THE COLONIZATION MECHANISM OF PINK SALMON POPULATIONS IN GLACIER BAY, ALASKA, BASED ON GENETIC DATA

A

THESIS

Presented to the Faculty

of the University of Alaska Fairbanks

in Partial Fulfillment of the Requirements

for the Degree of

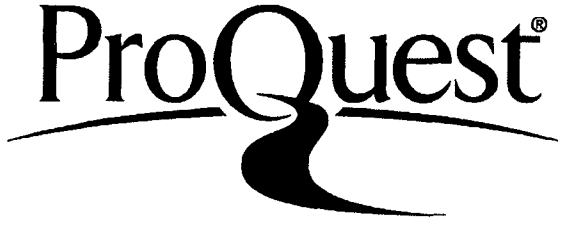
DOCTOR OF PHILOSOPHY

By

Christine M. Kondzela, M.S.

Fairbanks, Alaska

December 2010



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

THE COLONIZATION MECHANISM OF PINK SALMON POPULATIONS IN GLACIER BAY, ALASKA, BASED ON GENETIC DATA

By

Christine M. Kondzela

RECOMMENDED:

Dr. Richard Wilmot

Dr. Jerry Pella

Dr. William Smoker

Dr. Bruce Finney

Dr. A. J. Gharrett, Advisory Committee Chair

Dr. Trent Sutton, Program Head, Fisheries Division

APPROVED:

Dr. Michael Castellini, Interim Dean,
School of Fisheries and Ocean Sciences

Dr. Lawrence Duffy, Interim Dean of the Graduate School

Dec 7, 2010

Date

General Abstract

Following retreat of the last glacial advance in the early 1700s, pink salmon *Oncorhynchus gorbuscha* colonized many watersheds in Glacier Bay, Alaska. Streams in the lower Bay were populated first, and colonization proceeded up the Bay during the last 200 years. The objective of this study was to use analyses of genetic data—microsatellite and allozyme loci, and mitochondrial DNA haplotypes—to elucidate the colonization mechanism. The even- and odd-year broodlines served as replicate experiments; the mechanisms of colonization for the two broodlines were similar in most respects. The population genetic structure, based on allele/haplotype frequencies and genetic diversity (F_{ST}), suggested that in general, deglaciated streams were populated by colonists from nearby locations. The populations in lower Glacier Bay were likely established by colonists from populations outside Glacier Bay. In turn, the lower Bay populations contributed colonists to populations farther up the Bay, which subsequently provided colonists to the most recently deglaciated locations in the upper Bay, although in the even-year there appeared to be some contribution to the youngest populations from older populations, outside of or in lower Glacier Bay. Few genetically divergent donor sources contributed colonists based on the limited linkage disequilibrium, higher relatedness, and lower allelic diversity within Glacier Bay populations. The number of fish involved in initial colonization was not large, based on slightly reduced genetic diversity within Glacier Bay, but minimal founder effect signals precluded very small numbers of fish as well. Most of the genetic variation appeared early in the formation of populations and effective population size estimates were >100 fish in every population. Some gene flow after initial colonization is supported by the increased allelic diversity and decline in relatedness with population age, but heterogeneity within Glacier Bay suggested that gene flow must be limited among some populations. Colonization of the youngest streams coincided with the historically high abundance of pink salmon in Southeast Alaska during the 1990s; I speculate that the rapid expansion in the size of these populations subsequent to this study was the result of high survival rather than extensive gene flow.

Table of Contents

Page

Page

List of Figures

Page

Page

Page

Page

List of Tables

Page

Page

Page

List of Other Materials

List of Appendices

Page

Page

Acknowledgements

I would like to express my appreciation for the long-term support, including substantial financial support, from the Auke Bay Laboratories, National Marine Fisheries Service, NOAA, Juneau, Alaska, and from my supervisors, Drs. Richard Wilmot and Jeffrey Guyon. I am grateful for authorization of field collections in Glacier Bay from the National Park Service; the collections also serve in the development of coastwide genetic baselines in collaboration with other fisheries laboratories. Inquiry into the colonization of salmon in Glacier Bay started with a suggestion years ago from my graduate advisor, Dr. A. J. Gharrett, who at the time was my supervisor at NOAA. I deeply appreciate Dr. Gharrett's interest in the development of my fisheries career and for sharing his knowledge, counsel, and friendship.

Many people contributed in various ways to this study, without whom this project would never have been completed, and I thank them for their involvement. Chad Soiseth and Mark Schroeder provided field and administrative support through the Glacier Bay National Park and Preserve, National Park Service. Dr. Alexander Milner was kind enough to share both his experience and field time in Glacier Bay; his students, Elizabeth Flory and Kiernan Monaghan, provided abundance and spawning time information about salmon in some of the Glacier Bay streams. Conversation with Greg Streveler enlightened me about Glacier Bay stream development and the natural history of the area. Michele Masuda (NOAA) provided field assistance and computer programming for genetic data analysis. Laura Joralemon provided help with SAS programming and data formatting, as well as sample preparation in lab. Consultation with Dr. David Tallmon (UAS) provided insight about N_e, population bottlenecks, and colonization concepts. Don Ingledue and Kevin Monagle (ADFG) provided salmon escapement data for streams in northern Southeast Alaska and the Glacier Bay area. Dr. Robin Waples (NOAA) shared his knowledge of effective population size estimators as well as the use of several computer programs; his review of results from the N_e analyses is appreciated. I am grateful for editing assistance from Julie Gimbel.

Support in the lab was given by the following people: Wei Cheng, Dima Churikov,

Nick DeCovich, Bobette Dickerson, Elaine Forrest, Sara Gilk, Andy Gray, Chuck Guthrie, Chris Habicht, Sharon Hall, Carrie Hoover, Laura Joralemon, Kurt Kondzela, Zhuozhuo Li, John Pohl, and Sharon Wildes. Hanhvan Nguyen provided major assistance with extraction of DNA and analysis of the allozyme markers.

I would also like to thank the following people for their crucial assistance in the field: Dr. Richard Carlson, Dee Casey, Chuck Guthrie, Pat Harris, Michele Masuda, Hanhvan Nguyen, Claire Noll, Dr. James Olsen, John Pohl, Charley Russell, Sharon Wildes, and Dr. Richard Wilmot. I also thank the crews of the NOAA ships *Murre II* and *John N. Cobb*, with particular thanks to Strydr Nutting. Special thanks go to Pat Harris and Jeff Rounds for transportation and lodging for one (very wet!) field season onboard the sailing vessel *Pourquoi Rene*.

Last, but not least, I thank my husband, Kurt Kondzela, who is the most sanguine, patient person I know. Completion of this dissertation would not have been possible without his support through my years of graduate study.

General Introduction

Pink salmon *Oncorhynchus gorbuscha* are important ecological components of the North Pacific Ocean and adjacent seas, as well as the surrounding coastal freshwater environments of the Asian and North American continents (Heard 1991). They are the most abundant Pacific salmonid species and support valuable fisheries in Russia, Canada, and the United States (Beamish and Bouillon 1993). Pink salmon endured repeated advance and retreat of glaciations over much of their range, including the Aleutian Island chain, along Alaska's southern coast, and down along British Columbia to the Washington coast. Throughout the Pleistocene, most of this region was covered by the Cordilleran Ice Sheet that extended out to the exposed continental shelf producing Arctic-like conditions (Mann and Hamilton 1995; Kaufman and Manley 2004), although some ice-free areas may have served as refugia (Warner et al. 1982; Carrara et al. 2007). During this time, two broodlines developed, one from the other possibly multiple times (Churikov and Gharrett 2002), resulting in a two-year life history with reproductively isolated even- and odd-year lines. Some streams support runs of only one broodyear, whereas other streams have both; cyclic dominance occurs in some areas and can shift between the broodlines (Ricker 1962).

Salmonids are highly migratory organisms, and homing to natal sites is a central feature of their life histories (Groot and Margolis 1991). Homing occurs with great precision in most situations and is controlled by genetic factors (Quinn 1993) and learned components (olfactory imprinting; Harden Jones 1968; Cooper et al. 1976; Scholz et al. 1976; Quinn 1990). The advantages of homing are increased fitness acquired from local adaptation (Reisenbichler 1988; Smoker et al. 1998) and increased survival of offspring by returning to spawn in proven, good rearing habitat (Ricker 1972; McDowall 2001). The other life-history strategy is straying, thought to be in dynamic equilibrium with homing (Quinn 1984). Straying to new sites to spawn is an important component in the expansion of salmon to newly available favorable environments across geographic landscapes. It is a mechanism that prevents total loss from natural disasters, such as landslides and floods at the species and population level (e.g., Whitman et al. 1982;

Leider 1989), and provides an influx of genetic material after population bottlenecks (Gharrett and Smoker 1993). Due to the lack of multi-year age structure that occurs in other salmonids species, spatial straying provides the only means of buffering loss in pink salmon from natural disasters (Alexandersdottir 1987; Thorpe 1994). The low to moderate levels of population divergence (e.g., typical F_{ST} values are less than 0.05) exhibited by pink salmon are consistent with this limitation. Strength of population structure varies between the two broodlines and at different levels of geographical hierarchy across the range (Beacham et al. 1985, 1988; Gharrett et al. 1988; Zhivotovskii et al. 1990; Shaklee et al. 1991; Varnavskaya and Beacham 1992; Polyakova et al. 1993; Shaklee and Varnavskaya 1994; Olsen et al. 1998; Brykov et al. 1999; Seeb et al. 1999; Noll et al. 2001; Hawkins et al. 2002). The typically weak population structure within regions (e.g., Gharrett et al. 1988; Noll et al. 2001) suggests that proximate populations are more similar to each other than distant populations. In the short coastal streams of Southeast Alaska, a substantial portion of pink salmon adults can be comprised of probing fish from nearby streams (Jones and Thomason 1984; Maselko et al. 1999), a behavior that provides a mechanism for fish to move into nearby, available habitat. Based on tagging studies, fish that stray are far more likely to spawn in nearby systems (Thedinga et al. 2000). But even though the life history, biology, and ecology of pink salmon predicts this species would be more likely to stray than other salmonids species (Quinn 1985), pink salmon have the capacity to home with great precision and fidelity (McGregor et al. 1998; Gharrett et al. 2001).

Salmonids rapidly colonize new habitat, perhaps in part because of their phenotypic plasticity and ability to adapt quickly (e.g., Pavey et al. 2010; Ramstad et al. 2010), even in unstable, geologically young habitat. Lack of an extended freshwater residence in either the juvenile or adult life-history stages may be a factor in the rapid colonization of streams by pink salmon because of limited exposure to the environmental hazards of freshwater residence and absence of feeding in freshwater. Pink salmon have the capability of rapid expansion into unoccupied freshwater habitat (Great Lakes—Kwain and Lawie 1981, Kwain 1987, Dumont et al. 1988; Sashin Creek—Harry and Olson

1963, Vallion et al. 1981; Fraser River—Vernon 1962, Withler 1982; and northern Europe—Kossov et al. 1960, Neave 1965, Withler 1982).

Pink salmon populated numerous streams in Glacier Bay, Alaska, after the retreat of ice began in the early-1700s (Milner and Bailey 1989). There are many possible ways that new populations of pink salmon could have established in Glacier Bay after deglaciation. In order to determine the most likely colonization events, the genetic compositions of present day populations were examined and compared from the two broodlines, which serve as replicate experiments for analysis of the colonization process—the odd-year in Chapter 1, and the even-year in Chapter 2. The question of how colonization proceeded was addressed by asking a series of contrasting questions that were designed to look at extreme situations while anticipating that the actual mechanisms would lie somewhere within the framework of possibilities.

First, were populations derived from colonists from nearby or far away sources? If source populations were from nearby locations, the colonists would genetically resemble those populations more than distant populations. Descriptive analyses were followed by homogeneity testing and comparisons of estimates of genetic diversity (F_{ST}) to examine the temporal and spatial relationships of populations within and nearby Glacier Bay. Once we understood the probable location of donor sources, we focused on other aspects of colonization.

The second question asked if a single population or multiple source populations contributed colonists? If colonists came from multiple, genetically different sources, linkage disequilibrium should be present at the time of colonization (Gharrett and Zhivotovsky 2003) and persist for several generations (Hedrick 2005). There would also likely be higher allele richness and heterozygosity, as well as lower relatedness among fish (Ritland 2000) in new populations, than if new populations were derived from a single source population.

The third question asked if a small or large number of colonists were involved? For example, founder effects would be expected if a small number of fish colonized a location; and due to stochastic sampling of alleles, significant heterogeneity could result

among new populations (e.g., Gharrett and Thomason 1987; Spencer et al. 2000). Whereas, if a larger number of fish were involved, new populations would be genetically similar to the older source populations.

The fourth question asked if colonization was a one-time event or a process of recurrent gene flow? If colonization was a one-time event, genetic diversity would decrease and relatedness would increase over time, and heterogeneity among populations would develop. If gene flow is recurrent, then there should be an increase in both genetic diversity within populations and homogeneity among populations over time.

The objective of our study was to examine alternative colonization scenarios that we could expect to distinguish from the genetic signals. The possible colonization mechanism was narrowed within the contrasting boundaries outlined in the four questions above. The questions were addressed with an overlapping set of analyses, so the material in each chapter was organized around the questions to provide clarity in the interpretation of the results. Mitochondrial DNA variation was compared with nuclear variation (allozymes and microsatellites). Mitochondrial markers are haploid and maternally inherited in salmon, and thus have about one-quarter the effective population size of diploid nuclear markers, which should make them more sensitive to numbers of colonists and source populations (Nei and Tajima 1981; Ferris and Berg 1987; Avise 1994). In addition, because mtDNA is matrilineal, sex-related differences in dispersal may produce differences between mitochondrial and diploid nuclear markers. The number of genetic markers, sample sizes, and selected mtDNA regions and restriction endonucleases used in our study were chosen to avoid significant bias that can occur when estimating genetic diversity.

References

Alexandersdottir, M. 1987. Life history of pink salmon (*Oncorhynchus gorbuscha*) and implications for management in southeastern Alaska. Doctoral dissertation. University of Washington, Seattle.

Avise, J. C. 1994. Molecular markers, natural history and evolution. Chapman & Hall, New York.

Beacham, T. D., R. E. Withler, and A. P. Gould. 1985. Biochemical genetic stock identification of pink salmon (*Oncorhynchus gorbuscha*) in southern British Columbia and Puget Sound. Canadian Journal of Fisheries and Aquatic Sciences 42: 1474-1483.

Beacham, T. D., R. E. Wither, C. B. Murray, and L. W. Barner. 1988. Variation in body size, morphology, egg size, and biochemical genetics of pink salmon in British Columbia. Transactions of the American Fisheries Society 117:109-126.

Beamish, R. J., and D. R. Bouillon. 1993. Pacific salmon production trends in relation to climate. Canadian Journal of Fisheries and Aquatic Sciences 50:1002-1016.

Brykov, V. A., N. Polyakova, L. A. Skurikhina, A. D. Kukhlevsky, O. N. Kirillova, D. Churikov, A. I. Pudovkin, and A. J. Gharrett. 1999. Analysis of mtDNA indicates weak temporal genetic heterogeneity in pink salmon spawning runs in two rivers on Sakhalin Island. Journal of Fish Biology 55:617-635.

Carrara, P. E., T. A. Ager, and J. F. Baichtal. 2007. Possible refugia in the Alexander Archipelago of southeastern Alaska during the late Wisconsin glaciation. Canadian Journal of Earth Sciences 44:229-244.

Churikov, D., and A. J. Gharrett. 2002. Comparative phylogeography of the two pink salmon broodlines: an analysis based on a mitochondrial DNA genealogy. Molecular Ecology 11:1077-1101.

Cooper, J. C., A. T. Scholz, R. M. Horrall, A. D. Hasler, and D. M. Madison. 1976. Experimental confirmation of the olfactory hypothesis with homing, artificially imprinted coho salmon (*Oncorhynchus kisutch*). Journal of Fisheries Research Board of Canada 33:703-710.

Dumont, P., J. F. Bergeron, P. Dulude, Y. Mailhot, A. Rouleau, G. Ouellet, and J. P. Lebel. 1988. Introduced salmonids: where are they going in Quebec watersheds of the Saint-Laurent River? Fisheries 13(3):9-17.

Ferris, S. D., and W. J. Berg. 1987. The utility of mitochondrial DNA in fish genetics and fishery management. Pages 277-299 *in* N. Ryman, and F. Utter, editors. Population genetics & fishery management, University of Washington Press, Seattle.

Gharrett, A. J., S. Lane, A. J. McGregor, and S. G. Taylor. 2001. Use of a genetic marker to examine genetic interaction among subpopulations of pink salmon (*Oncorhynchus gorbuscha*). Genetica 111:259-267.

Gharrett, A. J., C. Smoot, A. J. McGregor, and P. B. Holmes. 1988. Genetic relationships of even-year northwestern Alaskan pink salmon. Transactions of the American Fisheries Society 117:536-545.

Gharrett, A. J., and W. W. Smoker. 1993. A perspective on the adaptive importance of genetic infrastructure in salmon populations to ocean ranching in Alaska. Fisheries Research 18:45-58.

Gharrett, A. J., and M. A. Thomason. 1987. Genetic changes in pink salmon (*Oncorhynchus gorbuscha*) following their introduction into the Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 44:787-792.

Gharrett, A. J., and L. A. Zhivotovsky. 2003. Migration. Pages 141-174 *in* E. M. Hallerman, editor. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland.

Groot, C., and L. Margolis. 1991. Pacific salmon life histories. University of British Columbia Press, Vancouver.

Harden Jones, F. R. 1968. Fish migration. Edward Arnold (Publishers) Ltd., London.

Harry Jr., G. Y., and J. M. Olson. 1963. Straying of pink salmon to Sashin Creek Little Port Walter Bay. Manuscript Report, MR 63-3, Department of Interior, U.S. Fish and Wildlife Service, Bureau of Commercial Fisheries, Biological Laboratory, Auke Bay, Alaska.

Hawkins, S. L., N. V. Varnavskaya, E. A. Matzak, V. V. Efremov, C. M. Guthrie III, R. L. Wilmot, H. Mayama, F. Yamazaki, and A. J. Gharrett. 2002. Population structure of odd-broodline Asian pink salmon and its contrast to the even-broodline structure. Journal of Fish Biology 60:370-388.

Heard, W. R. 1991. Life history of pink salmon (*Oncorhynchus gorbuscha*). Pages 119-230 *in* C. Groot, and L. Margolis, editors. Pacific salmon life histories. University of British Columbia Press, Vancouver.

Hedrick, P. W. 2005. Genetics of populations, 3rd edition. Jones and Bartlett Publishers, Boston.

Jones, J. D., and G. Thomason. 1984. U.S./Canada salmon stock interception research, southern southeastern Alaska pink salmon (*Oncorhynchus gorbuscha*) tagging study, 1982. Informational Leaflet No. 231, Alaska Department of Fish and Game, Division of Commercial Fisheries, Juneau.

Kaufman, D. S., and W. F. Manley. 2004. Pleistocene maximum and late Wisconsin glacier extents across Alaska, U.S.A. Pages 9-27 *in* J. Ehlers and P. L. Gibbard, editors. Quaternary glaciations - extent and chronology, part II, North America. Developments in Quaternary Science, Volume 2b, Elsevier, Netherlands.

Kossov, E. G., M. S. Lazarev, and L. V. Polikashin. 1960. Pink salmon in the basins of the Barents and White Seas. Rybnoe Khoziaistvo 36:20-25, (Fisheries Research Board of Canada, Translation Series No. 323).

Kwain, W. 1987. Biology of pink salmon in the North American Great lakes. Pages 57-65 *in* M. J. Dadswell, R. J. Klauda, C. M. Moffitt, and R. L. Saunders, editors. Common strategies of anadromous and catadromous fishes. American Fisheries Society Symposium 1, American Fisheries Society, Bethesda, Maryland.

Kwain, W., and A. H. Lawrie. 1981. Pink salmon in the Great Lakes. Fisheries 6(2):2-6.

Leider, S. A. 1989. Increased straying by adult steelhead trout, *Salmo gairdneri*, following the 1980 eruption of Mount St. Helens. Environmental Biology of Fishes 24:219-229.

Mann, D. H., and T. D. Hamilton. 1995. Late Pleistocene and Holocene paleoenvironments of the north Pacific coast. Quaternary Science Reviews 14:449-471.

Maselko, J. M., A. C. Wertheimer, and J. F. Thedinga. 1999. Estimating probing behavior of pink salmon. Pages 40-48 *in* S. Hawkins, rappoteur. Proceedings of the Northeast Pacific Pink & Chum Salmon Workshop, Juneau, Alaska.

McDowall, R. M. 2001. Anadromy and homing: two life-history traits with adaptive synergies in salmonid fishes? Fish and Fisheries 2:78-85.

McGregor, A. J., S. Lane, M. A. Thomason, L. A. Zhivotovsky, W. W. Smoker, and A. J. Gharrett. 1998. Migration timing, a life history trait important in the genetic structure of pink salmon. North Pacific Anadromous Fish Commission Bulletin 1:262-273.

Milner, A. M., and R. G. Bailey. 1989. Salmonid colonization of new streams in Glacier Bay National Park, Alaska. Aquaculture and Fisheries Management 20:179-192.

Neave, F. 1965. Transplants of pink salmon. Fisheries Research Board of Canada. Manuscript Report Series (biological) 830.

Nei, M., and F. Tajima. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97:145-163.

Noll, C., N. V. Varnavskaya, E. A. Matzak, S. L. Hawkins, V. V. Midanaya, O. N. Katugin, C. Russell, N. M. Kinas, C. M. Guthrie III, H. Mayama, F. Yamazaki, B. P. Finney, and A. J. Gharrett. 2001. Analysis of contemporary genetic structure of even-broodyear populations of Asian and western Alaskan pink salmon, *Oncorhynchus gorbuscha*. Fishery Bulletin 99:123-138.

Olsen J. B., L. W. Seeb, P. Bentzen, and J. E. Seeb. 1998. Genetic interpretation of broad-scale microsatellite polymorphism in odd-year pink salmon. Transactions of the American Fisheries Society 127:535-50.

Pavey, S. A., J. L. Nielsen, and T. R. Hamon. 2010. Recent ecological divergence despite migration in sockeye salmon (*Oncorhynchus nerka*). Evolution 64:1773-1783.

Polyakova, N. E., T. V. Malinina, A. V. Galeev, D. V. Zaikin, V. A. Brykov, and Y. P. Altukhov. 1993. Mitochondrial DNA variability in the pink salmon. Soviet Genetics 28(12):1610-1616.

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. Volume Proceedings of a NATO Advanced Research Institute on Mechanisms of Migration in Fishes, held December 13-17, 1982, in Acquafredda di Maratea, Italy. Plenum Press, New York.

Quinn, T. P. 1985. Homing and the evolution of sockeye salmon (*Oncorhynchus nerka*). Contributions in Marine Science 27:353-366.

Quinn, T. P. 1990. Current controversies in the study of salmon homing. Ethology Ecology and Evolution 2:49-63.

Quinn, T. P. 1993. A review of homing and straying of wild and hatchery-produced salmon. Fisheries Research 18:29-44.

Ramstad, K. M., C. A. Woody, and F. W. Allendorf. 2010. Recent local adaptation of sockeye salmon to glacial spawning habitats. Evolutionary Ecology 24:391-411.

Reisenbichler, R. R. 1988. Relation between distance transferred from natal stream and recovery rate for hatchery coho salmon. North American Journal of Fisheries Management 8:172-174.

