Service, 3301 C Street, Suite 300, Anchorage, AK 99503.

population's adaptive potential in the face of a fluctuating environment (Altukhov and Salmenkova 1987; Nelson and Soule 1987).

Recent conservation concerns have sparked research on eulachon (*Thaleichthys pacificus*) also known as candlefish due to a high oil content that enables a dried fish to burn when lit. Eulachon are an anadromous, forage fish from the family Osmeridae, with a geographic distribution from California to the Pribilof Islands in the Bering Sea. They are highly fecund, semelparous spawners (Moffitt 2002). Females broadcast lay an average of 35,000 eggs, while males release milt simultaneously. Eggs are fertilized in the water column, attach to river substrate, and hatch in 20 to 40 days. Larvae are immediately flushed to sea where they are dispersed by estuarine and ocean currents. After three to five years in the ocean, eulachon return to rivers to spawn, usually in the lower tidally influenced reaches. The extent that eulachon home to natal spawning sites is unknown (Hay and McCarter 2000). In the Pacific Northwest, eulachon spawning run strength and use of rivers for spawning are variable (Hay and McCarter 2000). This variability has been observed in Alaska as well. In Behm Canal, the strongest runs of fish have been reported to vary between the Eulachon, Unuk, and Klahini Rivers (Tisler, USFS, personal communication). Similar events are known to occur in adjacent rivers of the Copper River Delta (Moffitt, ADF&G, personal communication), the Yakutat Forelands (Gillikin, USFS, personal communication), upper Lynn Canal (Bachman, ADF&G, personal communication), and Berners Bay (Koski, NOAA, personal communication; USFS unpublished data).

Eulachon have long been an important food resource. Historically, Native Americans prized the oil that they rendered from the eulachon, which remained a solid at room temperature. This fat was widely traded through a network of "grease trails" between coastal and inland tribes. In recent times, in the Pacific Northwest, eulachon were caught in vast quantities in both subsistence and commercial fisheries, with commercial hauls often exceeding 1,000 metric tons a year from the Columbia River (NOAA 2009). This occurred until the early 1990's when eulachon abundance collapsed, leading to the proposal to list the southern distinct population segment of eulachon as threatened under the U.S. Endangered Species Act (ESA; NOAA 2009). In Alaska, eulachon have not been similarly exploited, though they are a popular subsistence and personal use fishery (Spangler et al. 2003; Joyce et al. 2004), and while commercial fisheries are currently limited, the collapse of eulachon in the Pacific Northwest has prompted interest (Moffitt 2002). An ESA ruling has not been proposed for Alaska eulachon, whose biomass has actually increased (Ormseth et al. 2008), but the collapse of the Behm Canal eulachon run (Ormseth et al. 2008; USFS, unpublished data) illustrates that Alaska eulachon are not immune to local perturbations. Though the cause of the Behm Canal crash is not clear (Ormseth et al. 2008), it may be a cyclical pattern of use by eulachon and not reflective of a stock collapse.

The population structure for eulachon within and among river systems in Alaska is unknown. In fact, relatively little is known about the life history of eulachon. This lack of knowledge combined with the variable spawning run strength and use of rivers for spawning complicates fishery management and assessment of stock collapses. Initial genetic studies of eulachon from the Pacific Northwest observed little population structure, which suggested that eulachon existed as large populations with low levels of genetic diversity (McLean et al.

1999; McLean and Taylor 2001). However, a more recent and thorough study of eulachon from this area observed that eulachon exhibited much higher levels of genetic diversity than originally thought and that significant differences occurred among eulachon from different rivers (Beacham et al. 2005). These contradictory results make it clear that an understanding of eulachon genetic population structure in Alaska is necessary to identify appropriate management units for maintenance of biodiversity and productivity. In this study, we investigate the genetic population structure of eulachon by assaying variation at 14 microsatellite loci to evaluate patterns of genetic diversity within and among 26 collections of eulachon distributed throughout Alaska.

Methods

Sample collection and laboratory analysis

Tissue samples from adult eulachon were collected from 26 freshwater spawning locations in Southcentral and Southeast Alaska (Table 1; Figure 1). Total genomic DNA was extracted from the tissue (~25mg) using proteinase K with the Dneasy™ DNA isolation kit (Qiagen Inc., Valencia, CA), quantified by fluorometry, and diluted to 30 ng/μl. Polymerase chain reaction (PCR) DNA amplification was used to assay genetic variation at the following microsatellite loci: *Tpa103*, *Tpa104*, *Tpa111*, *Tpa112*, *Tpa113*, *Tpa114*, *Tpa115*, *Tpa117*, *Tpa118*, *Tpa119*, *Tpa121*, *Tpa122*, *Tpa127*, *Tpa129* (Kaukinen et al. 2004). The PCR product was electrophoresed and visualized with the Applied Biosystems 3730 Genetic Analyzer utilizing a polymer denaturing capillary system. The sizes of bands were estimated and scored by the computer program GENEMAPPER® version 4.0. Applied Biosystems GeneScan™-600 LIZ® size standards, 20–600 bases, were loaded in all lanes to ensure consistency of allele scores. All scores were verified manually. Alleles were scored by two independent researchers, with any discrepancies being resolved by re-running the samples in question and repeating the double scoring process until scores matched. In addition, one column of samples from each DNA plate, representing a minimum of eight percent of the samples, was re-run for the full suite of loci following the above methods and compared to the original scores in order to check for plate transpositional errors.

Data analysis

Intra-population genetic diversity—Where multiple tests of the same hypothesis were performed, a sequential Bonferroni method was used to maintain the overall alpha at 0.05 (Rice 1989). The data were checked for duplicated genotypes using the program MICROSATELLITE TOOLKIT (Park 2001), and any duplicates were removed. The stocks and loci were assessed for conformance to Hardy-Weinberg and gametic phase equilibrium using the programs FSTAT 2.9.3 (Goudet 2001) and GENETIX 4.05 (Belkhir et al. 1996), respectively. This was done to test that the samples represented randomly mating, Mendelian populations. Significant tests of disequilibrium ($P < 0.05$) were compared to binomial expectations to determine if chance alone explained the results (Apostol et al. 1996). In addition, FSTAT 2.9.3 and GENETIX 4.05 were used to calculate estimates of allelic richness, percentage polymorphic loci (95%), observed and expected heterozygosity, and gene differentiation (G_{ST}). Estimates of effective number of alleles (Hartl and Clark 1997) for stocks and loci were calculated in Microsoft Excel™. A Mann-Whitney test (Zar 1984) was used to test for significant differ-

ences ($P < 0.05$) in observed levels of genetic diversity between regional groups. Contemporary estimates of effective population size (N_e) were calculated using the linkage disequilibrium method (Hill 1981; Bartley et al. 1992) in the program NEESTIMATOR 1.3 (Peel et al. 2004). Long-term estimates of N_e were calculated using Bayesian analysis in the program MIGRATE 3.0 (Beerli and Felsenstein 2001; Beerli 2006). Data for collections were pooled within regions for the N_e analysis.