Ricker, W. E. 1962. Regulation of the abundance of pink salmon populations. Pages 155-201 *in* N. J. Wilimovsky, editor. Symposium on pink salmon. H. R. MacMillan Lectures in Fisheries, Institute of Fisheries, University of British Columbia, Vancouver.

Ricker, W. E. 1972. Heredity and environmental factors affecting certain salmonid populations. Pages 27-160 *in* R. C. Simon and P. A. Larkin, editors. The stock concept in Pacific salmon. H. R. MacMillan Lectures in Fisheries, Institute of Fisheries, University of British Columbia, Vancouver.

Ritland, K. 2000. Marker-inferred relatedness as a tool for detecting heritability in nature. Molecular Ecology 9:1195-1204.

Scholz, A. T., R. M. Horrall, J. C. Cooper, and A. D. Hasler. 1976. Imprinting to chemical cues: the basis for home stream selection in salmon. Science 192:1247-1249.

Seeb, J. E., C. Habicht, W. D. Templin, L. W. Seeb, J. B. Shaklee, and F. M. Utter. 1999. Allozyme and mitochondrial DNA variation describe ecologically important genetic structure of even-year pink salmon inhabiting Prince William Sound, Alaska. Ecology of Freshwater Fish 8:122-140.

Shaklee, J. B., D. C. Klaybor, S. Young, and B. A. White. 1991. Genetic stock structure of odd-year pink salmon, *Oncorhynchus gorbuscha* (Walbaum), from Washington and British Columbia and potential mixed-stock fisheries applications. Journal of Fish Biology 39(Supplement A):21-34.

Shaklee, J. B., and N. V. Varnavskaya. 1994. Electrophoretic characterization of odd-year pink salmon (*Oncorhynchus gorbuscha*) populations from the Pacific coast of Russia, and comparison with selected North American populations. Canadian Journal of Fisheries and Aquatic Sciences 51(Suppl. 1):158-171.

Smoker, W. W., A. J. Gharrett, and M. S. Stekoll. 1998. Genetic variation of return date in a population of pink salmon: a consequence of fluctuating environment and dispersive selection? Alaska Fishery Research Bulletin 5:46-54.

Spencer, C. C., J. E. Neigel, and P. L. Leberg. 2000. Experimental evaluation of the usefulness of microsatellite DNA for detecting demographic bottlenecks. Molecular Ecology 9:1517-1528.

Thedinga, J. F., A. C. Wertheimer, R. A. Heintz, J. M. Maselko, and S. D. Rice. 2000. Effects of stock, coded-wire tagging, and transplant on straying of pink salmon (*Oncorhynchus gorbuscha*) in southeastern Alaska. Canadian Journal of Fisheries and Aquatic Sciences 57:2076-2085.

Thorpe, J. E. 1994. Salmonid flexibility: responses to environmental extremes. Transactions of the American Fisheries Society 123:606-612.

Vallion, A. C., A. C. Wertheimer, W. R. Heard, and R. M. Martin. 1981. Summary of data and research pertaining to the pink salmon population at Little Port Walter, Alaska, 1964-80. National Marine Fisheries Service, Northwest and Alaska Fisheries Science Center, Processed Report 81-10, Auke Bay, Alaska.

Varnavskaya, N. V., and T. D. Beacham. 1992. Biochemical genetic variation in odd-year pink salmon (*Oncorhynchus gorbuscha*) from Kamchatka. Canadian Journal of Zoology 70:2115-2120.

Vernon, E. H. 1962. Pink salmon populations of the Fraser River system. Pages 53-58 *in* N. J. Wilimosvky, editor. Symposium on pink salmon. H. R. MacMillan Lectures in Fisheries, Institute of Fisheries, University of British Columbia, Vancouver.

Warner, B. G., R. W. Mathewes, and J. J. Clague. 1982. Ice-free conditions on the Queen Charlotte Islands, British Columbia, at the height of late Wisconsin glaciation. Science 218:675-677.

Whitman, R. P., T. P. Quinn, and E. L. Brannon. 1982. Influence of suspended volcanic ash on homing behavior of adult chinook salmon. Transactions of the American Fisheries Society 111:63-69.

Withler, F. C. 1982. Transplanting Pacific salmon. Canadian Technical Report of Fisheries and Aquatic Sciences 1079, Nanaimo, British Columbia.

Zhivotovskii, L. A., M. K. Glubokovskii, R. M. Viktorovskii, A. M. Bronevskii, K. I. Afanas'ev, V. V. Efremov, L. N. Ermolenko, B. A. Kalabushkin, V. G. Kovalev, A. N. Makoedov, T. V. Malinina, S. P. Pustovoit, and G. A. Rubtsova. 1990. Genetic differentiation of pink salmon. Soviet Genetics 25(7):820-828.

Chapter 1

The colonization mechanism of odd-year pink salmon populations in Glacier Bay, Alaska, based on genetic data[1]

[1] Kondzela, C. M., R. L. Wilmot, and A. J. Gharrett. prepared. The colonization mechanism of odd-year pink salmon populations in Glacier Bay, Alaska, based on genetic data. Transactions of the American Fisheries Society X: xx-xx.

Abstract

Following the retreat of the last glacial advance in the early 1700s, salmonids colonized many of the deglaciated watersheds in Glacier Bay, Alaska. Genetic data (e.g., allozymes, microsatellites, and mitochondrial restriction fragment length polymorphisms [RFLPs]) from pink salmon *Oncorhynchus gorbuscha*, the most abundant and widespread of the Pacific salmon species that spawn within Glacier Bay, were used to elucidate the poorly understood colonization mechanism. For the odd-year broodline, our results suggested that deglaciated streams were populated by colonists from nearby locations. The populations outside Glacier Bay provided colonists to the first deglaciated locations in the lower Bay and these in turn contributed to the mid-Bay area, which subsequently provided migrants to the more recently deglaciated upper-Bay locations. More than one source may have contributed colonists (slightly elevated linkage disequilibrium, a deficit of heterozygosity expected under mutation-drift equilibrium, and heterogeneity among the younger populations), but contribution from many genetically divergent sources is unlikely (fewer microsatellite alleles and private alleles/haplotypes, and a higher level of relatedness in the younger populations). Colonization of odd-year pink salmon appears to be episodic with large numbers of fish, followed by low levels of gene flow. The effective population size estimates of the younger populations were typically several hundred fish and most of the genetic variation appeared early in the development of these populations. Although the number of fish involved in colonization was fairly large, the higher linkage disequilibrium, lower microsatellite allele richness, and higher levels of relatedness observed in the younger populations reduced the possibility of massive gene flow during initial colonization. Also, heterogeneity among populations exists within Glacier Bay even 40–60 generations after colonization began in the lower-Bay locations. Colonization of the youngest streams and the subsequent rapid expansion in the size of these populations coincided with the large increase in pink salmon escapement in Southeast Alaska during the 1990s.

Introduction

Pink salmon *Oncorhynchus gorbuscha* have probably existed as a species for more than 6 million years (Smith 1992). Along with other Pacific salmon species, pink salmon experienced the repeated glacier advances of the Pleistocene, a time with periods of cooler temperatures in the North Pacific Ocean and greatly restricted availability of freshwater habitat. During those advances, pink salmon likely survived in smaller refugia along the coast of western North America and eastern Asia (Warner et al. 1982; McPhail and Lindsey 1986). During warmer periods, pink salmon populations presumably expanded into available coastal and freshwater habitat. Although salmon approaching maturity have a strong propensity to home from ocean or coastal regions to freshwater natal sites, some straying must have occurred for range expansion and colonization to occur. Sometime during their evolution, possibly multiple times (Gharrett and Thomason 1987; Churikov and Gharrett 2002), a two-year life history developed that resulted in distinct even- and odd-year broodlines, which are virtually isolated reproductively. Today, pink salmon populations are distributed around the North Pacific Rim from Japan to Washington State (Heard 1991). Some locations have only one broodline, others have both, one of which may dominate; and the dominant broodline can change over time. In Southeast Alaska, both broodlines are abundant (Heinl and Geiger 2005).

Most of Southeast Alaska has been ice-free since about 13,000 years before present (BP) (Calkin 1988; Mann and Hamilton 1995), and streams in this region have probably sustained salmon populations for thousands of years (Finney et al. 2002). Glacier Bay was ice-free then; however, repeated advances and retreats of ice from the Fairweather Range into Glacier Bay occurred during the Holocene, with a significant advance during the Little Ice Age that covered the entire bay with ice by about 1700 AD (Goldthwait et al. 1966; Goodwin 1988; Mann and Hamilton 1995). Retreat of the ice since the early-1700s has exposed a large, complex fjord surrounded by land that is now carved with watersheds, which are fed by ice and snowmelt and abundant rainfall. Deglaciation and isostatic rebound will continue to influence the hydrology of streams in Glacier Bay for the foreseeable future (Milner et al. 2000). Many of these new watersheds provide

spawning and rearing habitat for pink salmon, chum salmon *O. keta*, sockeye salmon *O. nerka*, and coho salmon *O. kisutch* (Milner 1987; Milner and Bailey 1989).

There are many possible ways in which new populations of pink salmon could have established in Glacier Bay after deglaciation. In order to determine the most likely colonization events, we examined and compared the genetic compositions of present day populations. We approached the question of how colonization proceeded by asking a series of contrasting questions (Table 1.1). Once we understood whether populations were probably derived from colonists from nearby or far away sources, the possible ways in which colonization occurred was narrowed. We then focused on the other aspects of colonization—the number of source populations, the number of colonists, and the frequency of immigration into new systems.

Were source populations from locations nearby or farther away?

There are two time frames to consider when we ask about the sources of colonists: (1) what were the sources of the original colonists to Glacier Bay, that is, the fish that seeded the older Glacier Bay populations in the lower part of the bay about 150 years ago?; and (2) what were the sources of recent colonization events within the last 60 years? If the source populations were from nearby locations, the colonists would genetically resemble those populations more than distant populations, i.e., the allele or haplotype frequency distribution of colonists would be more similar to those of nearby populations (e.g., Hawkins et al. 2002), and F_{ST} values would be low. If the donors were from populations farther away, then colonists would have allele frequencies that differ from fish of older neighboring populations; new populations would be heterogeneous in comparison to nearby populations, which would result in higher F_{ST} values. Descriptive analyses, such as population trees and principal component analysis, followed by explicit homogeneity tests that examine spatial and temporal relationships, will be used to determine whether colonists in Glacier Bay are similar to nearby or more distant populations.

Did single or multiple source populations contribute colonists?

If colonists came from multiple, genetically different population sources, linkage disequilibrium should exist at the time of colonization (Gharrett and Zhivotovsky 2003). Linkage disequilibrium does not decay in a single generation, unlike the Wahlund effect, and will persist for several generations (Hedrick 2005). A composite measure (correlation coefficients) of linkage disequilibrium will be examined. There would also likely be higher allele richness and heterozygosity, as well as lower average relatedness and larger variance of average relatedness (Ritland 2000) among fish in new populations for several generations, than if new populations were derived from a single source population.

Did new populations start with small or large numbers of fish?

If a small number of fish colonized a location, we would expect founder effects with reduced genetic variation (e.g., heterozygosity, allele richness, and haplotype diversity would all be lower), small effective population size, and higher linkage disequilibrium (Hedrick 2005); whereas, if a large number of fish were involved, new populations would be genetically similar to older populations, have large effective population sizes, and exhibit lower linkage disequilibrium. Also, due to a stochastic sampling effect, allele frequencies in newer populations established by few colonists could result in significant heterogeneity among new populations (e.g., Gharrett and Thomason 1987). A very small number of colonists would result in a deficit of alleles and an excess of heterozygosity in a mutation-drift steady state situation; this bottleneck signal erodes with a half life of about $4N_e$ generations (Kimura 1983; Cornuet and Luikart 1996).

Did colonization occur as a one-time event or as a recurrent process?

If immigration into new populations was recurrent, then (depending on other factors such as number and source of colonists) there should be an increase in both genetic diversity (e.g., allele richness) within populations and homogeneity among populations as they advance in age. If colonization was a one-time event, heterogeneity among

populations would develop; and the relatedness of individuals within populations would increase over time, particularly if small numbers of colonists were involved.

Approach

The objective of our study was to examine alternative colonization scenarios that we could expect to distinguish from the genetic signals. We compared mtDNA variation with nuclear variation. Mitochondrial markers are haploid in salmon, have about one-quarter the effective population size of nuclear markers, and should be more sensitive to numbers of colonists and source populations (Nei and Tajima 1981a). In addition, because mtDNA is matrilineal, sex-related differences in dispersal may produce differences between mitochondrial and diploid nuclear markers. Clearly, the colonization mechanism would have been somewhere within the contrasting boundaries described above. By addressing those questions, we should be able to narrow the possibilities and possibly generate some other questions to pursue.

Methods

Study Site and Sample Collection

Most streams were sampled between late August and early September in 1991 and 1993, when spawned-out fish were present, usually during a short time period on a single day. Consequently, samples provided a temporal snapshot of multiple systems (Figure 1.1; Table 1.2).

Streams that developed in the wake of ice retreat in Glacier Bay are progressively younger from the mouth to the upper reaches of Glacier Bay. Three locations in the middle to lower section of Glacier Bay were sampled: the north head of Berg Bay, the north head of the north arm of Fingers Bay, and the head of Tyndall Cove (Figure 1.1). The mouths of these streams were exposed in approximately 1830, 1845, and 1875, respectively, and have been ice-free for less than 150 years (G. Streveler and C. Soiseth, National Park Service [NPS], unpublished data). These three streams are referred to as "medium-aged" throughout our study, although many years or even decades probably

separate their colonization events. The three medium-aged streams were sampled just above the mouth, from approximately the lower one-half kilometer reach of the stream. The N. Berg and Tyndall streams had thousands of pink salmon present at the time of collection; the smaller N. Fingers stream had an order of magnitude fewer pink salmon, based on our observations. The mouths of the three youngest streams sampled from upper Glacier Bay became ice free during the mid-1900s: Nunatak in 1935, Wolf Pt. in 1947, and Gull Lake outlet in 1955 (G. Streveler and C. Soiseth, NPS, unpublished data). Although census information for salmon in Glacier Bay streams is scant, the nearly continuous summer stream ecology studies in the Wolf Pt. and Nunatak systems since the late 1970s (e.g., Milner 1987; Milner et al. 2000, 2007) documented that pink salmon were first seen in Wolf Pt. in 1989 (approximately 20 spawned-out carcasses; Milner 1994) and in Nunatak in 1985 (>800 fish; Milner and Bailey 1989). Because relatively few fish were present in these young populations, sampling was conducted from the intertidal area upstream to either a barrier falls or a lake, but typically along less than two kilometers of stream reach. Populations from three streams adjacent to Glacier Bay that were ice free during the Holocene (E. Kahtaheena, Homeshore, and Spasski creeks) were sampled for comparison with the younger populations within Glacier Bay. Samples from the latter two streams were an extension of a pink salmon genetic baseline development project (e.g., Gharrett et al. 1990). Those two systems have large spawning populations of pink salmon; peak aerial counts ranged from thousands to more than 100,000 fish (Kevin Monagle, Alaska Department of Fish and Game [ADFG], personal communication). The E. Kahtaheena and Homeshore streams were sampled usually in their lower half kilometer reach; the pink salmon in Spasski were sampled approximately 2.5 kilometers upstream but about two weeks earlier than all other locations. The peak spawning count for Spasski is often one month earlier than the peak count for Homeshore, although there is considerable overlap in the overall spawning season (Kevin Monagle, ADFG, personal communication).

At most sampling sites, 100 adult spawners were collected (Table 1.2) with large dipnets or a beach seine and held in the water, typically for a few minutes in dipnets and

less than one hour in beach seines, until each fish was examined for maturity. Fish were sacrificed quickly by a blow to the head if they had spawned; and muscle, heart, liver, and eye tissues were collected and frozen in liquid nitrogen as soon as possible. For fish that had not spawned, a muscle plug, approximately 6 mm diameter by 20 mm long, was taken near the dorsal surface behind the dorsal fin, the adipose fin was collected, and the fish was released. Tissues for allozyme analysis were stored at -70°C. Tissues used for DNA analysis were either stored at -70°C or in DNA buffer (Seutin et al. 1991) at -20°C and can be stored indefinitely under these conditions.

Data Acquisition

Allozyme loci were analyzed with starch-gel electrophoresis (Aebersold et al. 1987; Seeb et al. 2000; Hawkins et al. 2002) for all available samples, typically 100 fish in each collection. Twenty-three enzymatic activities were used to resolve a total of 49 loci (Table 1.3).

Deoxyribonucleic acid was extracted from heart tissue or, when heart tissue was unavailable, from muscle, liver, or fin tissues. Total DNA, approximately 100 ng per µL, was isolated with the DNeasy®[2] animal tissue protocol of Qiagen, Inc. (Germantown, Maryland).

A subset of samples, typically the first 50 fish collected, was analyzed at nine microsatellite loci from each collection. The microsatellite loci were polymerase chain reaction (PCR) amplified in 10 µL reactions [1X buffer (Sigma-Aldrich® product P-2317), 1.5 mM $MgCl_2$, 0.09 mM each dNTP, 0.4 µM reverse primer, 0.4 µM forward primer (0.395-0.360 µM unlabeled, 0.005-0.040 µM labeled), 0.5 units *Taq* polymerase (Sigma-Aldrich® product D-4545), and approximately 100 ng total DNA]. Primers were labeled by a fluorescing dye, either IRDye® 700 or IRDye® 800 (LI-COR, Inc., Lincoln, Nebraska), which separates genotypes of co-migrating loci. With these two dyes, loci *One102* and *One109* were co-amplified, as were *One103* and *One111*. After

[2] Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA.

amplification, products from the following loci were combined into two groups for genotyping (1) *Oki10* and μ*Sat60*, and (2) *One103*, *One111*, and *Oneμ13*. Genotypes at *One103* were identified with two sets of primers, originally described for *One101* and *One103* (Olsen et al. 2000), both of which target the same repeat DNA sequence. The reactions were denatured at 95°C for 10 minutes, followed by 29 cycles at 95°C for 20 seconds, annealed at 50-63°C for 30 seconds, and extended at 72°C for 30 seconds, followed by a final incubation at 72°C for five minutes and 4°C for one minute (Table 1.4).

Microsatellite loci were genotyped from 0.25 mm gels [6.45% acrylamide/bisacrylamide (KB^{PLUS} 6.5% Gel Matrix, LI-COR, Inc., Lincoln, Nebraska), 0.074% ammonium persulfate, 0.074% tetramethylethylenediamine (TEMED)] on a LI-COR 4200 DNA Analyzer with SAGA™ software (LI-COR, Inc., Lincoln, Nebraska) and LI-COR 50-350 base pair (bp) DNA size standards to assist allele calls.

A subset of samples, typically the first 40 fish from each collection, was analyzed for mtDNA restriction site variation. To capture the underlying mtDNA genealogy of pink salmon that was described in Churikov and Gharrett (2002) and to minimize the effort necessary to identify informative sites, combinations of mtDNA regions and restriction endonucleases that targeted specific sites were chosen (Table 1.5). The ND1/ND2 (upper: 5'-TTTTCTAGTACGAAAGGACC-3', lower: 5'-ATTAAAGTG(A/C/T/G)TTGA(T/G)TTGCATTC-3') and ND3/ND4 (upper: 5'-TTACGCGTATAAGTGACTTCCAA-3', lower: 5'-TTTGGGTTCCTAAGACCAA-3') mtDNA regions (primer sequences were modified from Gharrett et al. 2001) were amplified in 50 μL reactions: 1X buffer (Sigma-Aldrich® product P-2317), 2 mM $MgCl_2$, 0.2 mM each dNTP, 0.2 μM each primer, 2 units *Taq* polymerase (Sigma-Aldrich® product D-4545), and approximately 500 ng total DNA. These reactions were denatured at 95°C for 2 minutes, followed by 30 cycles at 95°C for one minute, annealed at 50 or 53°C for one minute, and extended at 72°C for three minutes, followed by a final incubation at 72°C for three minutes and 4°C for one minute.

Subsamples of the amplified product were digested with restriction endonucleases under conditions recommended by the manufacturer--*Bst*N I, *Dde* I, and *Hin*f I for ND1/ND2 and *Rsa* I for ND3/ND4. In order to resolve additional key haplotypes, mtDNA regions in a few samples were digested with additional restriction endonucleases: region COI/COII/A8 (upper: 5'-TAATCGTCACAGCCCATGCCTTCGT-3', lower: 5'-GGTCAGTTTCAGGGTTCAGGTTTAGC-3') with enzymes *Sau*96 I and *Mbo* I; ND5/ND6 (upper: 5'-AACAGCTCATCCATTGGTCTTAGG-3', lower: 5'-TTACAACG(A/G)TGGTTTTTCA-3') with enzyme *Taq* I; and Cyt*b*/D-loop (upper: 5'-GAAAAACCA(C/T)CGTTGT(A/T)ATTCAACT-3', lower: 5'-TAGGGCCTCTCGTATAACCG-3') with enzyme *Sau*96 I (primer sequences were modified from Gharrett et al. 2001). The resulting fragments were electrophoretically separated on agarose gels (1.1% Synergel™, 0.8% agarose, 1X TBE [0.09 M Tris-base, 0.09 M boric acid, and 0.002 M EDTA (pH 8.0)] buffer, stained with ethidium bromide, and digitally photographed on a UV transilluminator. Polyacrylamide gels (10.64% acrylamide/bisacrylamide [AMRESCO® Acryl/BIS 29:1], 2X TAE [0.08 M Tris-base, 0.228% glacial acetic acid, and 0.002 M EDTA (pH 8.0)] buffer, 0.072% ammonium persulfate, 0.055% TEMED) stained with SYBR™ Green I (Molecular Probes Inc., Eugene, Oregon) were used to separate small haplotype fragments from a subset of samples. Fragment sizes were estimated with ProRFLP 2.38 software (DNA ProScan Inc., Nashville, Tennessee) and 25 and 100 bp commercial DNA ladders.

Data Analyses

Microsatellite genotypes were evaluated for large allele dropout, null alleles, and stuttering to detect potential genotyping errors with the software program MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004). Multiple testing was taken into account by using a sequential Bonferroni correction (Rice 1989) in which the initial Type I error (α) was adjusted by the number of tests within each collection.

Both microsatellite and allozyme data were tested for conformance to Hardy-Weinberg expectations (GENEPOP 3.4; Rousset and Raymond 1995). An excess of homozygotes ($F_{IS} > 0$; FSTAT 2.9.3.2; Goudet 1995) can indicate population mixtures, inbreeding, or the presence of null alleles at microsatellite loci. Isolocus (Allendorf and Thorgaard 1984) data cannot be tested for Hardy-Weinberg frequencies.

Departure from gametic equilibrium can occur when loci are physically linked, in population mixtures, or as a consequence of small population size. All nine microsatellite loci and allozyme loci that had common allele frequencies less than or equal to 0.95 in at least one collection or had at least two alleles in every collection were tested for composite genotypic equilibrium by permutation testing with the program GENETIX 4.05 (http://www.genetix.univ-montp2.fr/genetix/genetix.htm). For each locus-pair in each collection, 10,000 permutations of genotypes were used to test the null hypothesis of no correlation between alleles at two loci.

A reduced median network of pink salmon mtDNA haplotypes was constructed from data observed in our study and in Churikov and Gharrett (2002) by using the program NETWORK 4.1.1.2 (www.fluxus-engineering.com; Bandelt et al. 1995).

Population Structure

Several clustering analyses for constructing population trees were used to summarize patterns of genetic variation across the collections with data for all 49 allozyme loci and the 9 microsatellite loci (PHYLIP 3.5c; Felsenstein 1989). From allele frequencies, 10 trees were created with a continuous maximum likelihood method to check the confidence of branch topology (CONTML in PHYLIP with the Jumble and Global Rearrangement options chosen). A neighbor-joining (NJ) tree was constructed from unitized chord distances (Cavalli-Sforza and Edwards 1967; NEIGHBOR in PHYLIP) between collection pairs. Haplotype frequencies derived from the mtDNA RFLP data were used to estimate chord distances between populations and then construct NJ and UPGMA trees (NEIGHBOR in PHYLIP). The strength of the branching order in the NJ trees was examined with Monte Carlo re-sampling of allele or haplotype frequencies

1,000 times with replacement in each collection (A. J. Gharrett, unpublished FORTRAN program), calculating chord distances for each re-sampling iteration, and then summarizing with a consensus tree (CONSENSE in PHYLIP).

Arcsine square-root transformed allele frequencies of variable loci were used in a principal component analysis conducted with the procedure PRINCOMP in SAS® 6.12 software (SAS Institute Inc.). Rare alleles were removed to reduce spurious results. For the allozyme loci, all of the common alleles were included, as were all but one of the alternate alleles that were observed at a relative frequency equal to being seen at least twice in a sample the size of the smallest collection (typically about 40-50 fish) at a given locus. For most microsatellite loci there was not a distinct common allele; for these loci, alleles observed at least twice in at least one collection were retained. Only the common allele was retained for the bi-allelic locus *Ots208-2*.

Homogeneity of allozyme and microsatellite allele frequencies was tested with pseudo-exact tests and individual probabilities were pooled across loci by Fisher's method (GENEPOP 3.4; Raymond and Rousset 1995; Ryman et al. 2006). Homogeneity of mtDNA haplotype frequencies was tested with pseudo-exact tests in the software Arlequin 3.1 (Excoffier et al. 2005).

Population pairwise F_{ST} values (θ in Weir and Cockerham 1984) were estimated to further examine population divergence within and between each stream age category. Kruskal-Wallis one-way ANOVA on ranks was used to test for differences in the fixation indices with two comparisons: (1) the three within-age categories—old/old, med/med, new/new, and (2) the three between-age categories—med/old, new/med, new/old (SigmaStat 3.0.1; SPSS Inc.). For these analyses, the population pairwise F_{ST} values were estimated from the combined microsatellite and allozyme locus dataset with FSTAT 2.9.3.2 (Goudet 1995), and from the mitochondrial dataset with Arlequin 3.1 (Excoffier et al. 2005). F_{ST} (unordered alleles) for the mtDNA dataset was estimated rather than Φ_{ST} (ordered alleles), because the allele order for all the haplotypes in Churikov and Gharrett (2002) and this study needed to calculate Φ_{ST} was unavailable. However, even

if allele order was available, mutation events are unimportant at the time scale of this study.

The overall F_{ST} values were also estimated for each of the three marker types for each level of population hierarchy that had been tested for homogeneity (Genepop3.4; Arlequin 3.1). This involved two analyses, one with populations categorized by stream-age and the second with populations categorized by geographic location (mid-upper Bay or lower-outer Bay). In both analyses, F_{ST} values were calculated (1) between multi-year collections within a location, (2) among locations within stream-age category or location category with multi-year collections pooled, and (3) among stream-ages or locations with populations pooled within each stream age or location.

Measures of Differences among Populations

The effect of the three stream-age (new, medium, old) or the two location categories (middle-upper, lower-outside) on genetic measures was examined with non-parametric analyses, typically Kruskal-Wallis one-way ANOVA on ranks and Mann-Whitney rank sum tests (SigmaStat 3.0.1). We anticipated comparison of the colonization mechanism of the even-year and odd-year broodlines. To standardize the results of the two analyses, we used the same suite of allozyme loci, except for analyses that required only variable loci (e.g., linkage disequilibrium).

A suite of genetic diversity measures was estimated for the nine microsatellite and the forty-one allozyme loci that were variable in one or both pink salmon broodlines. Allele richness, estimated for each locus separately to maximize sample size (Hurlbert 1971; Petit et al. 1998), and unbiased expected heterozygosity (H_E; Nei 1987) was calculated in FSTAT 2.9.3.2 (Goudet 1995). Effective number of alleles (n_{eff}) was calculated from the equation: $n_{eff} = 1/(1-H_E)$ (eq 2.21, Hedrick 2005). Observed heterozygosity (H_O) of nuclear markers, the number of composite mtDNA haplotypes, and the number of private (observed in only one collection) alleles and haplotypes were counted in each collection. Residual genotypic linkage disequilibrium in allozyme and microsatellite loci that were variable in every collection was examined with correlation coefficients (GENETIX 4.05).

Mitochondrial diversity was examined with haplotype richness (the number of composite haplotypes adjusted for sample size; Hurlbert 1971; Petit et al. 1998), and haplotype diversity (h), the haploid analog of heterozygosity (Nei 1987; eq 8.5), which was estimated with the DA program in REAP 4.0 (McElroy et al. 1992). Nucleotide diversity and divergence within and among collections were not estimated from the mtDNA data because the regions of the genome analyzed and endonucleases used were chosen to maximize detection of potential genetic variation based on Churikov (2000), and do not represent a random sample of the mitochondrial variation.

Three methods were used to examine relatedness measures in each collection from individual genotypes for which both variable allozyme and microsatellite data were available. First, relatedness estimates, r_{xy}, between all possible pairs of individuals within collections were examined with a regression-type estimator (Lynch and Ritland 1999) that was calculated with the program IDENTIX 1.1 (ftp://162.38.181.25/pub/identix.zip). Variance estimates of the r_{xy} values were examined (1) to test the hypothesis of equal variance in collections from three stream-age categories and (2) to test the hypothesis of equal variance in observed and expected relatedness values. The expected values were estimated with a permutation re-sampling procedure of alleles (1,000 re-samplings) that assumed no relatedness. Second, the maximum-likelihood relatedness estimates between pairs of individuals within collections were calculated with the program ML-RELATE (Kalinowski et al. 2006). From genotype simulation (20,000 re-samplings) and likelihood ratio analysis, the different possible relationships between pairs of individuals were evaluated and used to determine the number of individuals possibly related to at least one other individual in the collection and the number of individuals unrelated to all others. Third, a feasible genealogy was computed for each collection that maximized the correlation between the pedigree and molecular co-ancestry matrices and eliminated incongruous assignments that can occur with the pair-wise relatedness methods. The genealogy was computed with the program MOL_COANC (Fernandez and Toro 2006) by using default control file parameters after determining that a range of modifications in these parameters produced nearly identical

results for one of the datasets, Nunatak 1991. From the genealogy, the number of unique parents for each collection was estimated.

Effective population size (N_e) was estimated with two of the most widely applied methods which, when both methods could be applied, provided independent information (Waples 1991). The first was a linkage disequilibrium moments-method based on the residual non-random association of alleles at different loci in finite populations; N_e was estimated with the program LDNE (Waples and Do 2008), which corrects for small sample size bias. The second was a temporal moments-method that requires samples from more than one generation and is based on the idea that allele frequencies change over time as a function of population size due to random genetic drift in finite populations (Nei and Tajima 1981b; Beaumont 2003). For the temporal method, N_e was estimated with the program SALMONNb (Waples 1989; Waples et al. 2006). The harmonic mean of the sample size and the mean F value (standardized variance in allele frequency change) from SALMONNb, along with plausible estimates for census (N = 1000 for Wolf Pt., N = 3000 for Nunatak, N = 5000 for Tyndall, and N = 10,000 for E. Kahtaheena), were used in equation 12 of Waples (1989) to estimate N_e; 95% confidence intervals for N_e were estimated from equation 16 (Waples 1989). Alleles that had frequencies less than 0.02 were pooled with the next least abundant allele for programs LDNE and SALMONNb to provide a balance between maximizing precision and minimizing bias (R. Waples, NOAA, personal communication).

With separate analyses for the allozyme and microsatellite datasets, each population was tested for recent bottlenecks from 10,000 iterations in the program BOTTLENECK 1.2.02 (Cornuet and Luikart 1996; Piry et al. 1999). All of the microsatellite loci and the 36 variable allozyme loci in the odd-year broodline were included in these analyses. The allozyme loci were assumed to mutate under the infinite alleles model (IAM), and the microsatellite loci (except *One111*) were assumed to mutate under the two-phase model (TPM; Di Rienzo et al. 1994). Two mutation scenarios were considered to examine the robustness of the TPM. The BOTTLENECK program was run 10,000 iterations for the TPM with 95% one-step and 5% multi-step mutations, with the variance among multiple

steps of 12 as recommended by Piry et al. (1999), and then with 70% one-step and 30% multi-step mutations, with a variance among multiple steps of 30—parameter settings that bring the model closer to the IAM.

Results

Allozyme Loci

Thirty-six allozyme loci in the odd-year pink salmon collections had 2-7 alleles per locus (Appendices 1.1 and 1.2); thirteen loci were monomorphic in all collections. Four isoloci (*sAAT-1,2*; *GPI-B1,2*; *sMDH-A1,2*; and *sMDH-B1,2*) exhibited low variability; and for subsequent analyses, all variation was assigned to one locus while the other locus was designated monomorphic (Gharrett and Thomason 1987; Waples 1988). Of the variable loci, 23 loci had a common allele frequency greater than or equal to 0.95, and 13 loci had a common allele frequency less than 0.95 in at least one collection. The most common allele at each locus was observed in every collection; some additional alleles occurred in single collections, other alleles in multiple collections. A total of 106 alleles were observed at the 36 variable loci; individual collections carried 60 to 74 alleles.

Sixteen alleles were limited to populations within Glacier Bay. Of these alleles, five were infrequent and only in the youngest, upper Bay collections (*sAH*115*; *GR*77*; *LDH-A1*90*; *LDH-A2*90*; *TPI-1*-180*). Seven alleles were unique to lower Bay collections, all rare (*mAH-1*125*; *GPI-A*110*; *GPI-B1,2*1*; *LDH-B1*150* and **47*; *sMDH-B2*143*; *TPI-1*-155*). The frequencies of the remaining four alleles that were present only within Glacier Bay collections ranged between 0.05-0.035 (*CK-A1*80* and **117*; *sMDH-A2*138*; *sMDH-B2*31*). Fourteen alleles unique to the three populations outside Glacier Bay were rare and most were singletons.

Microsatellite Loci

Most of the microsatellite loci were highly allelic and a large proportion of alleles were at low frequencies (Appendices 1.1 and 1.3). All nine loci were polymorphic in every collection. Twenty-one alleles, all at low frequency, were present only in the

Glacier Bay populations and nine alleles, all at low frequency, were observed only in the populations outside Glacier Bay. Interpretation of the microsatellite data did not appear to be affected by large allele dropout or stuttering, and evidence of null alleles existed only in one collection (Spasski 1991) at *Ots208-1* (F_{IS} = 0.098).

mtDNA RFLPs

Restriction endonuclease sites were inferred from single changes in mtDNA fragment patterns (Appendices 1.1 and 1.4). All five major lineages that were previously reported (Churikov and Gharrett 2002) were observed in the odd-year pink salmon samples (Figure 1.2; Tables 1.6 and 1.7)[3]. Ambiguities in the haplotype network were resolved by choosing the haplotype relationships presented in Churikov and Gharrett (2002), which were based on data from additional mtDNA regions and restriction endonucleases.

Four of the 13 haplotypes observed were new (Fig. 1.2; Tables 1.6 and 1.7). Two of the new haplotypes, "b" and "c", occurred at low frequency in several collections; the other new haplotypes, "a" and "e" were singletons. Between five and nine haplotypes were observed in each collection. The E* and AA* haplotypes were typically the most common haplotypes, as they were in the broad-scale geographic survey of Churikov and Gharrett (2002). The E* haplotype was the most common haplotype in the collections from upper Glacier Bay. The AA* haplotype was most abundant in three of four collections in mid-lower Glacier Bay. In the collections outside Glacier Bay, the most common haplotype was split about equally between E* and AA*. All three singletons (a, BO, e) occurred in older collections. Three haplotypes were restricted to collections within Glacier Bay (c, AO, A*); the "c" haplotype was primarily in the collections from the upper half of Glacier Bay, including Tyndall.

---

[3] Composite haplotype designations from Churikov and Gharrett (2002) were retained; a "*" following a haplotype, e.g., E*, represents the pool of haplotypes that were not distinguished from the mtDNA regions and restriction endonucleases of our study. New haplotypes were given lower case letter designations.

Hardy-Weinberg Equilibrium

Only seven (3.4%) of the 205 possible Hardy-Weinberg probability tests for allozymes were significant, fewer than expected by chance (5%) given the size of tests specified ($\alpha = 0.05$). Significant tests were spread across six collections and five loci, and there was no pattern of disequilibrium for either loci or collections (Appendix 1.2). No tests were statistically significant when multiple testing was taken into account.

Departure from Hardy-Weinberg equilibrium for microsatellite loci was about what would be expected by chance (6%) and was spread across six collections and five loci (Appendix 1.3). Most of the significant tests were due to excess homozygosity; however, none of the tests were significant after correcting for multiple testing.

Gametic Disequilibrium

For microsatellite and allozyme loci combined there were 164 significant locus-pair departures from gametic equilibrium out of 3682 possible comparisons (4.5%), about the proportion expected by chance alone (5%). No pattern of equilibrium departure was obvious; the percentage of allozyme-only, microsatellite-only and allozyme-microsatellite locus-pairs out of equilibrium was similar (4.4–4.7%). The fraction of locus-pairs out of equilibrium did not differ for collections from different stream-ages ($P = 0.257$, Kruskal-Wallis one-way ANOVA on ranks); however, it was larger in the collections from the mid-upper area of Glacier Bay (4.67%) than from the lower-outside Glacier Bay locations, 3.89% ($P = 0.014$, Mann-Whitney rank sum test). After sequential Bonferroni corrections for multiple testing, only one locus-pair was out of equilibrium for one collection (*Ots208-2* and *mAH-3*; Nunatak 1993).

Population Structure

Separate population tree and principal component analyses were made for the allozyme and microsatellite locus datasets (NJ trees in Appendix 1.5) and compared before a final analysis was conducted with the combined datasets. The continuous maximum likelihood trees (Appendix 1.6), NJ trees, and the consensus trees had similar

branching patterns with two basic clusters and short distances between the branch nodes. One cluster was collections from the mid and upper Glacier Bay locations, and the other cluster included collections from the lower and outside Glacier Bay locations (Figure 1.3A). The strength of the branching order on the NJ tree can be seen in the consensus tree from the bootstrapped allele frequency datasets (Figure 1.3B). In 97% of the trees from the bootstrap sampling, the lower-outside Glacier Bay collections clustered separately from the mid-upper Glacier Bay collections.

The UPGMA and NJ trees created from the mtDNA composite haplotype frequencies were similar in structure (not shown). The lower-outside Glacier Bay cluster was identical in the two trees, whereas some branch swapping occurred within the mid-upper Glacier Bay cluster of collections. The basic topology of the two clusters in the NJ tree (Figure 1.4A) is shared with the trees derived from the nuclear markers. A consensus tree from the bootstrap sampling of the mtDNA haplotype frequencies also clustered the lower-outside Glacier Bay collections, but the strength of branch order was weaker at many nodes on the tree (Figure 1.4B), as would be expected for smaller sample sizes and haploid data.

The principal component analyses that used 189 alleles from the combined datasets corroborate the topologies of the NJ trees (Figure 1.5). The first four components explain 51.6% of the genetic variation among collections. About half of that variation is described by the first principal component, which provides the primary contrast between the collections in the mid-upper Glacier Bay area (Gull, Wolf Pt., Nunatak, and Tyndall) and those in the lower and outside Glacier Bay areas (N. Berg, N. Fingers, E. Kahtaheena, Homeshore, and Spasski). The second component pulls out Gull 1993 from the other youngest populations in upper Glacier Bay, and separates the lower Glacier Bay collections of N. Fingers and N. Berg from the collections outside Glacier Bay. The third component pulls out N. Fingers from all other collections and indicates some differences among the youngest collections, and the fourth component reveals variation among the older collections (Appendix 1.7).

There were no single heavily weighted allele contributors or any distinct break in levels of contribution of alleles to the principal components; rather, many alleles contributed similar eigenvector weights within principal components (component loadings in Appendix 1.8). Alleles from both allozyme and microsatellite loci contributed to the larger component loadings.

Homogeneity

To examine homogeneity, collections from multiple years at a location were tested first, and then collections within stream-age categories (new, medium, old) were tested after pooling data from locations that were sampled in multiple years. Finally, populations across the three stream-age categories were tested after pooling collections within each stream age. Based on the apparent split between the younger populations from Tyndall northward and the older populations from N. Fingers southward that was evident in the tree and principal component analyses, homogeneity was also tested between collections from the mid-upper Bay and the lower-outside Bay locations.

Collections from locations sampled in multiple years were homogeneous for both allozyme and microsatellite datasets and for all but the Nunatak collections for the mtDNA dataset (Table 1.8). Differences in the frequencies of the E*, AA*, and AO haplotypes contributed to the heterogeneity between the multi-year collections from Nunatak (Table 1.7). The oldest populations outside of Glacier Bay were homogeneous in all three datasets. Heterogeneity was observed among the youngest collections in the upper Bay with the microsatellite loci and mtDNA haplotypes, but not with the allozyme loci. Heterogeneity was detected among the collections from the mid-lower Bay locations at the nuclear loci but not the mtDNA, primarily from differences between Tyndall, in the middle of Glacier Bay, and N. Fingers and N. Berg, the two collections closest to the mouth of Glacier Bay. When the collections were split into two location categories (1) mid-upper and (2) lower-outside Glacier Bay, the similarity of the N. Fingers and N. Berg collections with the collections outside Glacier Bay was evident (Table 1.9). Likewise, the similarity of the Tyndall collections to the other upper Bay

collections was apparent because the inclusion of the Tyndall collections did not substantially increase the heterogeneity among the youngest collections. Whether the collections were separated by the three stream-age or the two location categories, all three marker types displayed strong heterogeneity across the study site ($P < 0.0001$).

Although allele-frequency differences were not large between the younger and older populations, the differences that did exist and the presence or absence of alleles in certain populations contributed to the heterogeneity in the odd-year pink salmon populations. For example, the *CK-A1*80* and **117* alleles were observed only in the youngest populations and may either have been derived from an unsampled donor source or exist at low frequency in the older populations. The *GPI-A*86*, *LDH-B2*128*, and *PEPA*109* alleles occurred at higher frequency in the upper Bay populations and were either absent or at very low frequency in the lower-outside Bay populations. Likewise, the rare *GPI-B1,2*33*, **1*, and **54*, and *sMDH-A1,2*1* and **90* alleles were observed only in the older populations outside Glacier Bay.

Because there was little or no temporal heterogeneity for locations with data from multiple years, population pairwise fixation indices were estimated after pooling multiple years. The mitochondrial F_{ST} values for the six population groupings indicated more divergence among the oldest populations, as compared with the medium-aged or new populations; however, the differences were not significant (left-half of Figure 1.6A; $P = 0.511$). This is in contrast to the homogeneity among the older populations for all markers, and was driven by the divergence between Spasski and Homeshore ($F_{ST} = 0.0595$). Similarly, the large divergence between Tyndall and N. Berg ($F_{ST} = 0.0319$) in the medium-aged populations increased the mean value in the pairwise comparisons in the medium-age category. The youngest populations had small pairwise mitochondrial F_{ST} values (mean $F_{ST} = 0.0051$), even though significant heterogeneity in haplotype frequencies was observed among the three youngest populations (Table 1.8). One explanation for this discrepancy is that F_{ST} is a proportionate measure of variation (heterozygosity) that does not specify the identity of the alleles or haplotypes involved, whereas homogeneity testing does.

There was more divergence between the oldest and youngest populations (mean mitochondrial F_{ST} = 0.0444) than between either the oldest and medium-aged populations (mean F_{ST} = 0.0122) or between the youngest and medium-aged populations (mean F_{ST} = 0.0349), but the differences were not significant (right-half of Figure 1.6A; P = 0.075). The relatively large mean mitochondrial F_{ST} value between the youngest and oldest populations was attributable to the Spasski population— F_{ST} values between this population and the youngest populations averaged 0.0768 (Appendix 1.9). The largest pairwise mitochondrial F_{ST} value in the medium-aged and oldest population comparisons was also attributable to Spasski (with Tyndall).

The allozyme and microsatellite locus datasets produced a similar pattern of pairwise F_{ST} values and were combined for a final analysis. The population pairwise F_{ST} values for the three within-age categories (new, medium, old) did not differ from each other (left-half of Figure 1.6B; P = 0.382). However, the F_{ST} values were higher between the population pairs that are geographically farther apart (right-half of Figure 1.6B; P = 0.001): the mean F_{ST} value for the pairwise comparisons of the youngest and oldest populations (0.0063) was about twice that of the oldest and medium-aged populations (0.0032), and the medium-aged and youngest populations (0.0027). The relatively large average F_{ST} value for the youngest and oldest population pairwise comparisons was in part due to differences between the Spasski population and the Wolf Pt. and Nunatak populations (Appendix 1.9).

The three markers gave different results with regard to the pattern of fixation indices across population age (Table 1.8). The overall F_{ST} values from the mtDNA haplotypes increased with population age, but F_{ST} values from the microsatellite loci decreased with population age. The overall F_{ST} values increased from the new to the medium-aged populations and then decreased from the medium-aged to the older populations for the allozyme loci (opposite that of the pattern observed with plants, Giles and Goudet 1997). When the populations were split into the two location categories, the allozyme and microsatellite F_{ST} values decreased from the upper-mid Bay area to the lower-outside

Bay area, and the F_{ST} values from the mtDNA haplotypes had the opposite pattern (Table 1.9).

The fixation indices were roughly an order of magnitude higher for the mtDNA haplotypes than for the microsatellite and allozyme loci. This difference is reflective of the smaller effective population size of the mtDNA genome—one-quarter that of the nuclear genome—and due to its haploid, maternally inherited nature, mtDNA variation is more susceptible to founder/bottleneck effects and genetic drift than nuclear markers (Ferris and Berg 1987) and may reveal sex-associated differences in gene flow.

When eight additional populations from across northern SE Alaska (Figure 1.7; Appendix 1.10) were compared with the populations in the Glacier Bay study by using compatible data from 28 allozyme loci (S. Hawkins, NOAA, unpublished), they clustered with the old populations outside Glacier Bay in a principal component analysis (component 1) and on a NJ tree (Figure 1.8). The three oldest populations outside Glacier Bay and the additional eight northern SE Alaska populations were homogeneous ($P = 0.250$; pseudo-exact tests). Even though the maximum geographic distances among the eight northern SE Alaska populations are more than twice as long as the distances among the populations in the Glacier Bay study, the mean population pairwise F_{ST} value among the additional northern SE Alaska populations was similar to the mean pairwise F_{ST} value of the old populations in the Glacier Bay study and old-NSE grouping (Figure 1.9). The higher mean F_{ST} values observed among the new-old and new-NSE Alaska populations (Figure 1.9, far right) would be expected in the situation where colonists come from nearby sources. The genetic similarity of the older populations outside Glacier Bay with the other populations from northern SE Alaska suggests that there are many populations that potentially contributed colonists to Glacier Bay and the three "old" populations used for comparison with the Glacier Bay populations are representative of northern SE Alaska pink salmon populations.