Inter-population genetic diversity—Using PHYLIP 3.57 (Felsenstein 1989), replicate collection pairwise chord distance (Cavalli-Sforza and Edwards 1967) matrices were calculated from allele frequencies by bootstrapping over loci 1,000 times, wherefrom a consensus neighbor-joining (Saitou and Nei 1987) dendrogram was produced. To further investigate spatial relationships and assess whether gene flow was geographically restricted (i.e. isolation-by-distance), collection pairwise matrices of genetic distance (F_{ST}) and geographic distance (kilometers) were analyzed by standard linear regression and lowess smoothing. Significance of the correlation between the two matrices was determined by the Mantel test (Mantel 1967) with 10,000 randomizations using FSTAT 2.9.3.

GENEPOP 3.4 (Raymond and Rousset 1995) was used to conduct collection pairwise tests of allelic frequency homogeneity. Hierarchical likelihood ratio tests were conducted to determine the homogeneity of allelic frequencies among collections within areas, among areas, and between regions (G -test, Sokal and Rohlf 1995). The hierarchy used in these analyses was based on geographic location and population structure as depicted by neighbor-joining analysis. Alleles with expected overall counts of less than three were pooled with adjacent alleles to maintain asymptotic assumptions (Sokal and Rohlf 1995). The magnitude of heterogeneity within and between regions was compared using an approximate F -statistic (Smouse and Ward 1978).

A hierarchical unbiased gene diversity analysis was used to partition the genetic variation due to divergence among collections (Chakraborty and Leimar 1987; Nei and Chesser 1983). The relative proportions of the total diversity accounted for by the different levels of hierarchy were estimated by calculating coefficients of gene differentiation (G_{ST} -statistics). Significance of the G_{ST} -statistics was inferred from likelihood ratio tests of homogeneity (Chakraborty and Leimar, 1987). Gene flow or the effective number of migrants ($N_e m$) was estimated from G_{ST} -statistics assuming a hierarchical island model at equilibrium (Zhivotovsky et al. 1994). In addition, $N_e m$ between regions was estimated using MIGRATE 3.0.

Results

Intra-population genetic diversity

Significant ($P < 0.05$) Hardy-Weinberg disequilibrium was observed for *Tpa119* and *Tpa121*; therefore, these loci were dropped from further analyses of allele frequencies. All loci were in gametic phase equilibrium. Genetic diversity values varied for the loci, with the number of alleles ranging from 16–70, allelic richness from 4.0–13.5, effective number of alleles from 1.6–13.8, expected heterozygosity from 0.383–0.938, observed heterozygosity from 0.384–0.901, and G_{ST} from 0.002–0.011 (Table 2). Genetic diversity values for the collections ranged from 7.0–8.5 for allelic richness, 4.4–6.2 for effective number of alleles, 0.700–0.772

for expected heterozygosity, and 0.694–0.793 for observed heterozygosity (Table 3). Genetic diversity values were averaged over the collections by region and ranged from 7.6 (northern) to 7.9 (southern) for allelic richness, 5.4 (northern) to 5.7 (southern) for effective number of alleles, 0.737 (northern) to 0.757 (southern) for expected heterozygosity, and 0.727 (northern) to 0.735 (southern) for observed heterozygosity. Collections from the southern region had significantly ($P < 0.05$) higher levels of allelic richness and expected heterozygosity, but there were no significant differences between the regions for effective number of alleles and observed heterozygosity. Contemporary estimates of effective population size (N_e) and their 95% confidence intervals were 3525 (2578–5498) for the northern region and 2823 (2392–3427) for the southern region. Long-term estimates of N_e and their 95% credible intervals were 2458 (2400–2488) for the northern region and 2460 (2400–2488) for the southern region.

Inter-population genetic diversity

Neighbor-joining analysis revealed a consensus dendrogram with little bootstrap support; only collections from the Cook Inlet clustered together a majority of the time (Figure 2). The consensus dendrogram indicated a weak population structure aligned along broad geographic regions. Collections from Cook Inlet, Prince William Sound, and the Yakutat Forelands formed a northern region, while collections from upper Lynn Canal, Berners Bay, Stikine Strait, and Behm Canal formed a southern region. With the exception of Cook Inlet, little geographically based fine-scale genetic population structure was evident within the regions. Gene flow was geographically restricted overall ($r = 0.527$, $P < 0.0001$, Figure 3a) and between regions ($r = 0.333$, $P < 0.0001$, Figure 3b), but not within the northern ($r = 0.198$, $P > 0.05$, Figures 3c) or southern regions ($r = 0.044$, $P > 0.05$, Figures 3d). The significant IBD suggests that collections between regions are at or near migration-drift equilibrium, but equilibrium has not been approached within regions (Slatkin 1993).

Significant genetic divergence was observed in 59 of the 325 collection pairwise tests of allelic frequency homogeneity (Table 4). Only one of the significant pairwise tests was observed within regions, between Skagway and Upper Landing Slough (Table 4). The hierarchical analysis of allelic frequency homogeneity revealed highly significant genetic divergence between regions, but within regions genetic divergence was limited (Table 5). No genetic divergence was observed among collections within areas or among areas within regions (Table 5). However, overall additive effects resulted in significant genetic divergence for within southern areas, total northern region, and total within regions (Table 5). Levels of genetic divergence observed between regions were four times greater than within regions ($F_{18,432} = 4.2$, $P < 0.0001$). The northern and southern regions did not exhibit different levels of genetic divergence ($F_{198,234} = 1.1$, $P > 0.05$).

A total of 99.5% ($H_S/H_T=0.9950$) of the gene diversity was present in each collection (Table 6). Variation among collections accounted for only 0.5% ($G_{ST}=0.0050$) of the gene diversity, which, though low, was significant ($P < 0.05$). Variation between regions was significant ($P < 0.05$) as well and accounted for 0.24% ($G_{RT}=0.0024$) of the among collections gene diversity. Variation among areas within regions and among collections within areas accounted for 0.08% ($G_{LR}=0.0008$) and 0.18% ($G_{SL}=0.0018$) of the among collections gene diversity, respectively; however, these values were not significantly different ($P > 0.05$) from zero.

The overall G_{ST} estimate of $N_e m$ was 49 and ranged from 26 between regions to 156 among areas within regions. MIGRATE 3.0 estimates of $N_e m$ were 17 from the northern region to the southern region and 24 from the southern region to the northern region, for a total between regions $N_e m$ of 41.

Discussion

Intra-population genetic diversity

The analysis of eulachon genetic variation revealed that levels of heterozygosity (H_E , H_O) and number of alleles (A_R , A_E) are similar among the collections and regions (Table 3). Although the southern region has significantly higher levels of expected heterozygosity and allelic richness, the differences are slight and other measures of genetic diversity are not supportive of a latitudinal trend, which agrees with observations by McLean and Taylor (2001). These levels are high and signify strong evolutionary potential (Frankham et al. 2002). Moreover, the observed levels of genetic diversity (Table 2) are similar to those reported by Beacham et al. (2005), and further confirm that eulachon are not depauperate in microsatellite genetic diversity as suggested by McLean and Taylor (2001).

Mutation rates can vary widely among loci (Olsen et al. 2003); therefore, the dearth of microsatellite genetic diversity observed by McLean and Taylor (2001) is likely specific to those loci, as they postulate, and not indicative of the species, especially in light of the high levels of mitochondrial DNA (mtDNA) genetic diversity observed in eulachon (McLean et al. 1999). Because it is haploid and maternally inherited, mtDNA has an effective population size (N_e) that is one-fourth that of the nuclear genome, which makes it more vulnerable to loss of genetic diversity. The high levels of observed microsatellite genetic diversity from this study and Beacham et al. (2005) concur with mtDNA results from McLean et al. (1999), indicating that eulachon harbor sufficient genetic diversity to maintain potential for adaptive response.

The genetic health of a population can also be measured by estimating N_e , which is the size that the population behaves genetically in regards to loss of genetic diversity and is usually much less than census size. Populations are considered at risk of short- and long-term loss of genetic diversity if their N_e is below 50 and 500, respectively (Haller 2003a). The northern and southern regions both exceed these critical N_e values for loss of genetic diversity. However, the N_e estimates are not as high as expected when compared to census size (Nunney and Elam 1994; Nunney 1996) for which there is little data. Abundance for Copper River, based on total biomass and mean weight per fish, ranges from 43 million to 150 million fish (Moffitt et al. 2002). Using this number alone puts the N_e/N ratio at approximately 10^{-5} , which is quite low (Frankham 1995), but comparable to other highly fecund marine species (Hauser and Carvalho 2008).

Likely reasons for the low N_e/N ratio include unequal sex ratio, variance in reproductive success among individuals and locations, hierarchical population structure, and fluctuating population size (Nunney 1999; Haller 2003b). The agreement between the long-term and contemporary estimates of N_e suggests that fluctuating population size is not largely responsible for the low N_e/N ratio (Turner et al. 2002), and the recent population fluctuations that

have occurred in Behm Canal eulachon only reduce N_e moderately (Arnason 2004). Inbreeding due to hierarchical population structure can reduce N_e , but an analysis of the G_{ST} (0.005) and G_{IS} (0.023) values in an island-structured metapopulation model, using equation 16 from Nunney (1999), indicates that its affect is negligible. Unequal sex ratios may play a role; the mean annual percentage of eulachon males was 67% from 1998 to 2002 in the Copper River and was significantly different from 50:50 in all four years (Moffitt et al. 2002). In another river from the northern region, sex ratios were also skewed towards males (2:1; Spangler et al. 2003). This moderate inequality of sex ratios alone can not explain the low N_e/N ratio (Hallerman 2003b). Thus, the driving influence is likely a large variance in reproductive success among individuals and locations (Turner et al. 2002; Hallerman 2003b; Arnason 2004), which is not uncommon for a marine species with high fecundity and mortality (Type III survivorship). Eulachon must spawn at times that coincide with optimal river and ocean conditions for successful reproduction and recruitment. Variable environmental conditions are known to strongly affect recruitment in estuarine dependent species, leading to reproductive variances orders of magnitude greater than binomial expectations and resulting in low N_e/N ratios (Turner et al. 2002; Hallerman 2003b).

Inter-population genetic diversity

In Alaska, eulachon exhibit a low degree ($G_{ST} = 0.005$) of broad geographic scale genetic population structure. This structure is largely explained by two regional groups, with collections from the Yakutat Forelands, Prince William Sound, and Cook Inlet forming a northern region and collections from upper Lynn Canal, Berners Bay, Stikine Strait, and Behm Canal forming a southern region. The regions are similarly structured, without any difference in levels of divergence, whereas the level of divergence between regions is four times greater. Overall, there is a significant correlation between genetic and geographic distance, suggesting that gene flow is geographically restricted and follows an isolation-by-distance (IBD) model (Figure 3a), with geographic distance explaining 28% of the genetic variation. However, closer analysis reveals an absence of IBD within regions (Figure 3c, d) and that gene flow is primarily restricted by geographic distance between regions (Figure 3b). This results from a lack of heterogeneity within regions and suggests that while there is a barrier to gene flow between regions, no such impediment exists within regions. The counter clockwise Alaska gyre may facilitate larval dispersal from the Yakutat Forelands towards Prince William Sound and Cook Inlet and, therefore, gene flow among these three areas, whereas the Alexander Archipelago likely shelters Southeast Alaska eulachon from oceanic currents, restricting gene flow between regions, and allowing for the development of regional genetic divergence. However, the barrier to gene flow between regions is not that strong as the lack of bootstrap support and the high $N_e m$ estimates between regions indicate.

These results agree and contrast with observations of genetic variation collected primarily from Pacific Northwest eulachon (McLean et al. 1999; McLean and Taylor 2001; Beacham et al. 2005). All of the studies rejected the notion of panmixia, with regional structure observed by Beacham et al. (2005) and the present study, but only Beacham et al. (2005) observed fine-scale genetic population structure. Additionally, overall IBD was observed, except by McLean and Taylor (2001), who hypothesized that the lack of IBD for microsatellite variation when it was observed for mtDNA variation (McLean et al. 1999) was because of low

sample sizes and the one-fourth smaller N_e of mtDNA making it more sensitive to population structure. The differing mutational rates of loci do not allow for direct comparisons of F_{ST} or analogs (Olsen et al. 2004), but, with the same suite of loci, the degrees of population structure are identical between the present study and Beacham et al. (2005).

There are a number of potential reasons, acting separately or in combination, for the observation of fine-scale genetic population structure in eulachon from the Pacific Northwest (Beacham et al. 2005) and the absence of such structure in eulachon from Alaska. Eulachon likely survived Pleistocene glaciations in the Pacific refuge of which the Columbia River was the major freshwater habitat (McPhail and Lindsey 1970). Eulachon are not believed to have resided in the Bering Sea refuge because of their absence in Russia and the Far East (McPhail and Lindsey 1986). An mtDNA study supports the single Pacific refuge for eulachon (McLean et al. 1999), and the shallow mtDNA gene genealogy indicates a recent age of extant populations (Arnason 2004). Following deglaciation, eulachon likely underwent a northward range expansion, recolonizing newly available freshwater habitat. Therefore, insufficient time of separation, in comparison to the older southern populations, may explain the lack of divergence among Alaska eulachon collections within a region, which is suggested by the lack of within region migration-drift equilibrium, although the significant within region overall additive hierarchical effects and clustering of Cook Inlet collections may indicate emerging population structure. Augmenting or perhaps primarily responsible for this dearth of divergence could be higher rates of gene flow. Longer pelagic larval durations (PLD) are noted for northern latitude fish (Hauser and Carvalho 2008), and this has resulted in significantly reduced levels of genetic divergence in a color form of *Elacatinus evelynae* when compared to other color forms with shorter PLD (Taylor 2004). Alternatively, non-random sampling due to reproductive variance (sweepstakes effect) and statistical power may contribute to the perceived discrepancy. Allele frequencies can vary among cohort and location by random chance through the sweepstakes effect (Hedgecock 1994), resulting in genetic divergence that is not geographically influenced. With large sample sizes, even small deviations from random sampling can lead to significant results (Waples 1998). A lack of temporal stability was observed in the Beacham et al. (2005) study, which is supportive of a possible sweepstakes effect; however, regional divergence was three to six times greater than divergence among sampling years, indicative of a stable regional structure. Clearly, further genetic analyses across sampling years are required to vet the presence or absence of temporally stable, geographically based fine-scale genetic population structure in eulachon.