Testing of Population Structure

For the 204 locus-pairs for which there was variation in all collections, the correlation coefficients of linkage disequilibrium were larger in the collections within Glacier Bay than in collections outside Glacier Bay ($P < 0.001$, Kruskal-Wallis one-way ANOVA on ranks; Figure 1.10), although from a biological perspective the difference was slight—mean $R_{ij} = 0.109$ within Glacier Bay and 0.101 outside Glacier Bay.

Genetic Diversity.—Allele richness of allozyme loci did not differ across three stream-age categories ($P = 0.637$; Figure 1.11A); however, there were more microsatellite alleles in the collections from older populations ($P = 0.026$; Figure 1.11B). The older populations outside Glacier Bay averaged seven more microsatellite alleles across the nine loci than the youngest populations in the upper Bay. Heterozygosity was higher for the allozyme loci in the youngest populations ($P = 0.002$; Figure 1.12A). An opposite pattern was observed for the microsatellite loci, which had higher levels of heterozygosity in the oldest populations ($P = 0.030$; Figure 1.12B). We detected no differences in either haplotype richness ($P = 0.139$) or haplotype diversity ($P = 0.584$) in the collections across the three stream-age categories (Figure 1.13; Table 1.7). The number of private alleles and haplotypes (combined allozyme, microsatellite, and mtDNA datasets) was higher in the older populations ($P = 0.018$; Figure 1.14).

Relatedness.—The mean, median, and mode of Lynch and Ritland (1999) pairwise relatedness estimates were less than zero in all collections (Table 1.10) and did not differ by stream age ($P = 0.849$ for the mean values). The variance in average relatedness was higher in the younger collections for the observed values and marginally higher in the permuted values ($P = 0.024$ and 0.050, respectively; Figure 1.15). The observed variance of r_{xy} exceeded the permuted variance under the null hypothesis of no relatedness, particularly for the collections within Glacier Bay proper ($P = 0.001$; two-way ANOVA – the interaction of type of variance and stream age was not significant, $P = 0.155$). The observed variance of r_{xy} exceeded the expected values in the two Wolf Pt. and the 1991 Tyndall collections, and was higher than expected in all the other Glacier Bay collections (Table 1.10, percent rank of observed variance in relation to the permuted variance),

which may indicate the presence of genetically different groups of related fish in these younger populations.

The mean relatedness estimates from the maximum-likelihood method of Kalinowski et al. (2006) were near zero; the median and mode of the relatedness estimates were zero in all collections (Table 1.10). The mean r_{xy} did not differ among the three stream-age categories ($P = 0.271$); however, there was higher relatedness within collections from the mid-upper locations in Glacier Bay ($P = 0.008$) than within collections from the older populations in the lower section of and outside Glacier Bay (Appendix 1.11). Within collections, most of the fish were unrelated to each other based on their genotypes; however, the maximum-likelihood method determined that many, typically half or more, of the individuals within a collection had genotypes that were consistent with a half- or full-sib relationship to at least one other individual in the collection. The addition of data from more microsatellite loci would clarify this result: if the potential sibling relationships are real, then additional genotypes will support this; otherwise, additional genotypes will indicate non-sibling relationships. The significance of the differences in the proportion of unrelated and related individuals was ambiguous among collections from different stream ages ($P = 0.051$, chi-square), but the proportion did differ on a mid-upper Glacier Bay versus lower-outside Glacier Bay comparison ($P = 0.006$, chi-square). This difference was due primarily to the higher proportion of individuals that had genotypes consistent with being related to at least one other individual in the two Wolf Pt. collections. The number of unique parents (standardized by collection size) estimated from the genealogy for individuals within each collection (Fernandez and Toro 2006) was less in the collections from the youngest streams, ($P = 0.032$; Table 1.10, Figure 1.16). This corroborates the higher maximum-likelihood relatedness estimates in the younger populations.

Effective Population Size.—All of the N_e estimates derived from linkage disequilibrium for the combined allozyme and microsatellite datasets were large with the LDNE program (Table 1.11). A negative N_e estimate indicates that the disequilibrium can be explained by sampling error and the population cannot be distinguished from an

infinitely large population (Waples 1991). Separately, the allozyme and microsatellite datasets also estimated large effective population size for every location sampled. The allozyme-only dataset produced very large N_e estimates (not shown), larger than the microsatellite-only dataset, possibly due to many variable loci having allele frequencies near the 0 and 1 boundaries (Waples 1990). The temporal method employed by Waples (1989) and Waples et al. (2006) (SALMONNb program) also estimated large N_e for the four locations with data for two brood years. In our study, only one generation of drift was measured, which is the minimum time frame that can be used for the temporal method. For both methods, all of the 95% confidence intervals were broad, and included a minimum of 35-73 fish and a maximum of infinity. The estimation of N_e is influenced by the disequilibrium created with population mixing; when colonists come from a mixture of populations or from a population that still has some residual linkage disequilibrium, that disequilibrium will bias the estimate of N_e to values lower than expected under a migration-drift steady state (Waples 2004).

Bottlenecks.—The lack of a mode shift of the allozyme allele-frequency distributions (not shown; Luikart et al. 1998) and the presence of a heterozygosity deficiency (relative to expectations at a mutation-drift steady state) indicate that none of the populations experienced recent, severe bottlenecks (Figure 1.17A). The heterozygosity (H_E; unbiased estimate, Nei 1987) in 11 of the 13 collections was significantly less than expected (H_{EQ}) under mutation-drift steady state (Wilcoxon one-tail tests) for allozyme loci even though the infinite-allele model provides a conservative test for heterozygosity deficiency. The remaining two collections, Wolf Pt. 1991 and Gull 1993, exhibited a non-significant heterozygosity deficiency. Significant heterozygosity deficiency suggests that the populations in our study are not at mutation-drift equilibrium and may have experienced a recent expansion in population size (Luikart and Cornet 1998) or immigration (Cornuet and Luikart 1996), although Chakraborty et al. (1980) reported that many allozyme datasets have heterozygosity deficiency.

The results of the bottleneck analyses with the microsatellite dataset are less clear and depend on the mutation model invoked. When the two-phased model in the

BOTTLENECK program was run with 95% of the mutations one-step and 5% multiple-step, as recommended by Piry et al. (1999), eight of the twelve collections showed heterozygosity deficiency (Figure 1.17B). When the two-phased model was run with 70% and 30% one-step and multiple-step mutations, respectively, the mean H_{EQ} decreased and all but one collection had excess heterozygosity (not shown), a signature of a bottleneck. However, none of the Wilcoxon tests (two-tailed) for heterozygosity deficiency or excess under either mutation scenario in the two-phased model were significant and none of the collections exhibited the mode shift of the allele frequency distribution that is indicative of a severe bottleneck.

Discussion

Genetic data were used to narrow the possible ways in which pink salmon colonized freshwater habitat in Glacier Bay, Alaska, after retreat of the Neoglacial advance that culminated around 1700. Four questions were addressed in our study to examine the boundaries of possible colonization mechanisms posed in Table 1.1.

Were source populations from locations nearby or farther away?

This line of inquiry can be further split into two temporally distinct questions: (1) what were the sources of older Glacier Bay populations in the lower part of the bay about 150 years ago?; and (2) what were the sources of more recent colonization events in the upper part of Glacier Bay within the last 60 years? Our results suggest that deglaciated streams were populated by colonists from nearby locations. The populations outside Glacier Bay provided migrants to the lower Bay locations (N. Berg and N. Fingers) and these in turn contributed to the mid-Bay area (Tyndall). Migrants then moved from the mid-Bay (and possibly lower Bay) area into the upper-Bay locations (Nunatak, Wolf Pt., and Gull). There may be movement of fish amongst the youngest populations as well.

Analyses of all marker types suggested a break between the populations first established within Glacier Bay and those more recently formed farther up the Bay. The patterns of similarity/divergence shown in the NJ trees, which were supported by

homogeneity tests, showed that populations from the first two deglaciated locations, N. Berg and N. Fingers, were most similar to populations outside Glacier Bay. The medium-aged population further up the bay, Tyndall, tended to be intermediate between the two older populations in lower Glacier Bay and the younger populations farther up the Bay (Figures 1.3 and 1.4). This break in lower Glacier Bay that separated the younger populations from the older populations was also evident in the principal component analysis (Figure 1.5).

The population pairwise F_{ST} values were consistent with colonization from nearby sources (Figure 1.6). As expected with this pattern, the youngest and oldest populations, which were geographically the farthest apart, had the largest divergence. The large divergences for the new-old population grouping were driven in large part by the "old" Spasski population on northern Chichagof Island. This population is large for the region with annual escapements that range from 1,000s to >100,000 fish (K. Monagle, ADFG, unpublished data) and, unlike the other "old" locations, which were sampled near the mouths of the streams, the pink salmon from Spasski were collected about 2 kilometers upstream. Possibly more important is a temporal difference in spawning time—the Spasski samples were collected 1-3 weeks earlier than samples from any of the other locations. The peak of pink salmon spawning in Spasski Creek typically occurs about one month earlier than we observed in other systems (early September), although there is overlap in run times.

Changes in allele frequencies and an absence of alleles in new populations are the expected result of bottleneck or founder effects, where the number of colonizers is small enough that some alleles are lost or their frequencies differ from the donor source (e.g., Great Lakes pink salmon colonization event; Gharrett and Thomason 1987). Differences in allele/haplotype frequencies were observed in the Glacier Bay populations, but they were not extreme. The presence or absence of several alleles and haplotypes contributed to the heterogeneity observed among the odd-year pink salmon populations. Some mtDNA haplotypes and nuclear alleles were observed only within Glacier Bay, generally at low frequencies (Table 1.7, Appendices 1.1–1.3). The origin of those alleles remains a

question; they may exist in the older populations outside Glacier Bay at low frequencies or in populations farther away.

The populations sampled outside Glacier Bay were not necessarily the actual donor populations that provided colonists to Glacier Bay. Other potential donors could be any of the many hundreds of populations that swim past the entrance of Glacier Bay on migrations to their natal streams in northern SE Alaska (Halupka et al. 2000; Der Hovanisian and Geiger 2005). The genetic similarity of the three oldest populations outside Glacier Bay used in this study to other populations from around northern Southeast Alaska (Figures 1.8 and 1.9) supports this possibility.

Did single or multiple source populations contribute colonists?

We know that colonists were more likely to come from nearby locations, but it is less clear whether colonization involved donors from one or multiple sources. The difficulty in answering this question is due in part to the generally weak population structure of pink salmon populations in northern Southeast Alaska, which is the region that most likely provided the donor sources to Glacier Bay. Overall, the results rule out the possibility that a large number of populations contributed to the formation of the new populations in Glacier Bay, although it is likely that more than a single population was involved.

Some signals were ambiguous regarding the number of donor sources, whereas other signals were consistent with either a single donor source population (or genetically similar fish from multiple locations) or multiple, genetically divergent donor populations. For example, increased composite linkage disequilibrium in the younger populations relative to the older populations can be explained by colonists from multiple, genetically divergent sources (even if they are geographically close), e.g., Tyndall and N. Berg. However, disequilibrium would also be expected (1) if the number of colonists were small, or (2) if source populations had residual linkage disequilibrium because the signal can persist for several generations (Hedrick 2005). The first possibility can be ruled out (see the next question), but the second remains possible because there was slightly

elevated linkage disequilibrium in a couple of the medium-aged populations (Tyndall and N. Fingers). It may be difficult to distinguish the second possibility from multiple origins based only on linkage disequilibrium.

Younger populations had fewer microsatellite alleles and private alleles/haplotypes, which is consistent with a contribution from one population and reduces the possibility of a large number of populations contributing colonists (Figures 1.11B and 14). The slightly larger maximum likelihood relatedness estimates in the younger Glacier Bay populations (Appendix 1.11B) support the concept of fewer donor sources, but the higher variance of relatedness in the younger populations (Figure 1.15; Table 1.10) can be interpreted as the presence of several different groups of related individuals, which would be more likely to occur if multiple populations contributed to the formation of new populations. The low relatedness estimates indicate that few half- or full-sib relationships were detected in any population, a situation that is more likely to exist in the younger populations if they are formed with colonists from multiple populations.

We would expect heterozygosity to be lower if one population is the donor source, and higher if multiple, genetically divergent populations contribute colonists. The pattern of heterozygosities across stream age differed between the allozyme and microsatellite loci (Figure 1.12). This may be due to differences in the distribution of allele frequencies (a mathematical effect). The lower heterozygosity at allozyme loci in the older collections was attributable to more common alleles of very high frequency ($\geq$0.98 and <1.0) and a greater proportion of alleles of very low frequency (>0 and <0.02). The heterozygosity of the microsatellite loci was higher in the older populations because of an increase in the number of alleles and increases in the frequencies of less common alleles.

Heterogeneity existed among populations within Glacier Bay, but not among the oldest populations outside Glacier Bay, which is also consistent with colonists from multiple sources. Given the recent time frame in which Glacier Bay has been colonized (<200 years or <100 pink salmon generations), these populations are probably not near a migration-drift equilibrium. The lower Glacier Bay populations, N. Berg and N. Fingers, and the middle Glacier Bay population, Tyndall, are heterogeneous; if they all

contributed to the establishment of the newest populations farther up the Bay, the colonists were a mix of genetically divergent donors from nearby locations. Movement of fish among the youngest locations might further contribute to the heterogeneity present.

The interpretation of results from the bottleneck analyses regarding the number of source populations could be confounded with the number of colonists involved (see next question); however, the lack of a founder effect and a deficit of heterozygosity in all populations compared to that expected under drift-mutation steady state for the allozyme loci (but not the microsatellite loci) could be indicative of a mixture of populations and population expansion (Luikart and Cornuet 1998).

Did new populations start with small or large numbers of fish?

From our framework, we can rule out either extreme about the number of colonists involved in the formation of new populations. Neither a very small number (tens) nor a very large number (thousands) of pink salmon colonized new stream habitat in Glacier Bay. The number of colonists must have been more than a few fish because in the youngest populations there was no evidence of a bottleneck, linkage disequilibrium was limited, allozyme richness and mtDNA haplotype diversity and richness did not differ from older populations, and the effective population size (N_e) estimates were fairly large. On the other hand, massive gene flow also did not occur because microsatellite allele richness, the number of private alleles/haplotypes, and N_e estimates were lowest in the youngest populations. Also, the residual genotypic linkage disequilibrium was slightly higher (Figure 1.10) in the youngest populations, although the difference with the older populations outside the Bay from a biological standpoint was minimal.

Bottleneck analyses that compare the observed heterozygosity with that expected under mutation-drift equilibrium are best at detecting recent, severe, and permanent bottlenecks (Luikart and Cornuet 1998). We observed no evidence of a severe founder effect in the youngest Glacier Bay populations, which eliminated the possibility that streams were colonized by only a few fish.

Even though estimates of N_e may not be very accurate or precise, due in part to violation of assumptions such as no gene flow, they do provide useful insight about the numbers of fish involved in colonization. The N_e estimates for the Glacier Bay populations were typically lower than those for the populations outside Glacier Bay (Table 1.11), but they were still in the hundreds for the youngest populations, and the lower 95% confidence limit was typically 40-50 fish. If donors from multiple, genetically divergent source populations were present, as some analyses suggested, then linkage disequilibrium will downwardly bias the N_e estimates, and the real effective population sizes may be even larger than we estimated (Waples 2004). The power of the temporal and linkage disequilibrium methods to estimate effective population size increases with small effective population sizes; neither method is good at distinguishing large from very large populations, which was reflective of the negative values for some populations (Table 1.11; Waples 1989; Waples 1991).

A sufficient number of fish colonized new streams so that there was no measureable effect on mtDNA haplotype diversity or richness among the populations of the three stream-ages, even though the mtDNA genome has one-quarter the effective population size of nuclear markers and thus is affected more by population bottlenecks (Figure 1.13). On the other hand, there was approximately one less allele per microsatellite locus in the youngest populations (Figure 1.11B) and more heterogeneity within Glacier Bay than across a much larger area of northern SE Alaska (Figures 1.8 and 1.9). The heterogeneity of allele and haplotype frequencies among Glacier Bay populations and the loss of rare microsatellite alleles was evidence that massive movement (gene flow) of pink salmon did not occur during development of these populations.

Much of the genetic variation in the older populations was present in the younger populations; many fish must have been involved early in the colonization process. What little census information exists supports this conclusion, and indicates that new populations can become large quickly, within a few generations (Milner et al. 2007). Adult pink salmon were observed in Nunatak in 1985 (828 fish; Milner and Bailey 1989) during a year of large returns to Southeast Alaska (Heinl and Geiger 2005), but not in

previous surveys. The first observation of pink salmon in Wolf Pt. in 1989 and the increase in numbers of fish in the youngest streams coincided with the large increase in escapement to Southeast Alaska populations in the 1990s (Heinl and Geiger 2005). Additional fish continued to move into this system in the 1990s, although high survival of offspring may explain some of the increase in population size, as it must for the large increase in spawning run sizes that occurred in Southeast Alaska during this time period. More than 1,200 fish were counted in 1991 and the index of counts roughly doubled for the next couple of generations, exceeding 10,000 fish in 1997 (Milner et al. 2007). Less census information is available for Gull Lake. A few dozen pink salmon were first observed in this small system in 1989, with larger numbers seen in subsequent years (nearly 400 fish in 1991 and >1,000 fish in 1993; C. Soiseth, NPS, personal communication).

Did colonization occur as a one-time event or as a recurrent process?

Based on the youngest populations, i.e., Wolf Pt., Nunatak, and Gull Lake, most of the alleles and haplotypes observed in Glacier Bay appeared early in the development of the pink salmon populations. However, there was a trend of increased genetic diversity with population age. The increase in allele richness and heterozygosity of microsatellite loci, as well as the increase in the number of private alleles-haplotypes with population age, suggested that colonization of new populations is a recurrent process, rather than a one-time event. The decrease in relatedness estimates (Appendix 1.11B) and variance (Figure 1.15) with stream age, and a larger number of parents in the older populations (Figure 1.16) are also consistent with persistent gene flow.

Although recurrent immigration appears to be the source of additional alleles/haplotypes into Glacier Bay, gene flow is low enough that heterogeneity has been retained within Glacier Bay over several dozen generations. The N. Berg and N. Fingers populations in the lower section of Glacier Bay differ genetically from the population in the stream at the head of Tyndall Cove, which became ice free about 30-50 years later (Table 1.2). By using an approximate time of 40 years between the time streams become

ice free and pink salmon colonization commences (e.g., Nunatak and Wolf Pt.), we can assume Tyndall (ice free ca. 1875) was colonized in the early 1900s, and that about 40 generations of pink salmon have passed in the elapsed time between colonization and our sampling in 1991. N. Berg and N. Fingers still differ from the populations outside Glacier Bay after more than 60 generations since their colonization. The N_e estimates were fairly large in all of the populations, and allele frequency changes over time are slower in large populations than smaller populations (Hartl and Clark 1997)—perhaps this is why the differences have been retained within Glacier Bay. If the populations were large from their inception, subsequent gene flow would have limited short-term impact. Eventually, allele/haplotype frequencies will homogenize over time if gene flow exceeds the influence of drift. The older populations outside Glacier Bay were homogeneous for all three marker types (Table 1.8), and genetically similar at a suite of allozyme loci with populations across northern SE Alaska. These observations support the idea that gene flow, at least at the geographic scale of our study, may outweigh drift; and we can predict that eventually the populations within Glacier Bay will converge genetically.

Synthesis

Many of the new watersheds in Glacier Bay provide spawning and rearing habitat for Pacific salmon (Milner 1987; Milner and Bailey 1989), much as this area did prior to the Neoglacial advance of ice, according to oral history of local Tlingit people (Dauenhauer and Dauenhauer 1987; Catton 1995). Several rivers on the east side of Glacier Bay were named for their salmon resources by the Tlingit, who inhabited lower Glacier Bay hundreds of years ago before the final advance of ice, when the lower Bay was a large glacial outwash plain (Monteith et al. 2007).

Pink salmon colonization may be episodic with large populations developing over short time periods and limited subsequent immigration. Odd-year pink salmon populated recently deglaciated stream habitat in Glacier Bay relatively quickly and in large numbers with minimal founder effects, although heterogeneity may persist for many generations.

Colonizers appear to have been from nearby donor sources. Recurrent immigration into younger populations increases genetic diversity over time. The colonization of the youngest Glacier Bay streams and subsequent rapid expansion in the size of these populations coincided with the increase in pink salmon escapement in Southeast Alaska since the 1980s and the largest returns on record in the 1990s. The populations expanded rapidly either from high survival of founder offspring or recurrent immigration or both. The effective population size of pink salmon populations in Glacier Bay was relatively large very early in their formation, which is supported by limited census information (Milner and Bailey 1989; Milner 1994; Milner et al. 2007).

While populations are in their formative years, they will not be in migration-drift equilibrium because the balance between these processes takes time and may never fully equilibrate, given the environmental variation that salmon populations experience. The populations in the lower Bay that were the first to form after deglaciation still retain some signatures of colonization. Higher levels of linkage disequilibrium, an intermediate number of microsatellite alleles and private alleles/haplotypes, lower heterozygosity of microsatellite loci, an intermediate level of variance in relatedness, and differences in allozyme and microsatellite allele frequencies were present in the medium-aged populations. The youngest populations were probably maintained by fewer parents because they have an increased variance of relatedness, slightly higher linkage disequilibrium, and some reduced genetic variation, e.g., fewer microsatellite alleles. Although genetic diversity was lower in the newer populations, the difference in allele richness and heterozygosity was not large, i.e., the number of colonists was not very small. The fairly large estimates of N_e bear this out. Although most of the genetic variation appears to exist very early in the colonization process, the increase in allele richness from the newest to oldest populations suggested that some recurrent immigration is taking place, because the time periods are too short for mutation to be of significance.

The three marker types did not always provide consistent analysis outcomes, probably because of their unequal power in analyses that examine differences in genetic variation and diversity measures where there is not a strong founder effect. Significant

heterogeneity was observed among the youngest and medium-aged populations, but the three marker types were not always concordant (Table 1.8). Microsatellite loci appear to provide a more sensitive genetic diversity measure of allele richness than allozyme loci (Figure 1.11B). Their highly allelic nature with many low frequency alleles makes them more susceptible to founder effects. The younger populations in the upper Bay had fewer microsatellite alleles, but the difference with older populations was not large—approximately one less allele per microsatellite locus. The mitochondrial F_{ST} values were about an order of magnitude larger than the microsatellite and allozyme F_{ST} values, which may reflect the long-term consequence of a smaller effective population size for the mtDNA genome than for the nuclear genome, or possibly a higher stray rate for males. Even with a reduced effective population size, we did not detect a loss of genetic diversity in the mtDNA marker with colonization (Figure 1.13), although there were significant haplotype frequency differences among the youngest populations (Table 1.8).

It would be interesting to resample these systems in the future to see if the heterogeneity and linkage disequilibrium within Glacier Bay decreases over time, and whether the genetic diversity measures, such as allele richness increase. Samples from even newer systems, e.g., Stonefly Cr. in Wachusett Inlet, could be used to see whether the pattern of colonization we describe is consistent, or whether it applies only to a specific set of populations valid for a given point in time.

References

Aebersold, P. B., G. A. Winans, D. J. Teel, G. B. Milner, and F. M. Utter. 1987. Manual for starch gel electrophoresis: a method for the detection of genetic variation. NOAA Technical Report 61, National Marine Fisheries Service, Seattle.