In general, eulachon appear to be less influenced by genetic drift than by gene flow. Such a relationship is not unexpected for a species whose larvae have a limited freshwater existence and inadequately developed nervous systems when compared to Pacific salmon (Hay and McCarter 2000). This implies that imprinting to natal river may not be strong and that homing may be imprecise. Lack of freshwater imprinting has led to greater straying rates in pink salmon than in sockeye salmon (Quinn 1984). Evidence of homing imprecision is seen in the great variability in eulachon abundance and use of specific rivers for spawning. Many rivers do not have annual eulachon spawning runs, and, within an estuary, the majority of the spawning biomass has been observed to move among rivers, suggesting that eulachon may home only to estuarine systems (i.e., Cook Inlet, Prince William Sound, and Southeast Alaska; Spangler, unpublished data). The larval stage has a longer duration in estuaries, al-

lowing for the development of sensory organs and for the potential of imprinting, which, if it does occur, would likely be less precise than that of Pacific salmon (Hay and McCarter 2000).

Despite the indications of high gene flow, a degree of reproductive isolation and, therefore, demographic independence is supported by the observed regional divergence. Gene flow is restricted between regions either by distance through regional homing, as suggested by the IBD results, or by a biogeographical or oceanographic barrier (e.g., Alaska gyre, Alexander Archipelago). At the minimum, these two regions should be managed as separate units to maintain both productivity and evolutionary potential of eulachon. Further population structure may be developing within regions, but additional analyses of samples across years are necessary to rule out sweepstakes effect divergence and clarify the microevolutionary processes. Until such clarification, a precautionary approach to the management of eulachon would be wise when so little is known about their life history and ecology. In concurrence with McLean et al. (1999), it would be appropriate to at least conserve potential spawning habitat within the geographic distribution of eulachon because of high gene flow and variable use of rivers.

Conclusions

- 1) Alaska eulachon have high levels of genetic diversity and strong evolutionary potential.
- 2) A large variance in reproductive success is likely responsible for the low N_e/N ratio.
- 3) Alaska eulachon exhibit a low degree ($G_{ST} = 0.005$) of broad-scale, regionally based genetic population structure.
- 4) The northern and southern regions are demographically independent and should be managed separately.
- 5) Further structure may be emerging within regions, but additional analyses of samples across years are necessary to confirm.

Recommendation

- 6) Analyze additional samples collected across multiple years.

Acknowledgements

The U. S. Fish and Wildlife Service (USFWS) laboratory received \$68,259 in partial funding support from the USFWS, Office of Subsistence Management, through the Fisheries Resource Monitoring Program, project 07-610. We thank Andrea Medeiros, Publications Specialist, with the Office of Subsistence Management for editing and formatting this report.

References

- Altukhov, Y. P., and E. A. Salmenkova. 1987. Stock transfer relative to natural organization, management, and conservation of fish populations. Pages 333–343 *in* N. Ryman and F.

- Utter, editors. Population genetics and fishery management. University of Washington Press, Seattle, Washington.
- Apostal, B.L., W.C. Black IV, P. Reiter, and B.R. Miller. 1996. Stock genetics with RAPD-PCR markers: the breeding structure of *Aedes aegypti* in Puerto Rico. *Heredity* 76:325–334.
- Arnason, E. 2004. Mitochondrial cytochrome *b* DNA variation in the high-fecundity Atlantic cod: Trans-Atlantic clines and shallow gene genealogy. *Genetics* 166: 1871–1885.
- Bartley, D., M. Bagley, G. Gall, and B. Bentley. 1992. Use of linkage disequilibrium data to estimate effective size of hatchery and natural fish populations. *Conservation Biology* 6:365–375.
- Beacham, T. D., D. E. Hay, D. L. Khai. 2005. Population structure and stock identification of eulachon (*Thaleichthys pacificus*), an anadromous smelt, in the Pacific Northwest. *Marine Biotechnology* 7:363–372.
- Beerli, P. 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22:341–345.
- Beerli, P., and J. Felsenstein. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America* 98:4563–4568.
- Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Catch. 2004. GENETIX 4.05, software under Windows™ for the genetics of populations. University of Montpellier, Montpellier, France.
- Cavalli-Sforza, L. L., and A. W. F. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution* 21:550–570.
- Chakraborty, R., and O. Leimar. 1987. Genetic variation within a subdivided population. Pages 89–120 in N. Ryman and F. Utter, editors. Population genetics and fishery management. University of Washington Press, Seattle.
- Felsenstein, J. 1989. Phylogeny inference package (version 3.2). *Cladistics* 164:164–167.
- Frankham, R. 1995. Effective population size/adult population ratios in wildlife: a review. *Genetical Research* 66:95–107.
- Frankham, R., J. D. Ballou, and D. A. Briscoe. 2002. Introduction to Conservation Genetics. Cambridge University Press, Cambridge, England.
- Goudet, J. 1995. FSTAT a program of IBM PC compatibles to calculate Weir and Cockerham's (1984) estimators of *F*-statistics (version 1.2). *Journal of Heredity* 86:485–486.
- Hallerman, E. 2003a. Population viability analysis. Pages 403–417 in E. M. Hallerman, editor. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland.
- Hallerman, E. 2003b. Random genetic drift. Pages 197–214 in E. M. Hallerman, editor. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland.