Allendorf, F. W., and G. H. Thorgaard. 1984. Tetraploidy and the evolution of salmonid fishes. Pages 1-53 *in* B. Turner, editor. Evolutionary genetics of fishes. Plenum Press, New York.

Bandelt, H. J., P. Forster, B. C. Sykes, and M. B. Richards. 1995. Mitochondrial portraits of human populations. Genetics 141:743-753.

Beaumont, M.A. 2003. Conservation genetics. Pages 751-792 *in* D. J. Balding, M. Bishop, and C. Cannings, editors. Handbook of statistical genetics, 2nd edition, Volume 2. John Wiley & Sons, Ltd., West Sussex, England.

Boyer, S. H., D. C. Fainer, and M. A. Naughton. 1963. Myoglobin: inherited structural variation in man. Science 140:1228-1231.

Calkin, P. E. 1988. Holocene glaciation of Alaska (and adjoining Yukon Territory, Canada). Quaternary Science Reviews 7:159-184.

Catton, T. 1995. Land reborn: a history of administration and visitor use in Glacier Bay National Park and Preserve. National Park Service, Anchorage, Alaska.

Cavalli-Sforza, L. L., and A. W. F. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. Evolution 21:550-570.

Chakraborty, R., P. A. Fuerst, and M. Nei. 1980. Statistical studies on protein polymorphism in natural populations. III. distribution of allele frequencies and the number of alleles per locus. Genetics 94:1039-1063.

Churikov, D. Y. 2000. Mitochondrial DNA haplotype genealogies and population histories in the late Pleistocene: contrasts of pink salmon broodyears. Master's thesis. University of Alaska Fairbanks, Juneau.

Churikov, D., and A. J. Gharrett. 2002. Comparative phylogeography of the two pink salmon broodlines: an analysis based on a mitochondrial DNA genealogy. Molecular Ecology 11:1077-1101.

Clayton, J. W., and D. N. Tretiak. 1972. Amine-citrate buffers for pH control in starch gel electrophoresis. Journal Fisheries Research Board of Canada 29:1169-1172.

Cornuet, J. M., and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001-2014.

Dauenhauer, N. M., and R. Dauenhauer. 1987. Haa shuká, our ancestors: Tlingit oral narratives. University of Washington Press, Seattle.

Der Hovanisian, J. A., and H. J. Geiger. 2005. Stock status and escapement goals for salmon in southeast Alaska 2005. Alaska Department of Fish and Game, Special Publication No. 05-22, Douglas.

Di Rienzo, A., A. C. Peterson, J. C. Garza, A. M Valdes, M. Slatkin, and N. B. Freimer. 1994. Mutational processes of simple-sequence repeat loci in human populations. Proceedings of the National Academy of Sciences USA 91:3166-3170.

Estoup, A., P. Presa, F. Krieg, D. Vaiman, and R. Guyomard. 1993. $(CT)_n$ and $(GT)_n$ microsatellites: a new class of genetic markers for *Salmo trutta* L. (brown trout). Heredity 71:488-496.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1: 47-50.

Felsenstein, J. 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics 5: 164-166.

Fernandez, J., and A. Toro. 2006. A new method to estimate relatedness from molecular markers. Molecular Ecology 15:1657-1667.

Ferris, S. D., and W. J. Berg. 1987. The utility of mitochondrial DNA in fish genetics and fishery management. Pages 277-299 *in* N. Ryman, and F. Utter, editors. Population genetics & fishery management. University of Washington Press, Seattle.

Finney, B. P., I. Gregory-Eaves, M. S. V. Douglas, and J. P. Smol. 2002. Fisheries productivity in the northeastern Pacific Ocean over the past 2,200 years. Nature 416: 729-733.

Gharrett, A. J., A. K. Gray, and V. Brykov. 2001. Phylogeographic analysis of mitochondrial DNA variation in Alaskan coho salmon, *Oncorhynchus kisutch*. Fishery Bulletin 99:528-544.

Gharrett, A. J., and M. A. Thomason. 1987. Genetic changes in pink salmon (*Oncorhynchus gorbuscha*) following their introduction into the Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 44: 787-792.

Gharrett, A. J., R. B. Wilson, B. M. Baker, C. M. Guthrie, C. M. Kondzela, and H. R. Carlson. 1990. Preliminary report on genetic diversity of southern southeast Alaskan pink salmon populations. National Marine Fisheries Service, Northwest and Alaska Fisheries Centers Processed Report 90-03, Auke Bay, Alaska.

Gharrett, A. J., and L. A. Zhivotovsky. 2003. Migration. Pages 141-174 *in* E. M. Hallerman, editor. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland.

Giles, B. E., and J. Goudet. 1997. Genetic differentiation in *Silene dioica* metapopulations: estimation of spatiotemporal effects in a successional plant species. American Naturalist 149:507-526.

Goldthwait, R. P., F. Loewe, F. C. Ugolini, H. F. Decker, D. M. DeLong, M. B. Trautman, E. E. Good, T. R. Merrell III, and E. D. Rudolph. 1966. Soil development and ecological succession in a deglaciated area of Muir Inlet, southeast Alaska. Institute of Polar Studies Report No. 20, The Ohio State University Research Foundation, Columbus.

Goodwin, R. G. 1988. Holocene glaciolacustrine sedimentation in Muir Inlet and ice advance in Glacier Bay, Alaska, U.S.A. Arctic and Alpine Research 20:55-69.

Goudet, J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. Journal of Heredity 86:485-486.

Greig, C., D. P. Jacobson, and M. A. Banks. 2003. New tetranucleotide microsatellites for fine-scale discrimination among endangered chinook salmon (*Oncorhynchus tshawytscha*). Molecular Ecology Notes 3:376-379.

Halupka, K. C., M. D. Bryant, M. F. Willson, and F. H. Everest. 2000. Biological characteristics and population status of anadromous salmon in Southeast Alaska. General Technical Report PNW-GTR-468, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, Oregon.

Hartl, D. L., and Clark A. G. 1997. Principles of population genetics, 3rd edition. Sinauer Associates, Inc., Sunderland, Massachusetts.

Hawkins, S. L., N. V. Varnavskaya, E. A. Matzak, V. V. Efremov, C. M. Guthrie III, R. L. Wilmot, H. Mayama, F. Yamazaki, and A. J. Gharrett. 2002. Population structure of odd-broodline Asian pink salmon and its contrast to the even-broodline structure. Journal of Fish Biology 60:370-388.

Heard, W. R. 1991. Life history of pink salmon (*Oncorhynchus gorbuscha*). Pages 119-230 *in* C. Groot, and L. Margolis, editors. Pacific salmon life histories. University of British Columbia Press, Vancouver.

Hedrick, P. W. 2005. Genetics of populations, 3rd edition. Jones and Bartlett Publishers, Boston.

Heinl, S. C., and H. J. Geiger. 2005. Chapter 4: Pink salmon stock status and escapement goals in Southeast Alaska and Yakutat. Pages 153-189 *in* J. A. Der Hovanisian, and H. J. Geiger, editors. Stock status and escapement goals for salmon in Southeast Alaska 2005. Alaska Department of Fish and Game, Special Publication No. 05-22, Douglas.

Hoffman, S. H. 1982. Northern southeastern Alaska pink salmon (*Oncorhynchus gorbuscha*) tagging investigations, 1977-1980. Alaska Department of Fish and Game, Informational Leaflet No. 196, Juneau.

Holmes, R. S., and C. J. Masters. 1970. Epigenetic interconversions of the multiple forms of mouse liver catalase. FEBS Letters 11:45-48.

Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577-586.

Johnson, J., and K. Klein. 2009. Catalog of waters important for spawning, rearing, or migration of anadromous fishes—Southeastern region, effective June 1, 2009. Alaska Department of Fish and Game, Special Publication No. 09-04, Anchorage.

Kalinowski, S. T., A. P. Wagner, and M. L. Taper. 2006. ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. Molecular Ecology Notes 6:576-579.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, England.

Luikart, G., F. W. Allendorf, J. M. Cornuet, and W. B. Sherwin. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. Genetics 89:238-247.

Luikart, G., and J. M. Cornuet. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology 12:288-237.

Lynch, M., and K. Ritland. 1999. Estimation of pairwise relatedness with molecular markers. Genetics 152:1753-1766.

Mann, D. H., and T. D. Hamilton. 1995. Late Pleistocene and Holocene paleoenvironments of the north Pacific coast. Quaternary Science Reviews 14:449-471.

McElroy, D., P. Moran, E. Bermingham, and I. Kornfield. 1992. REAP—the restriction enzyme analysis package. Journal of Heredity 83:157-158.

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. John Wiley & Sons, New York.

Milner, A. M. 1987. Colonization and ecological development of new streams in Glacier Bay National Park, Alaska. Freshwater Biology 18:53-70.

Milner, A. M. 1994. Colonization and succession of invertebrate communities in a new stream in Glacier Bay National Park, Alaska. Freshwater Biology 32:387-400.

Milner, A. M., and R. G. Bailey. 1989. Salmonid colonization of new streams in Glacier Bay National Park, Alaska. Aquaculture and Fisheries Management 20:179-192.

Milner, A. M., E. E. Knudsen, C. Soiseth, A. L. Robertson, D. Schell, I. T. Phillips, and K. Magnusson. 2000. Colonization and development of stream communities across a 200-year gradient in Glacier Bay National Park, Alaska, U.S.A. Canadian Journal of Fisheries and Aquatic Sciences 57:2319-2335.

Milner, A. M., K. Monaghan, E. A. Flory, A. J. Veal, and A. Robertson. 2007. Ecological development of the Wolf Point Creek watershed; a 25-year colonization record from 1977 to 2001. Pages 3-7 *in* J. F. Piatt and S. M. Gende, editors. Proceedings of the fourth Glacier Bay science symposium. U.S. Geological Survey Scientific Investigations Report 2007-5047, Reston, Virginia.

Monteith, D., C. Connor, G. Streveler, and W. Howell. 2007. Geology and oral history—complementary views of a former Glacier Bay landscape. Pages 50-53 *in* J. F. Piatt and S. M. Gende, editors. Proceedings of the fourth Glacier Bay science symposium. U.S. Geological Survey Scientific Investigations Report 2007-5047, Reston, Virginia.

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

Nei, M., and F. Tajima. 1981a. DNA polymorphism detectable by restriction endonucleases. Genetics 97:145-163.

Nei, M., and F. Tajima. 1981b. Genetic drift and estimation of effective population size. Genetics 98:625-640.

Olsen, J. B., P. Bentzen, M. A. Banks, J. B. Shaklee, and S. Young. 2000. Microsatellites reveal population identity of individual pink salmon to allow supportive breeding of a population at risk of extinction. Transactions of the American Fisheries Society 129:232-242.

Petit, R. J., A. El Mousadik, and O. Pons. 1998. Identifying populations for conservation on the basis of genetic markers. Conservation Biology 12:844-855.

Piry, S., G. Luikart, and J-M. Cornuet. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity 90:502-503.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): population genetics 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.

Ridgway, G. J., S. W. Sherburne, and R. D. Lewis. 1970. Polymorphism in the esterases of Atlantic herring. Transactions of the American Fisheries Society 99:147-151.

Ritland, K. 2000. Marker-inferred relatedness as a tool for detecting heritability in nature. Molecular Ecology 9:1195-1204.

Rousset, F., and M. Raymond. 1995. Testing heterozygote excess and deficiency. Genetics 140:1413-1419.

Ryman, N., S. Palm, C. Andre, G. R. Carvalho, T. G. Dahlgren, P. E. Jorde, L. Laikre, L. C. Larsson, A. Palme, and D. E. Ruzzante. 2006. Power for detecting genetic divergence: differences between statistical methods and marker loci. Molecular Ecology 15:2031-2045.

Schaal, B. A., and W. W. Anderson. 1974. An outline of techniques for starch gel electrophoresis of enzymes from the American oyster *Crassostrea virginica* Gmelin. Georgia Marine Science Center, University System of Georgia, Skidaway Island.

Scribner, K. T., J. R. Gust, and R. L. Fields. 1996. Isolation and characterization of novel salmon microsatellite loci: cross-species amplification and population genetic applications. Canadian Journal of Fisheries and Aquatic Sciences 53:833-841.

Seeb, L. W., C. Habicht, W. D. Templin, K. E. Tarbox, R. Z. Davis, L. K. Brannian, and J. E. Seeb. 2000. Genetic diversity of sockeye salmon of Cook Inlet, Alaska, and its application to management of populations affected by the Exxon Valdez oil spill. Transactions of the American Fisheries Society 129:1223-1249.

Seutin, G., B. N. White, and P. T. Boag. 1991. Preservation of avian blood and tissue samples for DNA analysis. Canadian Journal of Zoology 69:82-90.

Shaklee, J. B., F. W. Allendorf, D. C. Morizot, and G. S. Whitt. 1990. Gene nomenclature for protein-coding loci in fish. Transactions of the American Fisheries Society 119:2-15.

Shaw, C. R., and R. Prasad. 1970. Starch gel electrophoresis of enzymes - a compilation of recipes. Biochemical Genetics 4:297-320.

Smith, C. T., B. F. Koop, and R. J. Nelson. 1998. Isolation and characterization of coho salmon (*Oncorhynchus kisutch*) microsatellites and their use in other salmonids. Molecular Ecology 7:1614-1616.

Smith, G. R. 1992. Introgression in fishes: significance for paleontology, cladistics, and evolutionary rates. Systematic Biology 41:41-57.

Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535-538.

Waples, R. S. 1988. Estimation of allele frequencies at isoloci. Genetics 118:371-384.

Waples, R. S. 1989. A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics 121:379-391.

Waples, R. S. 1990. Conservation genetics of Pacific salmon. III. Estimating effective population size. Journal of Heredity 81:277-289.

Waples, R. S. 1991. Genetic methods for estimating the effective population size of cetacean populations. Pages 279-300 *in* A. R. Hoelzel, editor. Genetic ecology of whales and dolphins. Report of the International Whaling Commission, Special Issue 13, Cambridge, UK.

Waples, R. S. 2004. Salmonid insights into effective population size. Pages 295-314 *in* A. P. Hendry and S. C. Stearns, editors. Evolution illuminated: salmon and their relatives. Oxford University Press, Oxford, UK.

Waples, R. S., and C. Do. 2008. LDNE: a program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8:753-756.

Waples, R. S., M. Masuda, and J. Pella. 2006. SALMONNb: a program for computing cohort-specific effective population sizes (N_b) in Pacific salmon and other semelparous species using the temporal method. Molecular Ecology Notes 7:21-24.

Warner, B. G., R. W. Mathewes, and J. J. Clague. 1982. Ice-free conditions on the Queen Charlotte Islands, British Columbia, at the height of late Wisconsin glaciation. Science 218:675-677.

Webb, E. C. 1992. Enzyme nomenclature 1992: recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzymes. Academic Press, San Diego, California.

Weir, B. S., and C. C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358-1370.

Table 1.1. Expected genetic signals for the contrasting possibilities that define the framework boundaries of the colonization mechanism of pink salmon in Glacier Bay, Alaska. Some genetic measures are confounded with or driven by other factors (*).

	Expected differences in new populations						Medium-aged and older populations (relative to younger populations)	
	Source location		Number sources		Number colonists		Colonization frequency	
	Near	Far	One	Many	Few	Many	Once	Recurrent
Between populations								
Homogeneity	similar	different						
FST	lower	higher						
Within populations								
Linkage disequilibrium	*	*	lower	higher	higher	lower	*	*
Allele-haplotype richness	*	*	lower	higher	lower	higher	lower	higher
Private alleles-haplotypes	*	*	lower	higher	lower	higher	lower	higher
Heterozygosity or effective number alleles	*	*	lower	higher	lower	higher	lower	higher
Effective population size	*	*	*	*	lower	higher	*	*
Relatedness	*	*	higher	lower	*	*	higher	lower
Bottleneck signal	*	*	*	*	yes	no	*	*

Table 1.2. Location of sample collections of odd-year pink salmon within and near Glacier Bay, Alaska. Water body number (Johnson and Klein 2009) is given below the location name. Years before present (BP). The name of the nearest geographic feature or the colloquial name was used for watersheds without official names. Abbreviated names used in figures in parentheses.

Location	Latitude	Longitude	*n*	Collection date	Stream age[a]
Youngest populations, upper Glacier Bay, systems ice-free less than 60 years					
Wolf Point (WO)	59.00	-136.16	62	4, 5 Sep 1991	1947
114-77-11000			100	30 Aug 1993	
Nunatak (NU)	58.98	-136.10	100	3 Sep 1991	1935
114-77-10350			100	31 Aug 1993	
Gull Lake outlet (GU)	58.95	-136.29	100	1, 2 Sep 1993	1955
114-77-10900					
Medium-aged populations, mid-lower Glacier Bay, systems ice-free less than 150 years					
Tyndall (TY)	58.59	-136.36	100	7 Sep 1991	ca. 1875
114-73-10150			99	5 Sep 1993	
Fingers Bay (NF) north head of north arm	58.60	-136.23	42	13 Sep 1991	ca. 1845
114-72-10170					
N. Berg Bay (NB)	58.52	-136.24	100	6 Sep 1993	ca. 1830
114-71-10320					
Oldest populations, outside Glacier Bay, systems ice-free during Neoglaciation					
E. Kahtaheena (EK)	58.41	-135.57	100	30 Aug 1991	>13,000 BP
114-23-10240			100	8 Sep 1993	
Homeshore (HM)	58.31	-135.37	100	7 Sep 1993	>13,000 BP
114-25-10100					
Spasski (SP)	58.09	-135.28	100	20 Aug 1991	>13,000 BP
114-27-10300					

[a]Estimated date stream mouth exposed from deglaciation (G. Steveler and C. Soiseth, National Park Service, personal communication).

Table 1.3. Protein coding loci analyzed in odd-year pink salmon within and near Glacier Bay, Alaska. Enzyme numbers (Webb 1992) and nomenclature (Shaklee et al. 1990), the tissue(s) and buffer(s) in which they were resolved, level of variability observed, and F_{ST} values.

Enzyme	Enzyme Number	Locus	Tissue[a]	Buffer[b]	Level of Variability[c]	F_{ST}[d]
Aspartate aminotransferase	2.6.1.1	*sAAT-1,2**	H,M	7,3	2,2	n/a
		*sAAT-3**	E	9	3,3	0.007
		*sAAT-4**	L	8	3,3	0.001
		*mAAT-1**	H,M	7,3	2,2	0.001
		*mAAT-2**	L	4	2,3	0.001
Aconitate hydratase	4.2.1.3	*mAH-1**	H	6,7	2,2	n/a
		*mAH-2**	H	6,7	1,2	n/a
		*mAH-3**	H,M	6,7,5	3,2	0.017
		*mAH-4**	H,M	6,7,5	3,3	0.016
		*sAH**	L	4	2,2	-0.002
Alanine aminotransferase	2.6.1.2	*ALAT**	M	2	2,1	0.003
Creatine kinase	2.7.3.2	*CK-A1**	M	1	3,1	0.006
		*CK-A2**	M	1	1,1	n/a
		*CK-B**	E	9	1,1	n/a
		*CK-C1**	E	9	2,3	-0.001
		*CK-C2**	E	9	2,2	n/a
Formaldehyde dehydrogenase	1.2.1.1	*FDHG**	L	8	3,2	0.004
Fumarate hydratase	4.2.1.2	*FH**	M	5	2,3	n/a
Glycerol-3-phosphate dehydrogenase	1.1.1.8	*G3PDH-1**	M	2,3	3,3	-0.003
Glucose-6-phosphate isomerase	5.3.1.9	*GPI-A**	H,M	1	2,1	0.006
		*GPI-B1,2**	H,M	1	2,3	0.001
Glutathione reductase	1.6.4.2	*GR**	E	5	2,1	n/a
Isocitrate dehydrogenase	1.1.1.42	*mIDHP-1**	H	7	1,2	n/a
		*mIDHP-2**	H	7	2,1	0.002
L-Lactate dehydrogenase	1.1.1.27	*LDH-A1**	M	1	2,2	n/a
		*LDH-A2**	M	1	2,1	n/a
		*LDH-B1**	E,H,M	1	2,2	n/a
		*LDH-B2**	E,H,L,M	1,4	2,2	0.009
		*LDH-C**	E	9	1,1	n/a
Malate dehydrogenase	1.1.1.37	*sMDH-A1,2**	L	8	2,3	0
		*sMDH-B1,2**	L,M	8,2,3	3,2	0.001

Table 1.3. Continued.

Malic enzyme	1.1.1.40	*mMEP-1**	H,M	7,3,5	3,3	0.002
Mannose-6-phosphate isomerase	5.3.1.8	*MPI**	H	2	2,2	0.016
Peptidase:						
Cytosol non-specific dipeptidase (glycyl-leucine)	3.4.-.-	*PEPA**	H,M	2	3,2	0.013
Tripeptide aminopeptidase (leucyl-glycyl-glycine)	3.4.-.-	*PEPB-1**	E,H	9,1	2,3	0.003
Proline dipeptidase	3.4.13.9	*PEPD-2**	H,M	2	3,3	-0.001
Leucyl-tyrosine peptidase	3.4.-.-	*PEPLT**	H,M	2	3,3	0
Phosphogluconate dehydrogenase	1.1.1.44	*PGDH**	L,M	4,3	3,3	-0.001
Phosphoglucomutase	5.4.2.2	*PGM-2**	H,M	6,1,3	3,2	0.011
Pyruvate kinase	2.7.1.40	*PK-2**	E	5	2,1	-0.001
Superoxide dismutase	1.15.1.1	*sSOD-1**	H,L	2,8	1,3	n/a
Triose-phosphate isomerase	5.3.1.1	*TPI-1**	E	9	2,1	0.001
		*TPI-2**	E	9	1,2	n/a
		*TPI-3**	E	9	1,2	n/a
		*TPI-4**	E	9	1,1	n/a
Over all loci						0.003

[a]M = muscle; H = heart; L = liver; E = eye.

[b]1 = lithium hydroxide (Ridgway et al. 1970)
2 = Tris-EDTA-borate (Boyer et al. 1963)
3 = amine-citrate, pH 6.1 (Clayton and Tretiak 1972)
4 = amine-citrate, pH 6.5 (Clayton and Tretiak 1972)
5 = Tris-citrate, pH 7.0 (Shaw and Prasad 1970)
6 = Tris-citrate discontinuous (Schaal and Anderson 1974)
7 = amine-citrate-EDTA, pH 7.2 (modified from Clayton and Tretiak 1972); 2 mL 0.04M nicotinamide adenine dinucleotide added to heart gels
8 = amine-citrate-EDTA, pH 7.4 (modified from Clayton and Tretiak 1972)
9 = Tris-glycine (Holmes and Masters 1970).

[c]First number for odd-year, second number for even-year broodlines
1 = monomorphic
2 = low (frequency of most prevalent allele > 0.95)
3 = high (frequency of most prevalent allele ≤ 0.95).

[d]Weir and Cockerham (1984) theta; n/a, no variation or alternate alleles present in only one collection.

Table 1.4. Microsatellite loci, repeat type, annealing temperature, number of alleles (n_a), allele size range, unbiased expected heterozygosity (H_E), F_{ST}, and primer source used for odd-year pink salmon collected within and near Glacier Bay, Alaska.

Locus	Type	Temperature °C	n_a	Size range	H_E	F_{ST}[a]	Source[b]
One102	tetra-	63	25	211 - 307	0.925	0.005	1
One103	tetra-	58	38	169 - 317	0.957	0.003	1
One109	tetra-	63	18	126 - 194	0.907	0.006	1
One111	tetra-	58	19	189 - 269	0.676	0.006	1
One µ*13*	di-	50	3	141 - 147	0.144	-0.001	2
Ots208-1	tetra-	50	29	123 - 235	0.932	0.004	3
Ots208-2	di-	50	2	291 - 293	0.496	-0.001	3
µ*Sat60*	di-	59	7	107 - 121	0.443	0.004	4
Oki10	di-, tetra-	63	56	179 - 333	0.947	0.003	5
Over all loci						0.004	

[a] Weir and Cockerham (1984).

[b] 1 = Olsen et al. (2000)
2 = Scribner et al. (1996)
3 = Greig et al. (2003)
4 = Estoup et al. (1993)
5 = Smith et al. (1998).

Table 1.5. Mitochondrial DNA restriction sites examined in odd-year pink salmon as shown in Figure 1.2.

Site	mtDNA region	Restriction endonuclease	Position of binary code[a]
1	ND1/ND2	*Bst* N I	1
2	ND1/ND2	*Bst* N I	3
3	ND1/ND2	*Dde* I	1
4	ND1/ND2	*Dde* I	2
5	ND1/ND2	*Dde* I	3
6	ND1/ND2	*Hin* f I	1
7	ND1/ND2	*Hin* f I	2
8	ND1/ND2	*Hin* f I	3
9	ND3/ND4	*Rsa* I	1
10	ND3/ND4	*Rsa* I	2
11	COI/COII/A8	*Sau* 96 I	1
12	COI/COII/A8	*Sau* 96 I	2
13	COI/COII/A8	*Mbo* I	1
14	COI/COII/A8	*Mbo* I	1,2
15	COI/COII/A8	*Mbo* I	3
16	ND5/ND6	*Taq* I	1
17	ND5/ND6	*Taq* I	2
18	Cyt*b* /D-loop	*Sau* 96 I	1
19	Cyt*b* /D-loop	*Sau* 96 I	2
20	Cyt*b* /D-loop	*Sau* 96 I	1,2

[a]See Appendix 1.4.

Table 1.6. Odd-year pink salmon composite mtDNA haplotypes derived from haplotypes of specific mtDNA regions cut with restriction endonucleases (see Appendix 1.4). Composite haplotypes denoted in upper case letters match Churikov and Gharrett (2002); those in lowercase letters indicate new haplotypes. Letters in parentheses denote presumed haplotypes.

Composite	Region	ND1/ND2			ND3/ND4	COI/COII/A8		ND5/ND6	Cyt*b*/D-loop
haplotype[a]	Enzyme	*Bst* N I	*Dde* I	*Hin* f I	*Rsa* I	*Sau* 96 I	*Mbo* I	*Taq* I	*Sau* 96 I
A*		a	a	a	a	(a)	(a)	(a)	a
E*		a	a	a	b	b	(a)	(a)	(a)
F*		a	a	a	b	c	(a)	(a)	(a)
M*		a	a	b	a	(a)	(a)	(a)	(b)
AA*		b	a	b	(a)	(a)	a	a	(a)
AG*		b	a	b	(a)	(a)	b	a	(a)
AM*		b	b	b	(a)	(a)	(a)	(a)	(a)
AO		b	c	b	(a)	(a)	(a)	(a)	(a)
BO		b	a	b	(a)	(a)	c	a	(a)
a		a	a	d	a	(a)	(a)	(a)	(a)
b		a	d	a	a	(a)	(a)	(a)	(a)
c		d	a	a	c	b	(a)	(a)	(a)
e		b	a	b	(a)	(a)	e	a	(a)

[a]A* = A,G,H,I,V,AQ,BE,BF
E* = E,K,L,AR,BL,BN,BQ
F* = F,Q
M* = M,N,O,AP,BJ,BK,BZ
AA* = AA,AD,AE,AF,AJ,AT,AW,BA,BC,BD,BR,BT,CA
AG* = AG,AH,AI,BP
AM* = AM,AN,AS, BY,CC.

Table 1.7. Number of odd-year pink salmon samples (n), composite mtDNA haplotype frequencies, number of haplotypes per collection (n_c), haplotype richness (h_{rich}), and haplotype diversity (h). Upper case letters match Churikov and Gharrett (2002)[a]; lowercase letters indicate new haplotypes.

Location	n	A*	E*	F*	M*	AA*	AG*	AM*	AO	BO	a	b	c	e	n_c	h_{rich}	h
Youngest populations, upper Glacier Bay																	
Wolf Pt. 1991	40	0	0.450	0	0.050	0.250	0	0	0	0	0	0.125	0.125	0	5	4.999	0.7192
Wolf Pt. 1993	40	0	0.400	0	0.025	0.200	0.025	0.075	0.100	0	0	0.075	0.100	0	8	7.900	0.7872
Nunatak 1991	39	0.077	0.564	0	0.051	0.077	0.103	0	0.026	0	0	0.051	0.051	0	8	7.974	0.6680
Nunatak 1993	40	0	0.425	0	0	0.325	0.075	0	0.125	0	0	0	0.050	0	5	4.999	0.7077
Gull Lake 1993	39	0	0.538	0	0.051	0.231	0.154	0	0	0	0	0	0.026	0	5	4.974	0.6464
Medium-aged populations, mid-lower Glacier Bay																	
Tyndall 1991	40	0.150	0.250	0.025	0.125	0.300	0.025	0.025	0	0	0	0.025	0.075	0	9	8.800	0.8218
Tyndall 1993	40	0.025	0.475	0	0.075	0.150	0.100	0.025	0.025	0	0	0.025	0.100	0	9	8.800	0.7423
N. Fingers 1991	39	0	0.385	0	0.103	0.410	0.026	0.026	0	0	0	0.051	0	0	6	5.949	0.6869
N. Berg 1993	39	0.026	0.231	0.026	0.051	0.436	0.154	0.026	0.026	0	0	0	0.026	0	9	8.872	0.7463
Oldest populations, outside Glacier Bay																	
E. Kahtaheena 1991	40	0	0.275	0.025	0.050	0.550	0	0.050	0	0.025	0	0.025	0	0	7	6.847	0.6308
E. Kahtaheena 1993	39	0	0.385	0	0.103	0.308	0.077	0.051	0	0	0	0.051	0	0.026	7	6.974	0.7544
Homeshore 1993	39	0	0.487	0	0.051	0.359	0.026	0.051	0	0	0	0.026	0	0	6	5.949	0.6437
Spasski 1991	38	0	0.211	0	0.079	0.395	0.237	0.053	0	0	0.026	0	0	0	6	6.000	0.7539

[a]A* = A,G,H,I,V,AQ,BE,BF
E* = E,K,L,AR,BL,BN,BQ
F* = F,Q
M* = M,N,O,AP,BJ,BK,BZ
AA* = AA,AD,AE,AF,AJ,AT,AW,BA,BC,BD,BR,BT,CA
AG* = AG,AH,AI,BP
AM* = AM,AN,AS, BY,CC.

Table 1.8. Fixation indices and probability of pseudo-exact tests used to test homogeneity of odd-year pink salmon collections within and near Glacier Bay, Alaska, for three stream-age categories. Based on nuclear locus allele frequencies and mtDNA haplotype frequencies. Probability values for the nuclear loci were pooled across loci with Fisher's method.

	Allozyme		Microsatellite		mtDNA	
Location	F_{ST}	P	F_{ST}	P	F_{ST}	P
Youngest populations, upper Glacier Bay						
Wolf Pt. 1991	-0.0024	0.9997	-0.0011	0.5099	-0.0085	0.2425
Wolf Pt. 1993						
Nunatak 1991	-0.0018	0.9991	0.0015	0.3025	0.0470	0.0083
Nunatak 1993						
Gull Lake 1993						
Among New[a]	0.0005	0.1761	0.0030	<0.0001	0.0050	0.0106
Medium-aged populations, mid-lower Glacier Bay						
Tyndall 1991	-0.0026	0.9834	0.0022	0.1978	0.0367	0.1102
Tyndall 1993						
N. Fingers 1991						
N. Berg 1993						
Among Medium[a]	0.0027	0.0009	0.0026	0.0003	0.0210	0.0791
Oldest populations, outside Glacier Bay						
E. Kahtaheena 1991	-0.0003	0.9921	0.0003	0.1375	0.0328	0.1654
E. Kahtaheena 1993						
Homeshore 1993						
Spasski 1991						
Among Old[a]	0.0010	0.8645	0.0006	0.3997	0.0217	0.0688
Among Stream Age[b]	0.0038	<0.0001	0.0031	<0.0001	0.0211	<0.0001

[a]Multi-year collections pooled within locations.

[b]Populations pooled within stream-age.

Table 1.9. Fixation indices and probability of pseudo-exact tests used to test homogeneity of odd-year pink salmon collections within and near Glacier Bay, Alaska, for two location categories (mid-upper and lower-outer). Based on nuclear locus allele frequencies and mtDNA haplotype frequencies. Probability values for the nuclear loci were pooled across loci with Fisher's method.

	Allozyme		Microsatellite		mtDNA	
Location	F_{ST}	P	F_{ST}	P	F_{ST}	P
Youngest populations, mid-upper Glacier Bay						
Wolf Pt. 1991	-0.0024	0.9997	-0.0011	0.5099	-0.0085	0.2425
Wolf Pt. 1993						
Nunatak 1991	-0.0018	0.9991	0.0015	0.3025	0.0470	0.0083
Nunatak 1993						
Gull Lake 1993						
Tyndall 1991	-0.0026	0.9834	0.0022	0.1978	0.0367	0.1102
Tyndall 1993						
Among mid-upper Bay[a]	0.0010	0.0993	0.0027	<0.0001	0.0072	0.0043
Oldest populations, lower-outer Glacier Bay						
N. Fingers 1991						
N. Berg 1993						
E. Kahtaheena 1991	-0.0003	0.9921	0.0003	0.1375	0.0328	0.1654
E. Kahtaheena 1993						
Homeshore 1993						
Spasski 1991						
Among lower-outer Bay[a]	0.0006	0.2145	0.0017	0.0100	0.0091	0.1473
Between mid-upper and lower-outer[b]	0.0054	<0.0001	0.0035	<0.0001	0.0349	<0.0001

[a]Multi-year collections pooled within locations.

[b]Populations pooled within location category.

Table 1 10 Pairwise relatedness coefficients (r_{xy}) of odd-year pink salmon from moment and maximum-likelihood (ML) estimators The observed variance of r_{xy} is ranked with the permuted variance of r_{xy} from the allele resampling procedure Number (n) of fish whose genotypes were consistent with being either unrelated or related (half- or full-sib) to at least one other fish within respective collections The number of parents required to produce the genotypes present within each collection based on genealogies reconstructed from the molecular marker data

	Lynch and Ritland (1999)				Kalinowski et al (2006)			Fernandez and Toro (2006)		
		Obs	Perm	% Rank		*n* individuals		Collection	*n*	Parent collection
Location	Mean r_{xy}[a]	var r_{xy}	var r_{xy}	obs var	ML mean r_{xy}[b]	Unrelated	Related	size	parents	ratio
Youngest populations upper Glacier Bay										
Wolf Pt 1991	-0 0199	0 0061	0 0049	> highest	0 0340	8	43	51	39	0 76
Wolf Pt 1993	-0 0213	0 0060	0 0051	> highest	0 0327	9	39	48	32	0 67
Nunatak 1991	-0 0205	0 0056	0 0050	99 9	0 0321	16	34	50	36	0 72
Nunatak 1993	-0 0202	0 0056	0 0052	94 6	0 0323	13	37	50	37	0 74
Gull Lake 1993	-0 0204	0 0053	0 0045	99 9	0 0351	16	34	50	37	0 74
Medium-aged populations mid-lower Glacier Bay										
Tyndall 1991	-0 0205	0 0059	0 0049	> highest	0 0350	12	38	50	41	0 82
Tyndall 1993	-0 0179	0 0051	0 0047	97 0	0 0346	18	39	57	43	0 75
N Fingers 1991	-0 0249	0 0051	0 0044	99 1	0 0307	21	20	41	35	0 85
N Berg 1993	-0 0204	0 0048	0 0045	91 2	0 0314	15	35	50	41	0 82
Oldest populations outside Glacier Bay										
E Kahtaheena 1991	-0 0205	0 0045	0 0045	51 5	0 0316	21	29	50	39	0 78
E Kahtaheena 1993	-0 0205	0 0046	0 0044	72 5	0 0311	15	35	50	38	0 76
Homeshore 1993	-0 0203	0 0041	0 0043	23 2	0 0303	23	27	50	43	0 86
Spasski 1991	-0 0205	0 0053	0 0047	99 0	0 0329	10	40	50	38	0 76

[a] Median and mode were less than zero for all collections

[b] Median and mode were zero for all collections

Table 1.11. Estimates of effective population size (N_e) based on nine microsatellite loci and the variable allozyme loci for odd-year pink salmon collections within and near Glacier Bay, Alaska. Estimates represent the N_e of the previous parental brood-year. 95% confidence intervals are given in parentheses.

	SALMONNb[a]	LDNE[b]
Location	Temporal	Linkage disequilibrium
Youngest populations, upper Glacier Bay		
Wolf Pt. 1991		303 (37 - ∞)
Wolf Pt. 1993	240 (55 - ∞)	585 (41 - ∞)
Nunatak 1991		472 (38 - ∞)
Nunatak 1993	943 (73 - ∞)	951 (44 - ∞)
Gull Lake 1993		-42209 (45 - ∞)
Medium-aged populations, mid-lower Glacier Bay		
Tyndall 1991		436 (46 - ∞)
Tyndall 1993	141 (45 - ∞)	416 (49 - ∞)
N. Fingers 1991		994 (35 - ∞)
N. Berg 1993		4541 (46 - ∞)
Oldest populations, outside Glacier Bay		
E. Kahtaheena 1991		-1316 (45 - ∞)
E. Kahtaheena 1993	268 (57 - ∞)	2228 (42 - ∞)
Homeshore 1993		-1156 (50 - ∞)
Spasski 1991		524 (39 - ∞)

[a] Waples et al. (2006).

[b] Waples and Do (2008).

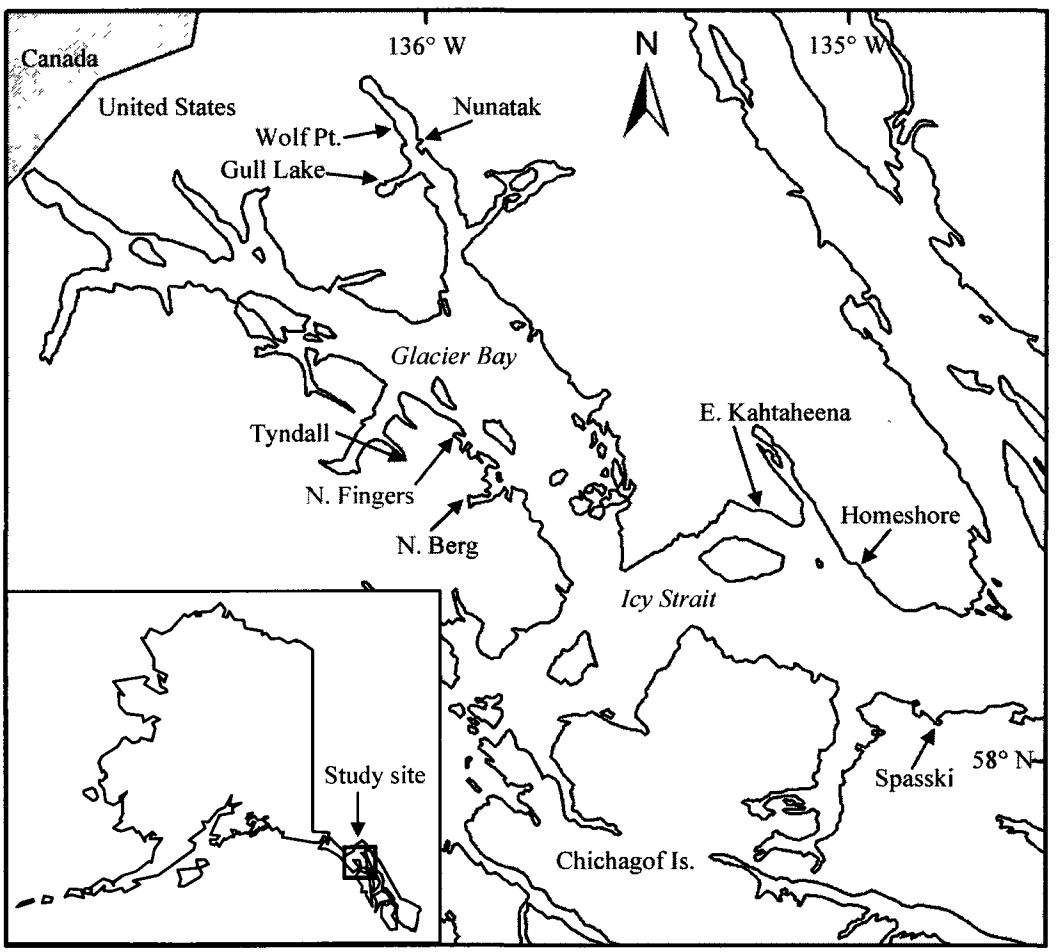


Figure 1.1. Location of odd-year pink salmon collections within and near Glacier Bay, Alaska.

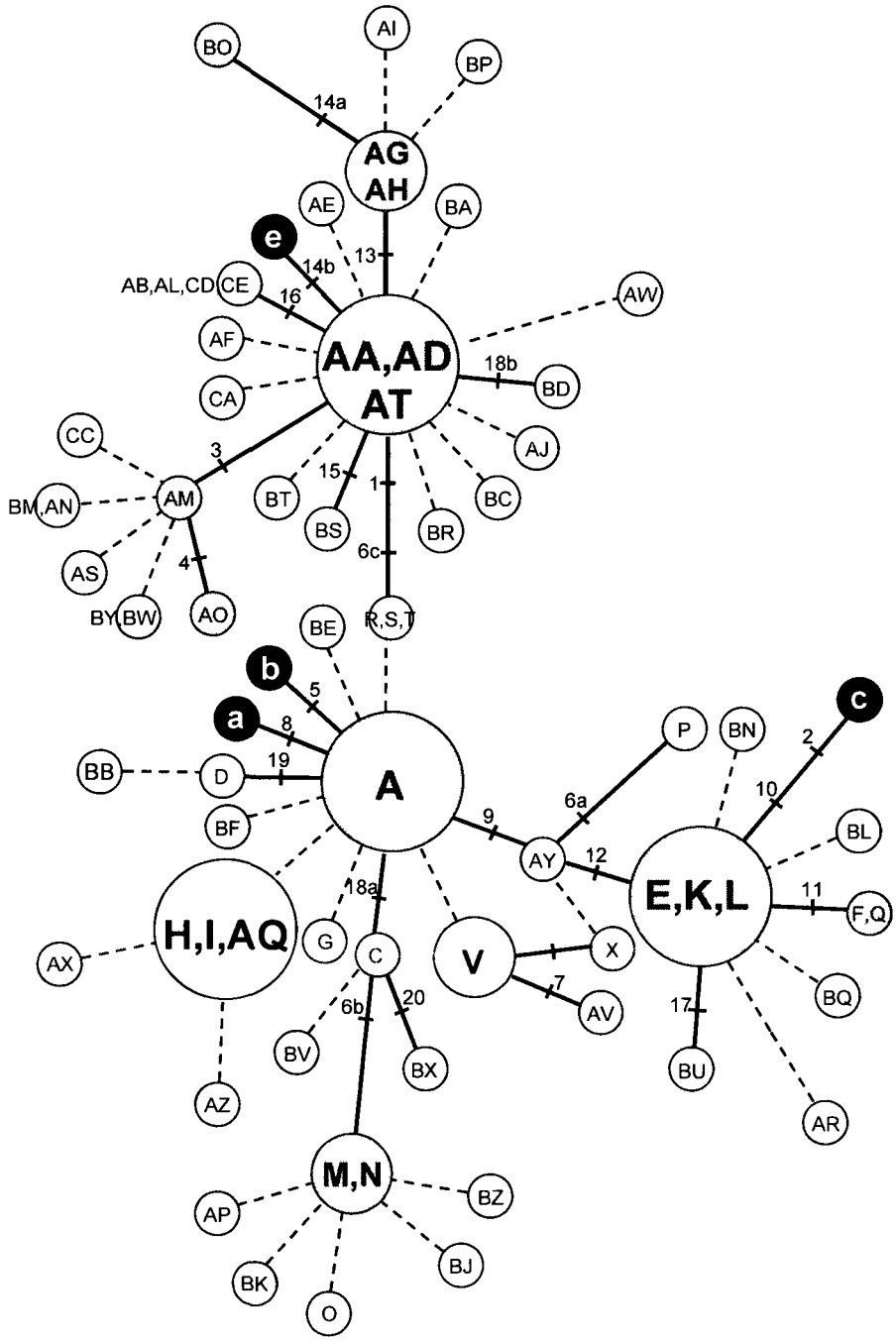


Figure 1.2. Restriction endonuclease sites examined in odd-year pink salmon. Observed haplotypes in gray and black (new) circles. Dashed lines indicate sites not examined. Details in Tables 1.5 and 1.6.

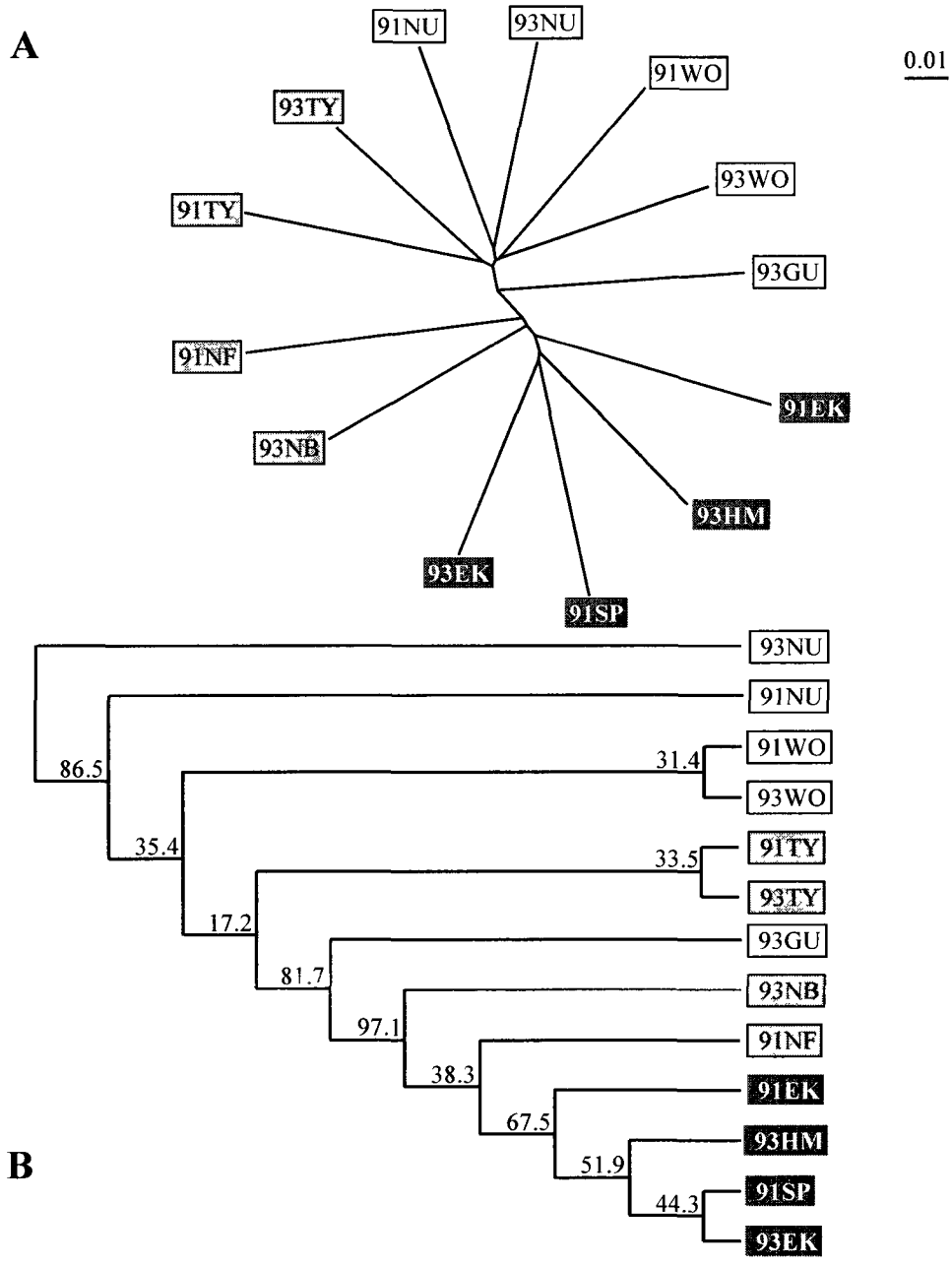


Figure 1.3. Neighbor-joining trees from nuclear allele frequencies for odd-year pink salmon populations within and near Glacier Bay, Alaska. Trees made with chord distances (Cavalli-Sforza and Edwards 1967) from 9 microsatellite and 49 allozyme loci: **(A)** neighbor-joining tree, and **(B)** consensus neighbor-joining tree from 1,000 bootstrap re-samplings of allele frequencies. Numbers at nodes indicate percentage the group occurred among the trees.