- Hartl, D. L., and A. G. Clark. 1997. Principles of population genetics. Sinauer Associates, Sunderland, Massachusetts.
- Hauser, L., and G. R. Carvalho. 2008. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries* 9:333–362.
- Hay, D. E., and P. B. McCarter. 2000. Status of the eulachon *Thaleichthys pacificus* in Canada. Department of Fisheries and Oceans Canada, Canadian Stock Assessment Secretariat, Research Document 2000/145, Ottawa.
- Hedgcock, D. 1994. Does variance in reproductive success limit effective population size of marine organisms? Pages 122–134 in A. Beaumont, editor. Genetics and evolution of aquatic organisms. Chapman and Hall, London.
- Hill, W. G. 1981. Estimation of effective population size from data on linkage disequilibrium. *Genetical Research* 38:209–216.
- Joyce, T. L., M. B. Lambert, and S. Moffitt. 2004. Eulachon subsistence harvest opportunities. U.S. Fish and Wildlife Service, Office of Subsistence Management, Fisheries Resource Monitoring Program, Final Report No. 02-075, Anchorage, Alaska.
- Kaukinen, K. H., K. J. Supernault, K. M. Miller. 2004. Development of microsatellite loci in Eulachon (*Thaleichthys pacificus*). *Molecular Ecology Notes* 7:632–634.
- Mantel, N. A. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27:209–220.
- McLean, J. E., D. E. Hay, and E. B. Taylor. 1999. Marine population structure in an anadromous fish: life-history influences patterns of mitochondrial DNA variation in the eulachon, *Thaleithys pacificus*. *Molecular Ecology* 8:143–158.
- McLean, J. E., and E. B. Taylor. 2001. Resolution of population structure in a species with high gene flow: microsatellite variation in the eulachon (Osmeridae: *Thaleithys pacificus*). *Marine Biology* 139:411–420.
- McPhail, J. D., and C. C. Lindsey. 1970. Freshwater fishes of North-western Canada and Alaska. *Fisheries Research Board of Canada Bulletin* 173: 1–381.
- McPhail, J. D., and C. C. Lindsey. 1986. Zoogeography of the freshwater fishes of Cascadia (the Columbia system and rivers north to the Stikine). Pages 615–637 in C. H. Hocutt and E. O. Wiley, editors. The zoogeography of North American freshwater fishes. Wiley, Chichester, UK.
- Moffitt, S., B. Marston, and M. Miller. 2002. Summary of eulachon research in the Copper River Delta, 1998–2002. Alaska Department of Fish and Game, Regional Information Report No. 2A02-34, Anchorage.
- Nei, M., and R. K. Chesser. 1983. Estimation of fixation indices and gene diversities. *Annals of Human Genetics* 47:253–259.
- Nelson, K., and M. Soule. 1987. Genetical conservation of exploited fishes. Pages 345–368 in N. Ryman and F. Utter, editors. Population genetics and fishery management. University of Washington Press, Seattle, Washington.

- NOAA (National Oceanic and Atmospheric Administration). 2009. Endangered and threatened wildlife and plants: proposed threatened status for southern distinct population segment of eulachon. Federal Register 74:10857–10876.
- Nunney, L. 1996. The influence of variation in female fecundity on effective population size. *Biological Journal of the Linnean Society* 59:411–425.
- Nunney, L. 1999. The effective size of a hierarchically structured population. *Evolution*: 53:1–10.
- Nunney, L., and D. R. Elam. 1994. Estimating the effective population size of conserved populations. *Conservation Biology* 8:175–184.
- Olsen, J. B., C. Habicht, J. Reynolds, and J. E. Seeb. 2004. Moderately and highly polymorphic microsatellites provide discordant estimates of stock divergence in sockeye salmon, *Oncorhynchus nerka*. *Environmental Biology of Fishes* 69:21–36.
- Olsen, J. B., S. J. Miller, W. J. Spearman, and J. K. Wenburg. 2003. Patterns of intra- and inter-population genetic diversity in Alaskan coho salmon: implications for conservation. *Conservation Genetics* 4:557–569.
- Ormseth, O. A., L. Conners, M. Guttormsen, and J. Vollenweider. 2008. Forage fishes in the Gulf of Alaska. National Marine Fisheries Service, Alaska Fisheries Science Center. Available at: <http://www.afsc.noaa.gov/REFM/docs/2008/GOAforage.pdf>
- Park, S. D. E. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection. Doctoral dissertation. University of Dublin, Ireland.
- Peel, D., J. R. Ovenden, and S. L. Peel. 2004. NeEstimator: software for estimating effective population size, Version 1.3. Queensland Government, Department of Primary Industries and Fisheries.
- Quinn, T. P. 1984. Homing and straying in Pacific salmon. Pages 357–362 in J. D. McCleave, G. P. Arnold, J. J. Dodson, and W. H. Neill, editors. *Mechanisms of migration in fishes*. Plenum Press, New York.
- Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2) population genetic software for exact tests and ecumenicism. *Journal of Heredity* 86: 248–249.
- Rice, W. R. 1989. Analyzing tables of statistical tests. *Evolution* 43:223–225.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406–425.
- Smouse, P. E., and R. H. Ward. 1978. A comparison of the genetic infrastructure of the Ye'cuana and the Yanomama: a likelihood analysis of genotypic variation among populations. *Genetics* 88:611–631.
- Sokal, R. R., and F. J. Rohlf. 1995. *Biometry*. W.H. Freeman. New York.
- Spangler, E. K., R. E. Spangler, and B. L. Norcross. 2003. Eulachon subsistence use and ecology investigations. U.S. Fish and Wildlife Service, Office of Subsistence Management, Fisheries Resource Monitoring Program, Final Report No. 00-041, Anchorage, Alaska.

- Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47:264–279.
- Taylor, M. S. 2004. Geography, coloration, and speciation in a genus of neotropical reef fishes (Gobiidae: *Elacatinus*). Doctoral dissertation. Louisiana State University, Baton Rouge.
- Turner, T. F., J. P. Wares, and J. R. Gold. 2002. Genetic effective size is three orders of magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). *Genetics* 162:1329–1339.
- Waples, R. S. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Heredity* 89: 438–450.
- Zar, J. H. 1984. *Biostatistical Analysis*, 2nd edition. Prentice Hall, Englewood Cliffs, New Jersey.
- Zhivotovsky, L. A., A. J. Gharrett, A. J. McGregor, M. K. Glubokovsky, and M. W. Feldman. 1994. Gene differentiation in Pacific salmon (*Oncorhynchus sp.*): facts and models with reference to pink salmon (*O. gorbuscha*). *Canadian Journal of Fisheries and Aquatic Sciences* 51(Supplement 1):223–232.

Table 1. Geographic region, area, collection location, collection label, year collected, and number (N) of samples.

Region	Area	Collection	Label	Year	N	
North	Cook Inlet	Kenai	1	2004	55	
		Yetna	2	2004	50	
		Susitna	3	2004	25	
	Prince William Sound	38 mile		4	2003	48
			Ibeck	5	2003	55
		Yakutat Forelands	Alaganik	6	2003	55
			Esker	7	2003	55
			Tawah	8	2003	55
	South	Upper Lynn Canal	Lost	9	2003	14
			Situk	10	2003	55
			Seal	11	2003	55
			Awkwe	12	2003	55
Chilkat			13	2003	51	
Chilkoot			14	2003	55	
Tiaya			15	2003	55	
Skagway			16	2005	55	
Berners Bay			Berners	17	2003	55
			Lace	18	2003	55
		Antler	19	2003	55	
Stikine Strait		Stikine	20	2003	55	
Behm Canal		Eulachon	21	2003	55	
		Unuk	22	2003	55	
	Klahini	23	2003	55		
	Matney Slough	24	2003	55		
	Upper Landing	25	2003	55		
	Side Channel	26	2003	55		

Table 2. Results across all collections for each locus: number of alleles, allelic richness (A_R), effective number of alleles (A_E), unbiased expected heterozygosity (H_E), observed heterozygosity (H_O), and measure of genetic divergence (G_{ST}).