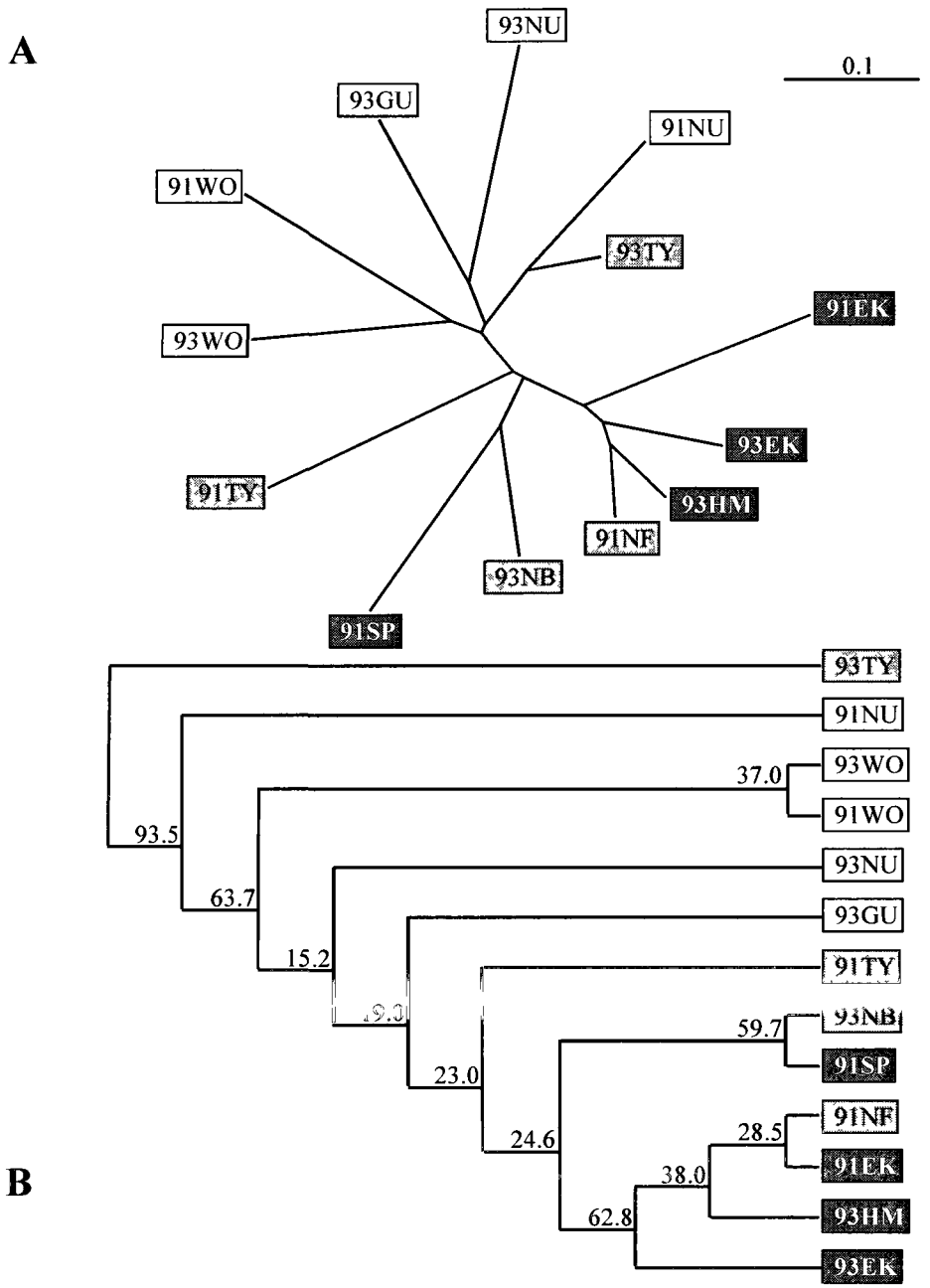


Figure 1.4. Neighbor-joining trees from mtDNA haplotype frequencies for odd-year pink salmon populations within and near Glacier Bay, Alaska. Trees made with chord distances (Cavalli-Sforza and Edwards 1967): **(A)** neighbor-joining tree, and **(B)** consensus neighbor-joining tree from 1,000 bootstrap re-samplings of haplotype frequencies. Numbers at nodes indicate percentage the group occurred among the trees.

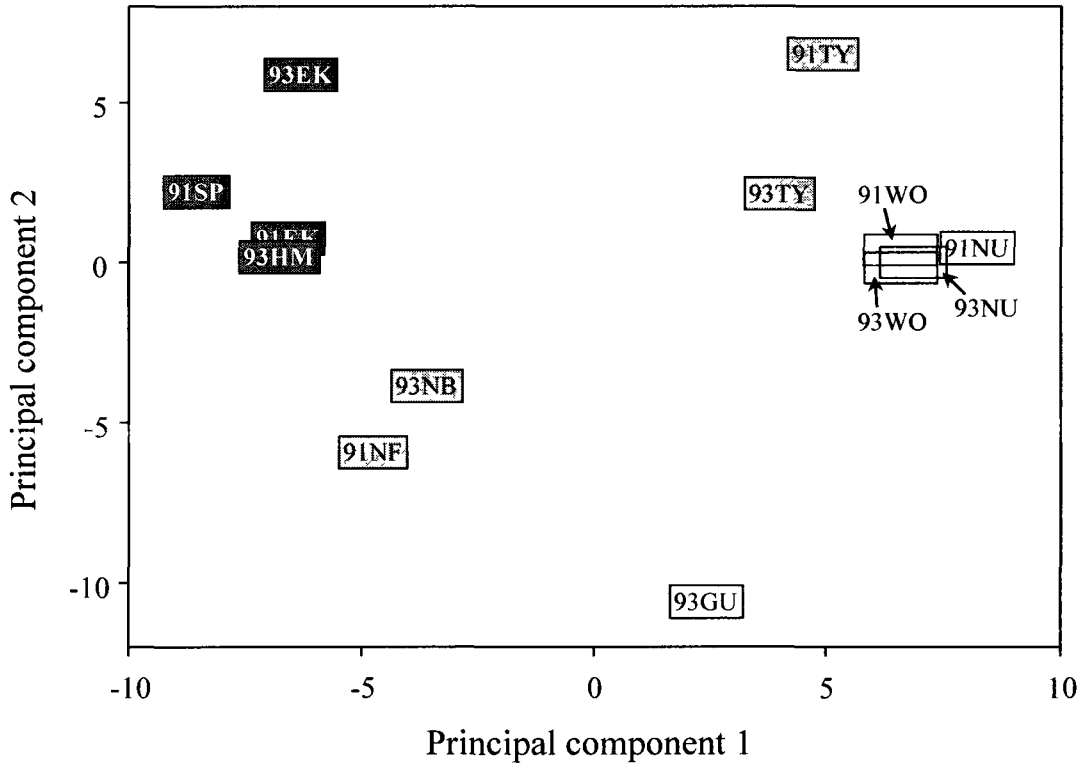


Figure 1.5. The first two principal components based on 51 allozyme and 138 microsatellite allele frequencies for the odd-year pink salmon collections within and near Glacier Bay, Alaska.

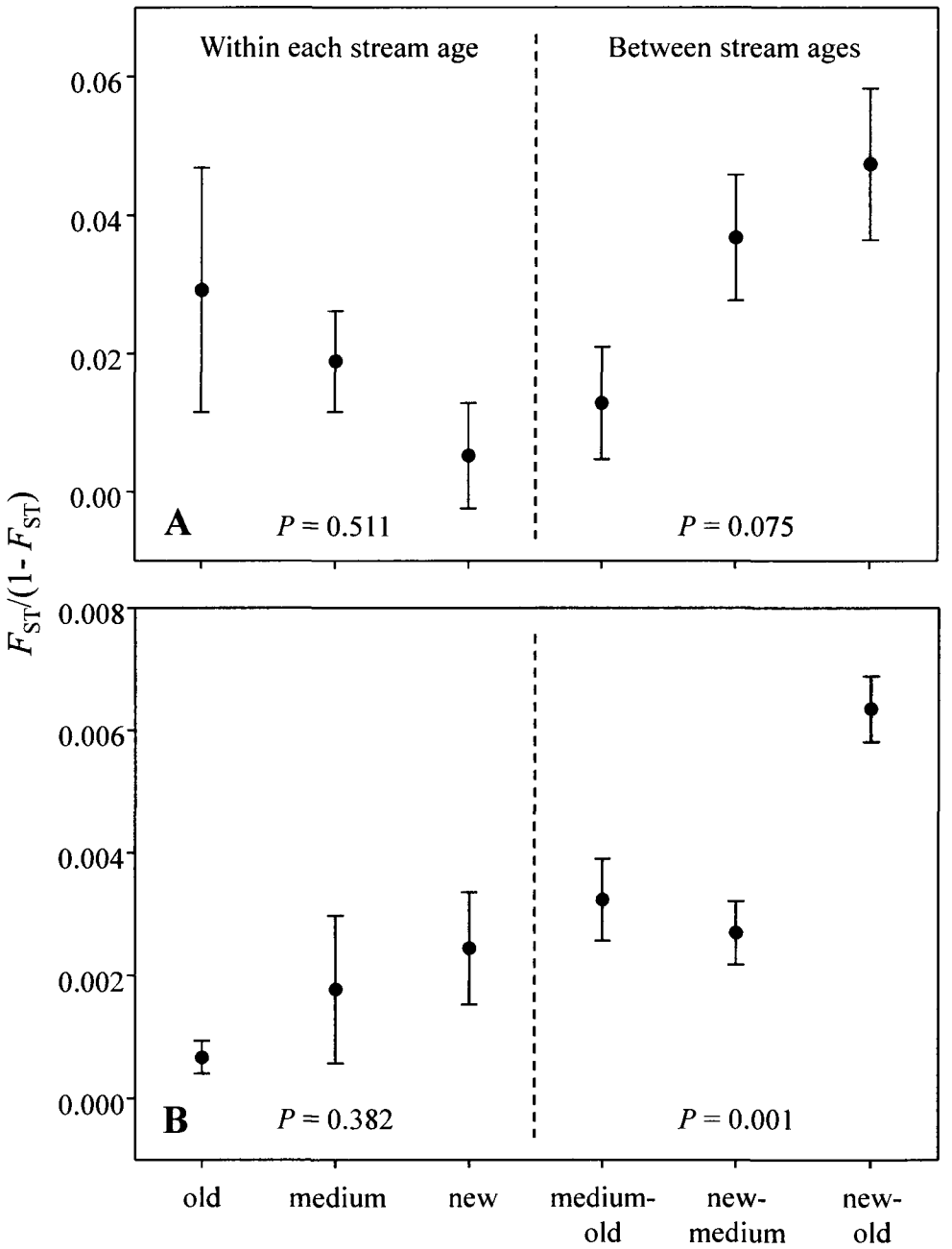


Figure 1.6. Pairwise fixation indices for odd-year pink salmon populations from three stream-age categories. Mean and SE for **(A)** mtDNA and **(B)** microsatellite and allozyme loci. Multiple year collections pooled within location. Within stream-age categories on the left; between stream-age categories on the right. *P*-values from Kruskal-Wallis ANOVA on ranks.

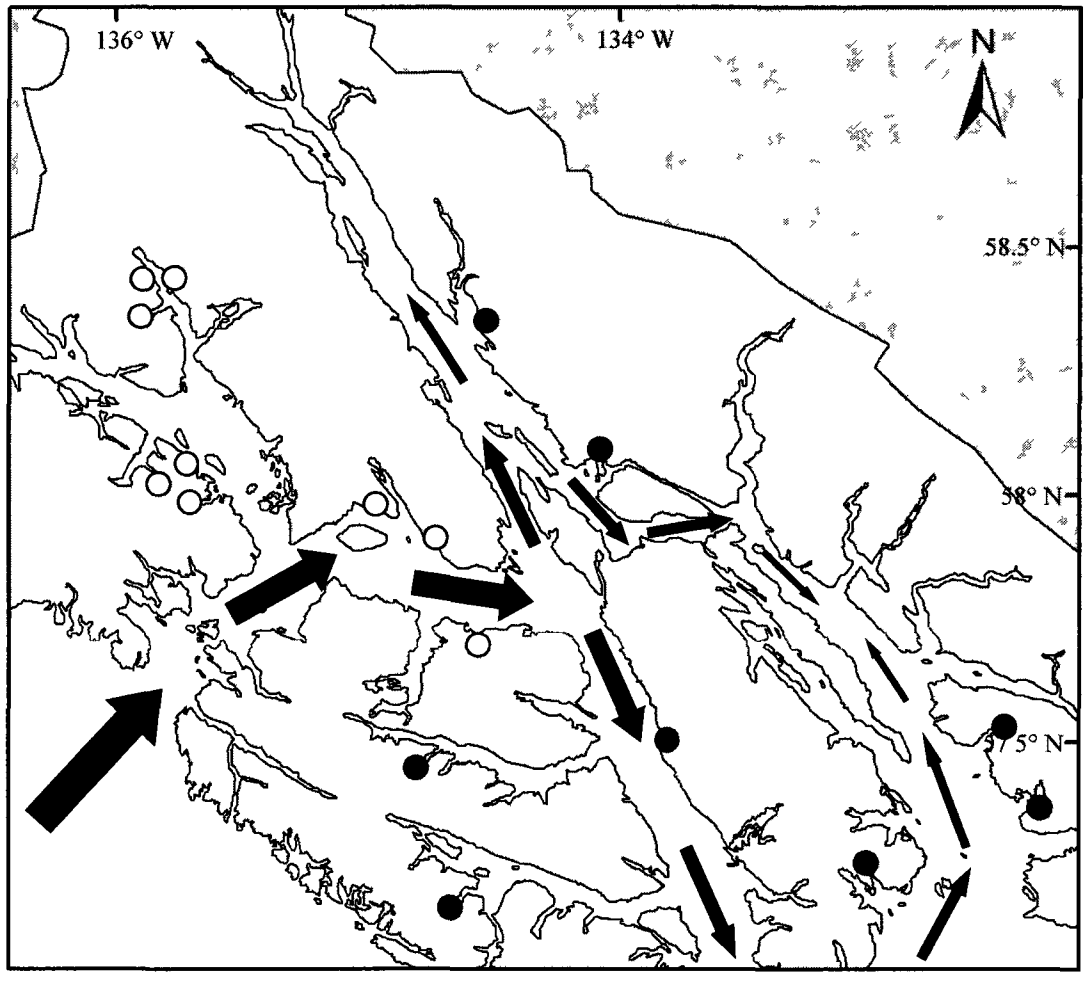


Figure 1.7. Location of odd-year pink salmon collections: (1) within and near Glacier Bay, Alaska (open circles), and (2) from eight additional populations from northern Southeast Alaska for comparison (solid circles). Arrows indicate major migration routes for salmon returning to coastal waters (Hoffman 1982).

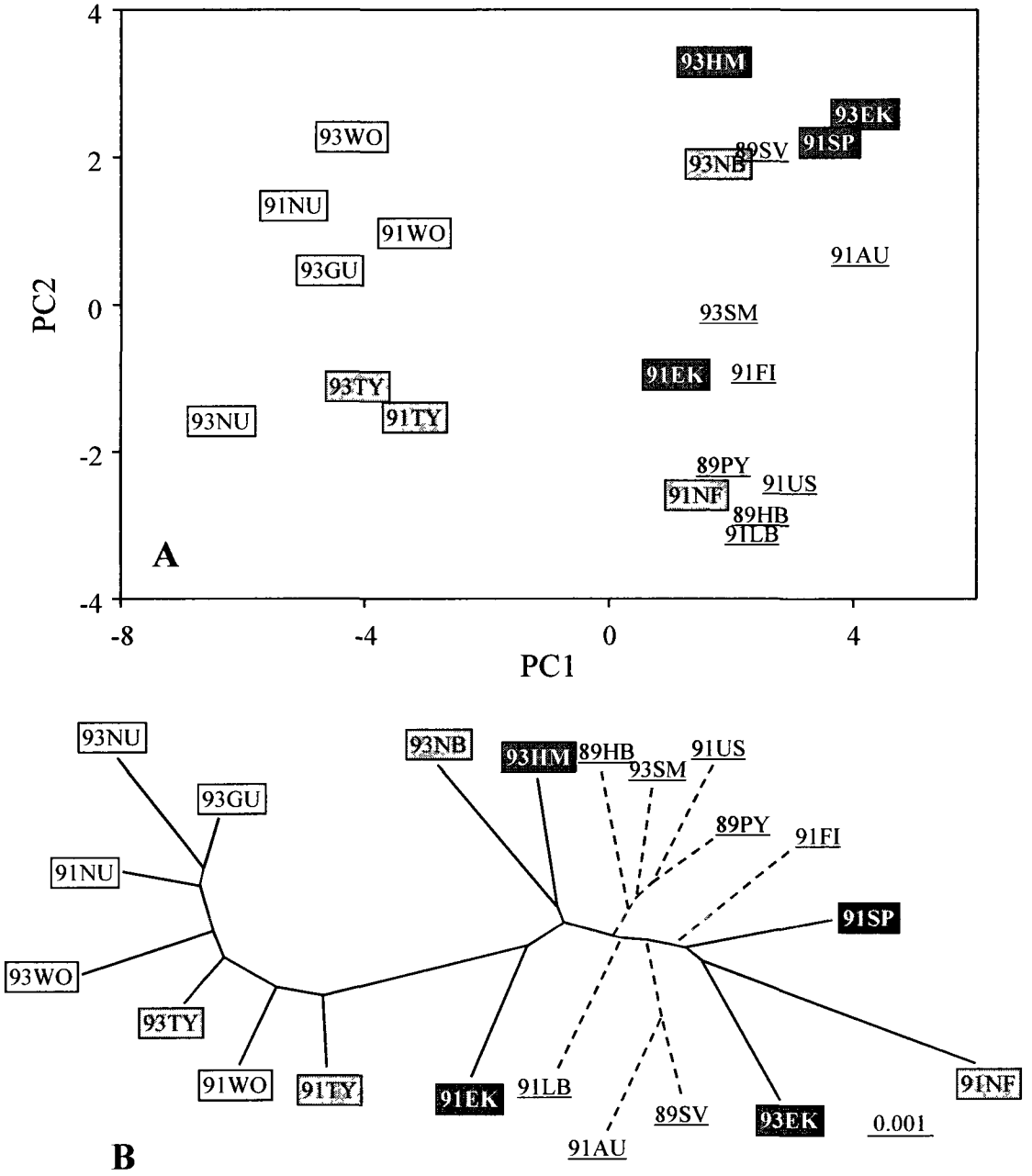


Figure 1.8. Population structure analyses for odd-year pink salmon populations within and near Glacier Bay, Alaska, and eight additional populations (dashed lines/underlined) from northern Southeast Alaska based on 28 allozyme loci. **(A)** principal component analysis, and **(B)** neighbor-joining tree from chord distances (Cavalli-Sforza and Edwards, 1967).

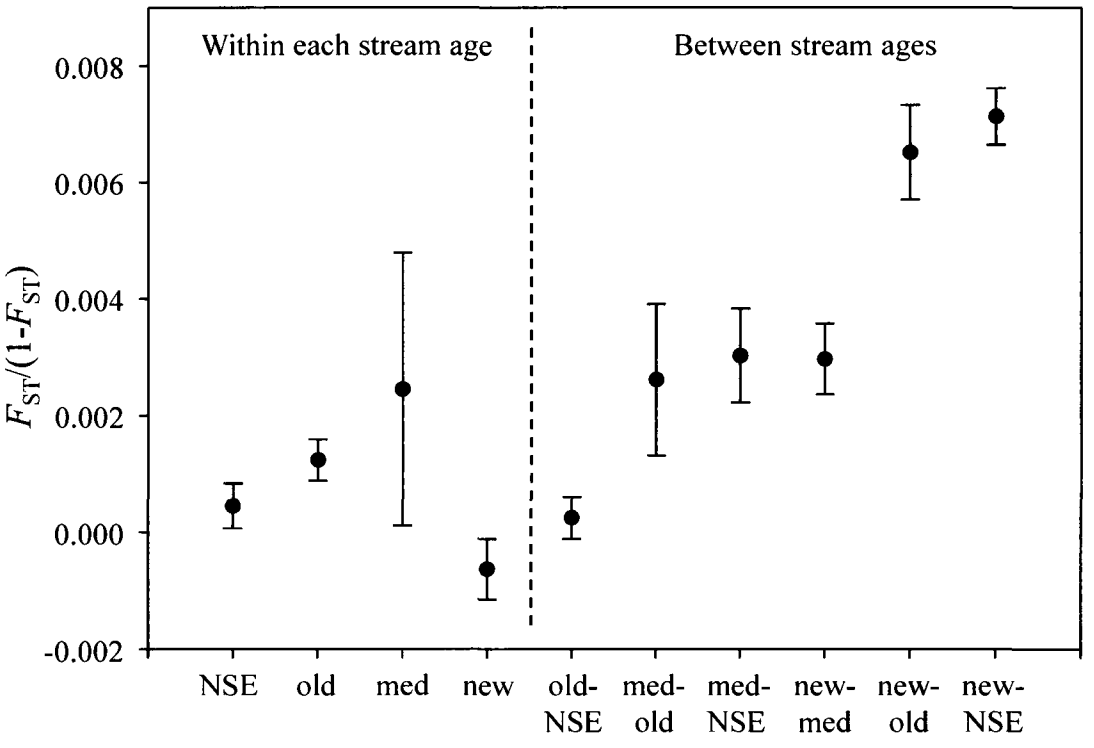


Figure 1.9. Pairwise fixation indices for odd-year pink salmon populations within and near Glacier Bay, and eight populations from northern Southeast Alaska from 28 variable allozyme loci. Mean and SE for three stream-age categories—old, medium (med), new; multiple year collections pooled within location. Within stream-age categories on the left; between stream-age categories on the right.