Locus	Alleles	A_R	A_E	H_E	H_O	G_{ST}
<i>Tpa103</i>	40	13.5	13.8	0.938	0.901	0.005
<i>Tpa104</i>	45	9.1	4.7	0.784	0.718	0.007
<i>Tpa111</i>	16	6.6	5.1	0.808	0.820	0.009
<i>Tpa112</i>	25	9.0	7.5	0.873	0.816	0.002
<i>Tpa113</i>	32	8.1	5.8	0.835	0.839	0.002
<i>Tpa114</i>	17	7.0	3.2	0.683	0.705	0.009
<i>Tpa115</i>	16	4.9	3.1	0.687	0.691	0.004
<i>Tpa117</i>	20	4.8	1.6	0.383	0.384	0.002
<i>Tpa118</i>	23	4.0	2.2	0.535	0.506	0.009
<i>Tpa119</i>	20	5.7	2.6	0.606	0.511	0.011
<i>Tpa121</i>	70	12.5	11.3	0.922	0.631	0.006
<i>Tpa122</i>	41	10.5	8.7	0.892	0.833	0.002
<i>Tpa127</i>	25	7.3	3.2	0.678	0.699	0.009
<i>Tpa129</i>	30	10.2	8.0	0.882	0.863	0.004

Table 3. Results for each collection across all loci: mean sample size (N), percentage polymorphic loci at the 95% criterion ($\%P$), allelic richness (A_R), effective number of alleles (A_E), unbiased expected heterozygosity (H_E), and observed heterozygosity (H_O). The N listed here indicates the mean number of samples across loci where data were successfully collected, which may differ from the N in Table 1.

Collection	N	$\%P$	A_R	A_E	H_E	H_O
Kenai	49	100.0	7.6	5.6	0.729	0.699
Yetna	46	100.0	7.6	5.6	0.734	0.694
Susitna	24	100.0	7.5	5.2	0.717	0.697
38 mile	46	100.0	7.7	5.7	0.747	0.715
Ibek	52	100.0	8.0	5.9	0.758	0.748
Alaganik	52	100.0	8.0	5.9	0.764	0.743
Esker	52	100.0	7.6	5.6	0.741	0.723
Tawah	50	100.0	7.4	5.3	0.729	0.697
Lost	12	100.0	7.1	4.4	0.713	0.710
Situk	45	100.0	7.7	5.8	0.758	0.793
Seal	41	100.0	7.0	4.6	0.702	0.750
Awkwe	50	100.0	7.8	5.8	0.755	0.753
Chilkat	50	100.0	8.2	6.0	0.771	0.782
Chilkoot	53	100.0	7.9	5.4	0.745	0.707
Tiaya	52	100.0	7.9	5.9	0.769	0.717
Skagway	52	100.0	7.9	5.8	0.762	0.721
Berners	55	100.0	8.0	5.7	0.772	0.721
Lace	55	100.0	8.0	5.8	0.759	0.734
Antler	52	100.0	8.1	5.9	0.755	0.729
Stikine	50	100.0	7.9	5.4	0.769	0.710
Eulachon	45	100.0	7.7	5.4	0.763	0.762
Unuk	54	100.0	8.2	6.2	0.766	0.747
Klahini	54	100.0	8.1	5.9	0.764	0.735
Matney Slough	51	100.0	7.7	5.4	0.736	0.742
Upper Landing	45	100.0	7.1	5.0	0.700	0.750
Side Channel	54	100.0	8.5	6.0	0.768	0.735

Table 4. Collection pairwise tests of allele frequency homogeneity. * = significant test ($P < 0.05$); NS = not significant. Reference Table 1 for collection names.

Collection	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
1																											
2	NS																										
3	NS	NS																									
4	NS	NS	NS																								
5	NS	NS	NS	NS																							
6	NS	NS	NS	NS	NS																						
7	NS	NS	NS	NS	NS	NS																					
8	NS																										
9	NS																										
10	NS																										
11	NS																										
12	NS																										
13	NS	*	*	NS																							
14	NS	*	NS	NS	NS	NS	NS	*	NS	NS	NS	NS	NS														
15	NS																										
16	*	*	*	NS	*	NS	NS	NS	NS																		
17	*	NS	*	NS	NS	NS	NS	NS	*	NS																	
18	NS	NS	*	NS	NS	NS	NS	*	NS	*	NS																
19	NS	NS	*	NS	NS	NS	NS	*	NS																		
20	*	NS	NS	NS	NS	NS	NS	*	NS																		
21	*	*	*	NS	NS	*	*	*	NS	NS	*	*	NS														
22	*	*	NS	NS	NS	NS	*	*	NS	NS	*	NS															
23	*	*	*	NS	NS	*	NS	*	NS	NS	*	NS															
24	*	*	*	NS																							
25	*	*	*	*	*	*	*	*	NS	*	NS	*	NS														
26	*	*	NS	NS	NS	NS	NS	*	NS																		

Table 5. Hierarchical tests of homogeneity based on 12 loci.

Source of variation		df	G-test
Northern region	CI	36	40.0
	PWS	36	43.6
	YF	90	117.7
Within Northern areas		162	201.3
Among Northern areas		36	61.5
Total Northern region		198	262.8*
Southern region	ULC	54	73.0
	BB	36	41.3
	SS	No test	
	BC	90	122.8
Within Southern areas		180	237.0*
Among Southern areas		54	49.4
Total Southern region		234	286.5
Total within regions		432	549.2*
Between regions		18	97.0*
Total		450	646.2*

* $P < 0.05$

Table 6. Hierarchical gene diversity analysis based on 12 loci.

Source of variation	Gene diversity	G_{ST} -statistics	$N_e m$
Average within collections	$H_S = 0.7480$	$H_S / H_T = 0.9950$	
Average among collections within areas	$D_{SL} = 0.0014$	$G_{SL} = 0.0018$	135
Average among areas within regions	$D_{LR} = 0.0006$	$G_{LR} = 0.0008$	156
Average between regions	$D_{RT} = 0.0018$	$G_{RT} = 0.0024^*$	26
Total gene diversity	$H_T = 0.7518$	$G_{ST} = 0.0050^*$	49

* $P < 0.05$

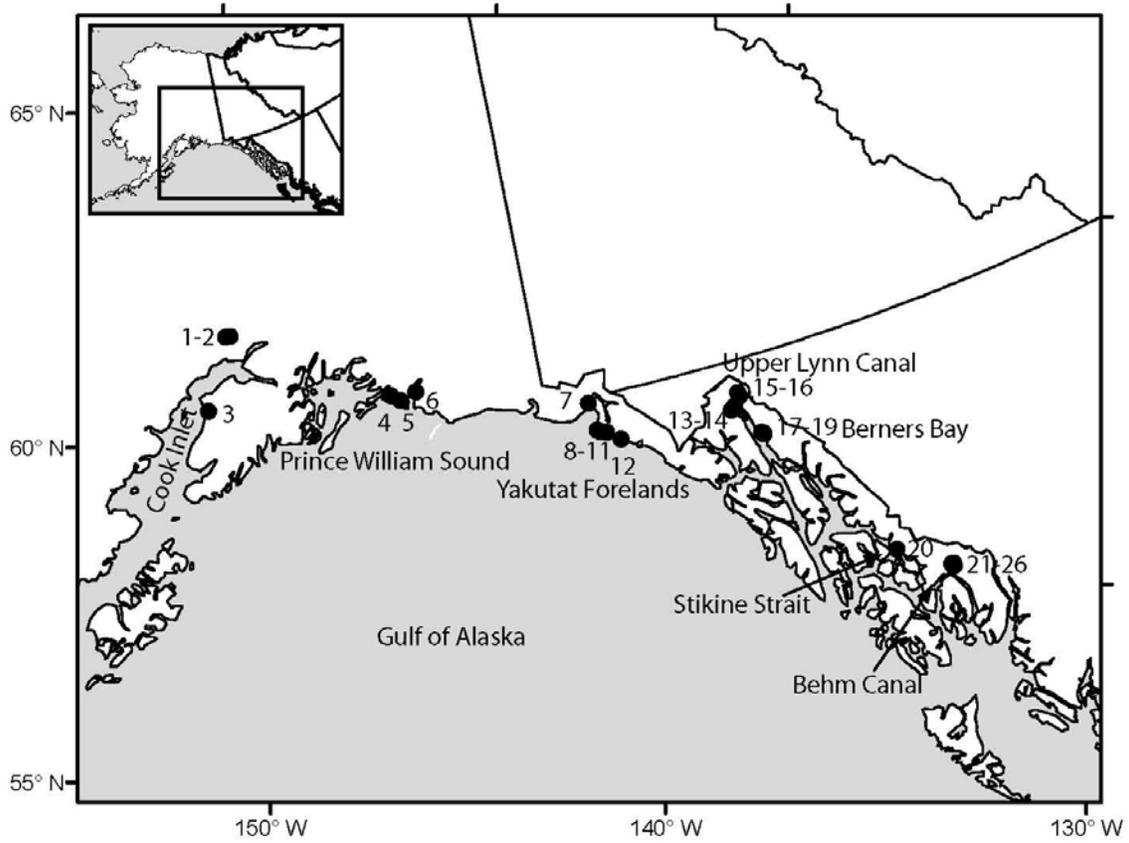


Figure 1. Collection locations, reference Table 1 for collection names.

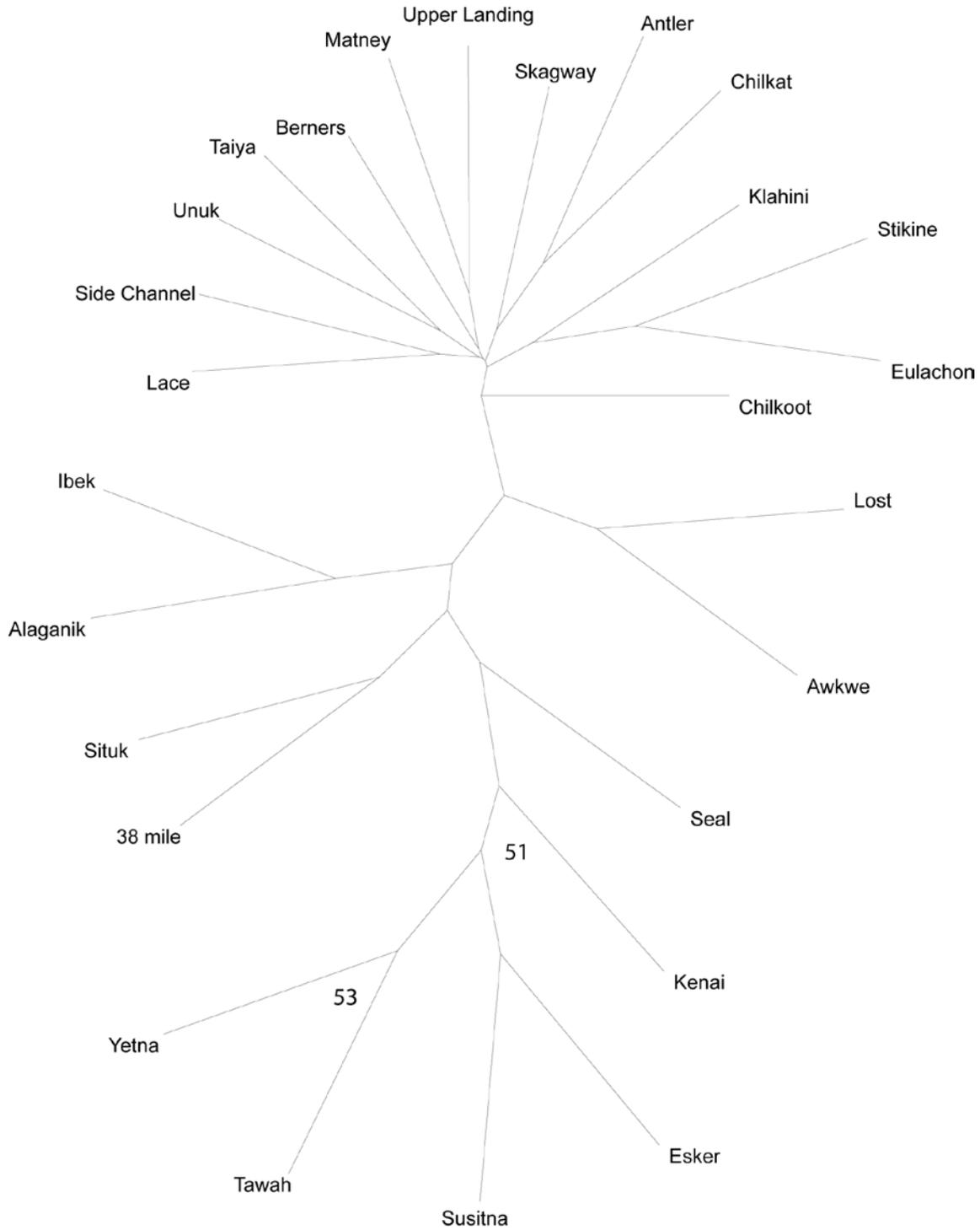
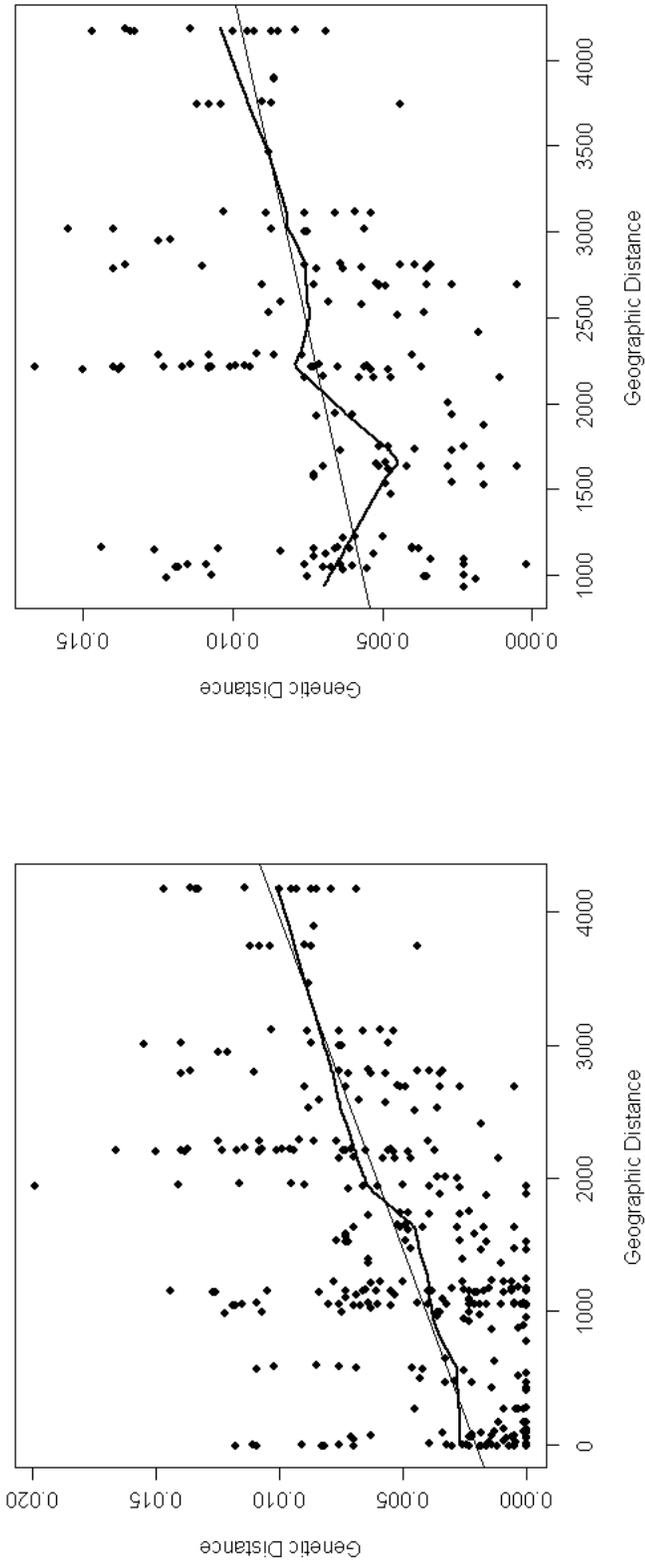


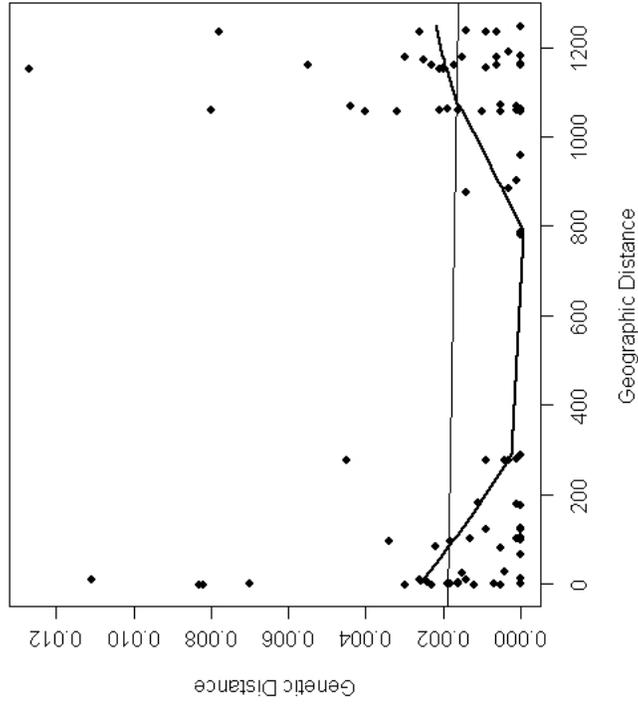
Figure 2. Consensus neighbor-joining dendrogram of Cavalli-Sforza and Edwards (1967) chord distances calculated from allele frequencies at 12 loci after 1000 bootstrap iterations.



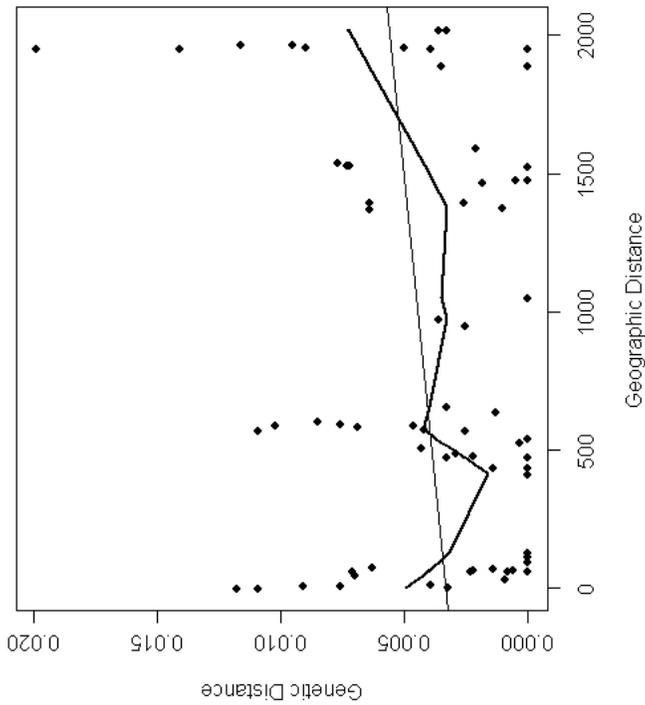
A. Among all collections ($y = 0.000002x + 0.0021$,
 $R^2 = 0.28, P < 0.05$)

B. Between regions ($y = 0.000001x + 0.0044$,
 $R^2 = 0.11, P < 0.05$)

Figure 3. Collection pairwise scatter plots of genetic distance (F_{ST}) on geographic distance (kilometers). Linear and lowest trend lines are displayed.



D. Within southern region ($y = 0.0000002x + 0.00019$,
 $R^2 = 0.002, P > 0.05$)



C. Within northern region ($y = 0.000001x + 0.00033$,
 $R^2 = 0.04, P > 0.05$)

Figure 3. Continued.