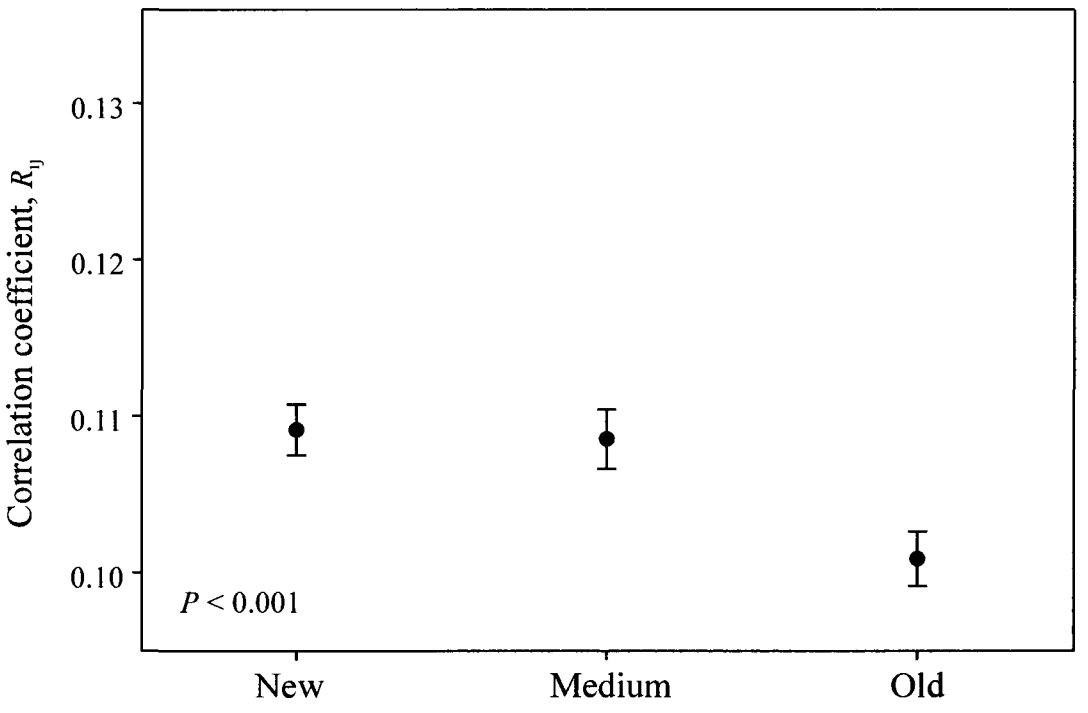


Figure 1.10. Linkage disequilibrium correlation coefficients for 204 locus-pairs for odd-year pink salmon collections from three stream-age categories (mean, SE). *P*-values from Kruskal-Wallis ANOVA on ranks.

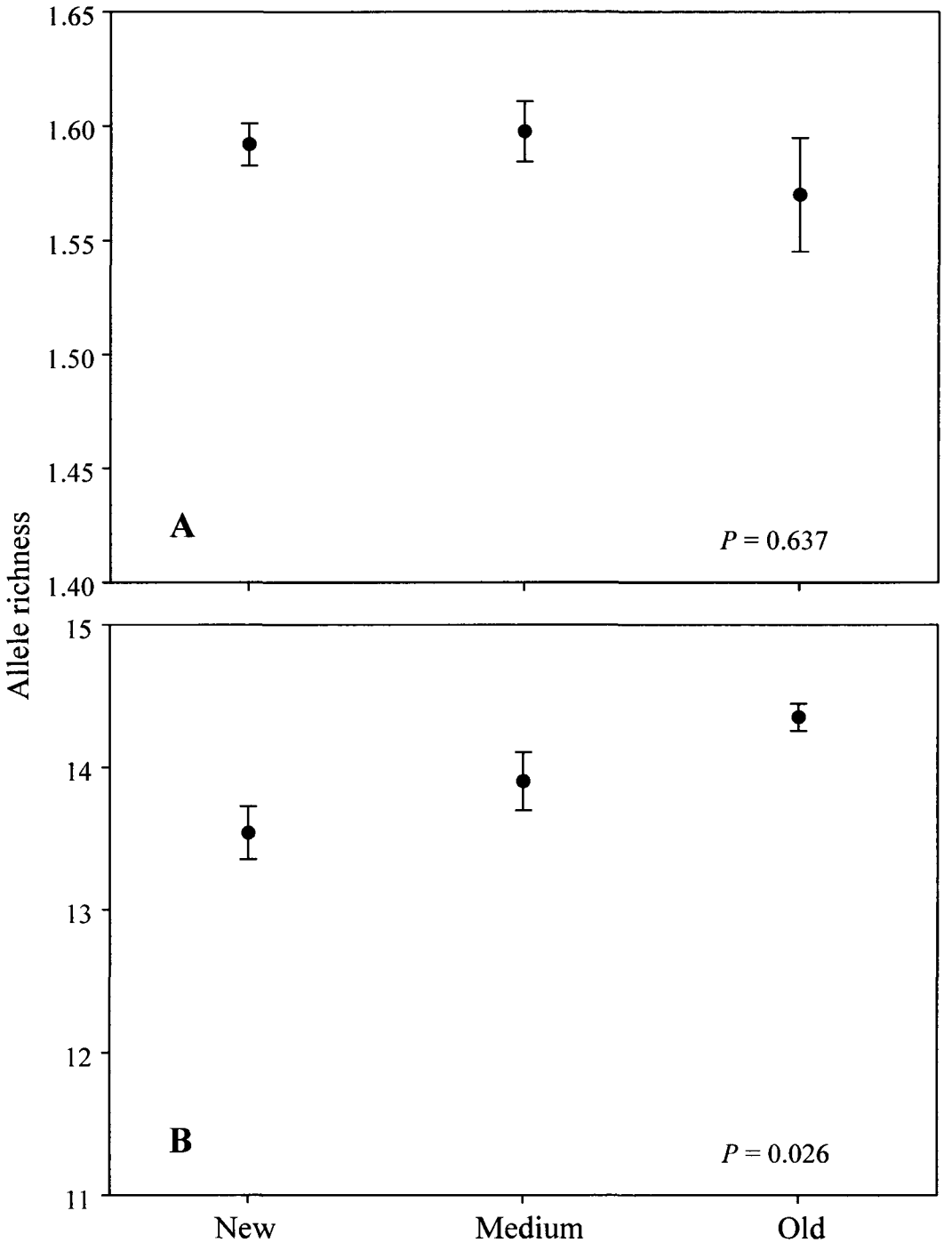


Figure 1.11. Allele richness per locus for odd-year pink salmon collections from three stream-age categories. Mean and SE for **(A)** 41 allozyme loci and **(B)** 9 microsatellite loci. *P*-values from Kruskal-Wallis ANOVA on ranks.

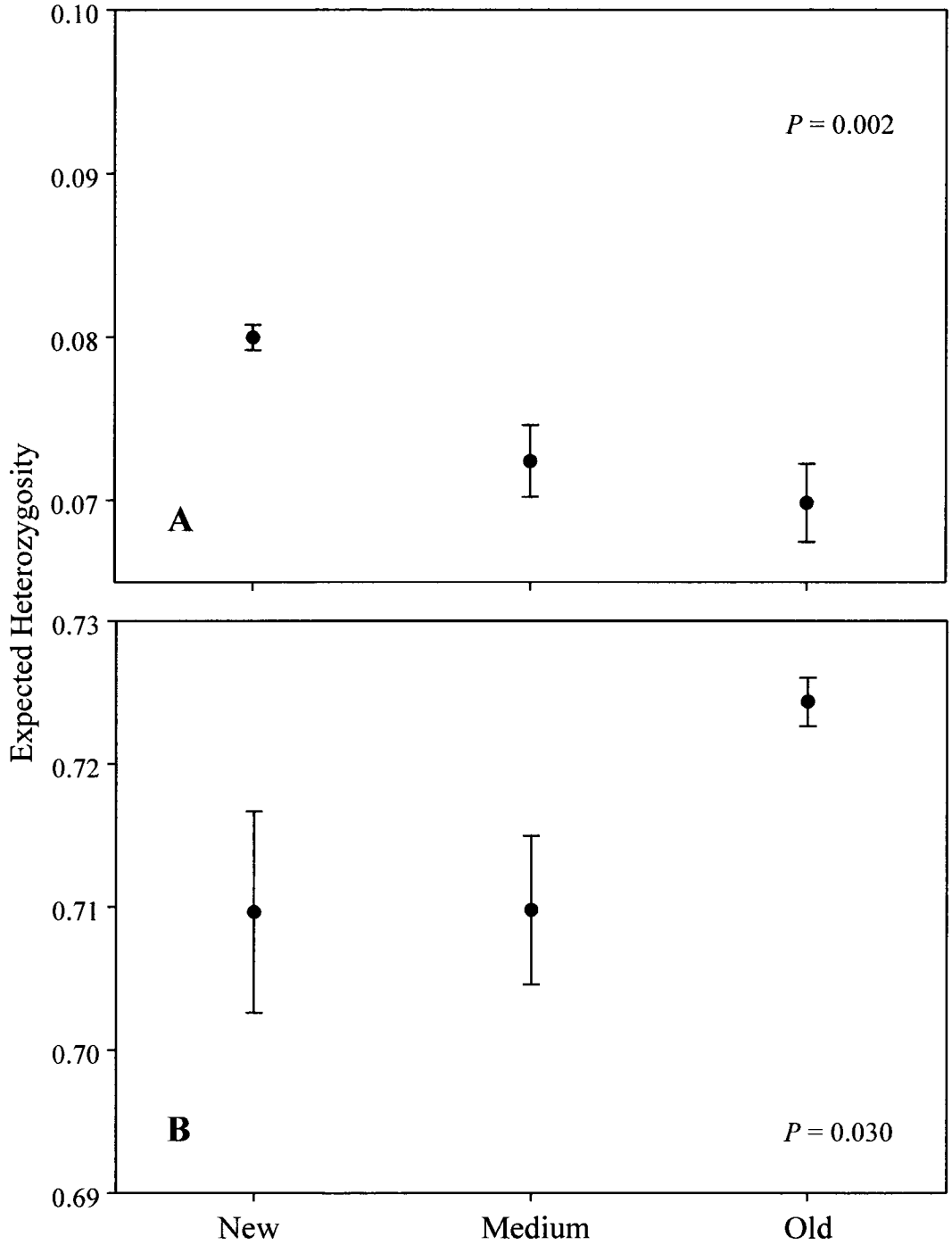


Figure 1.12. Expected heterozygosity for odd-year pink salmon collections from three stream-age categories. Mean and SE for **(A)** 41 allozyme loci and **(B)** 9 microsatellite loci. *P*-values from Kruskal-Wallis ANOVA on ranks.

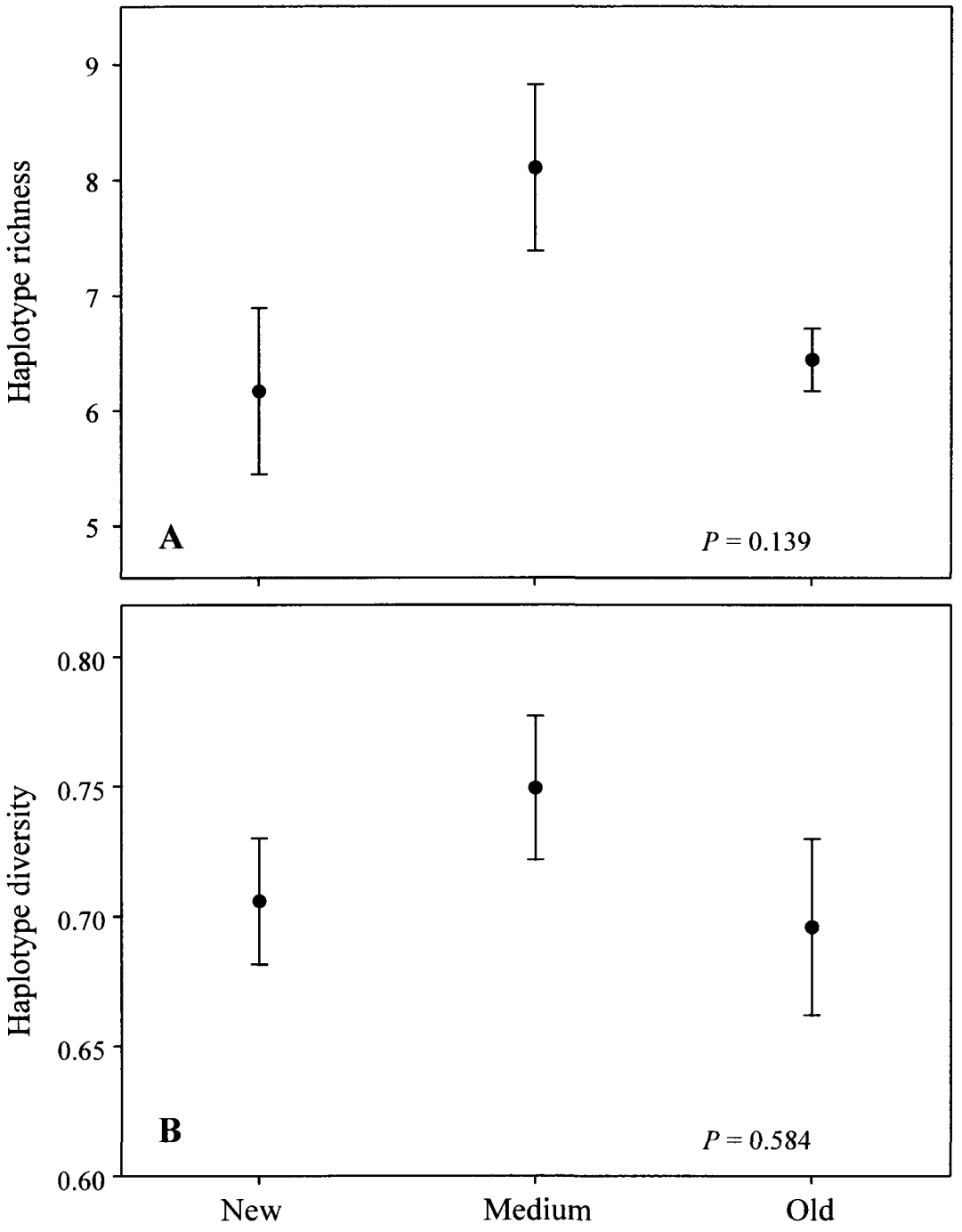


Figure 1.13. Measures of mtDNA variation for odd-year pink salmon collections from three stream-age categories. Mean and SE for **(A)** haplotype richness and **(B)** haplotype diversity. *P*-values from Kruskal-Wallis ANOVA on ranks.

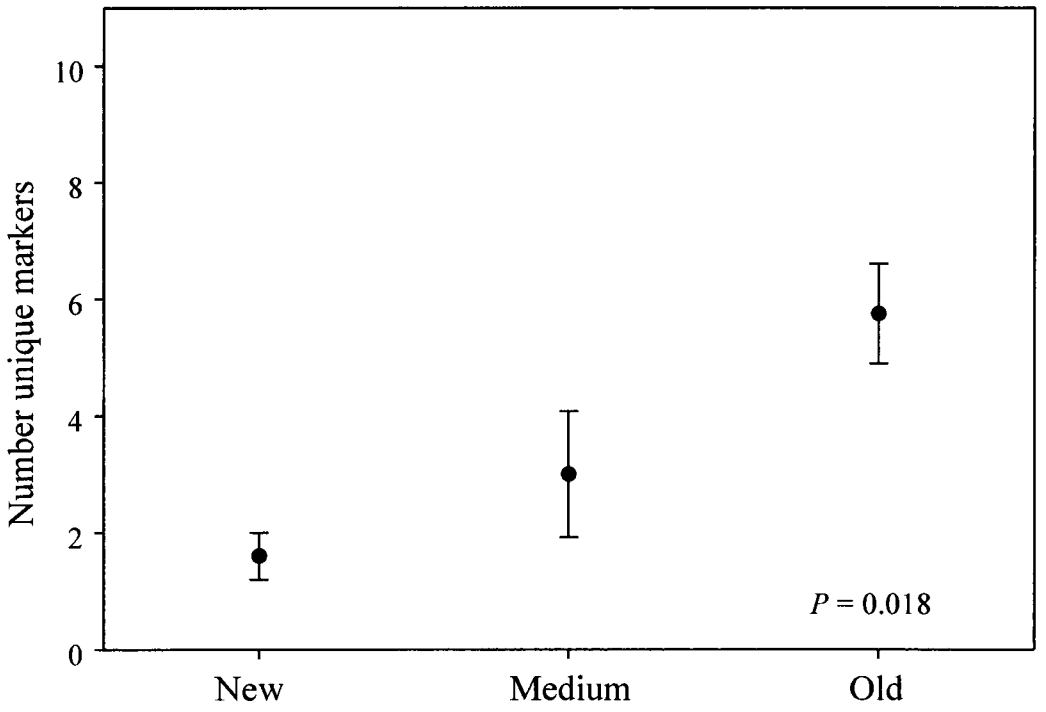


Figure 1.14. Number of private alleles and composite haplotypes in the microsatellite, variable allozyme, and mtDNA datasets for odd-year pink salmon collections. Mean and SE from three stream-age categories; *P*-value from Kruskal-Wallis ANOVA on ranks.

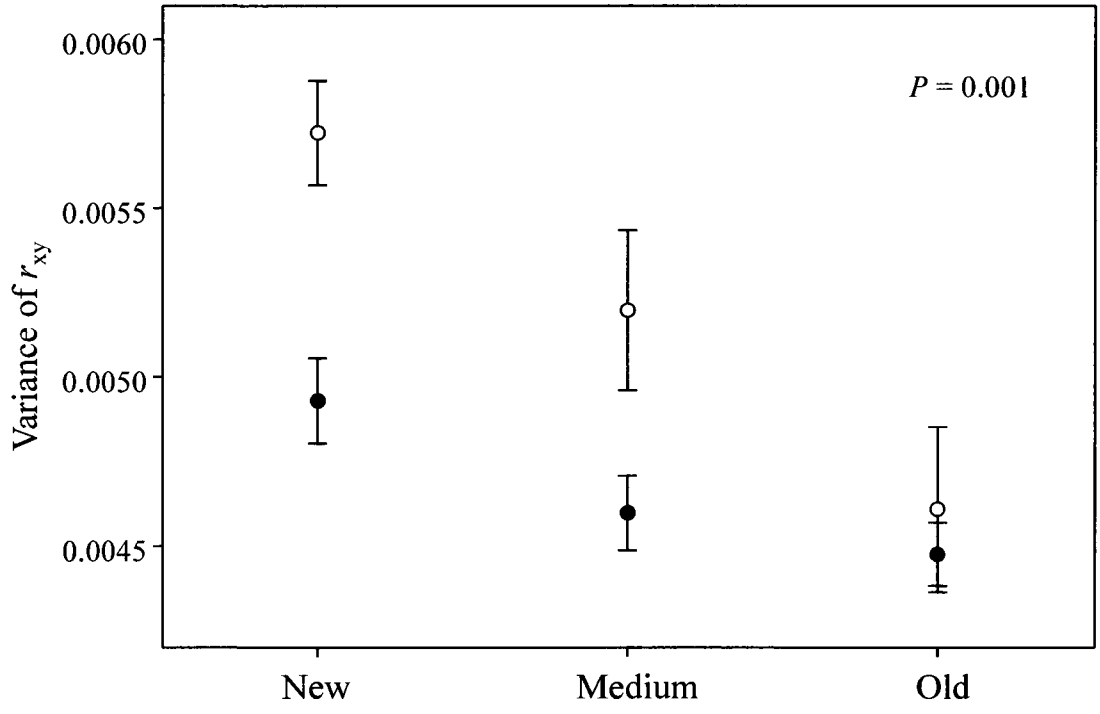


Figure 1.15. Variances of Lynch and Ritland (1999) relatedness estimates (r_{xy}) for odd-year pink salmon collections from three stream-age categories. The mean observed (open circles) and permuted (solid circles) variances differed across stream-ages (P = 0.024 and 0.050, respectively). Observed variance was greater than the permuted variance (P = 0.001, ANOVA).

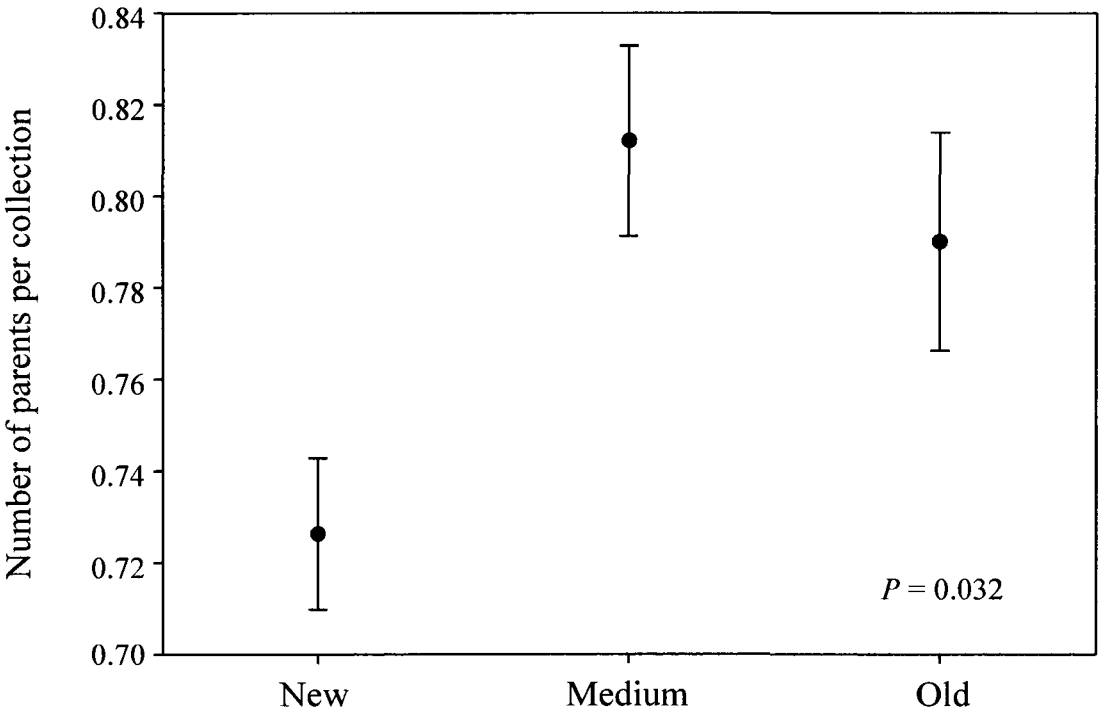


Figure 1.16. Estimated number of parents per collection for the odd-year pink salmon from three stream-age categories (mean, SE). Standardized by collection sample size; from the genealogy that generated the highest correlation between the pedigree and molecular co-ancestries (Fernandez and Toro, 2006). *P*-value from Kruskal-Wallis one-way ANOVA on ranks.

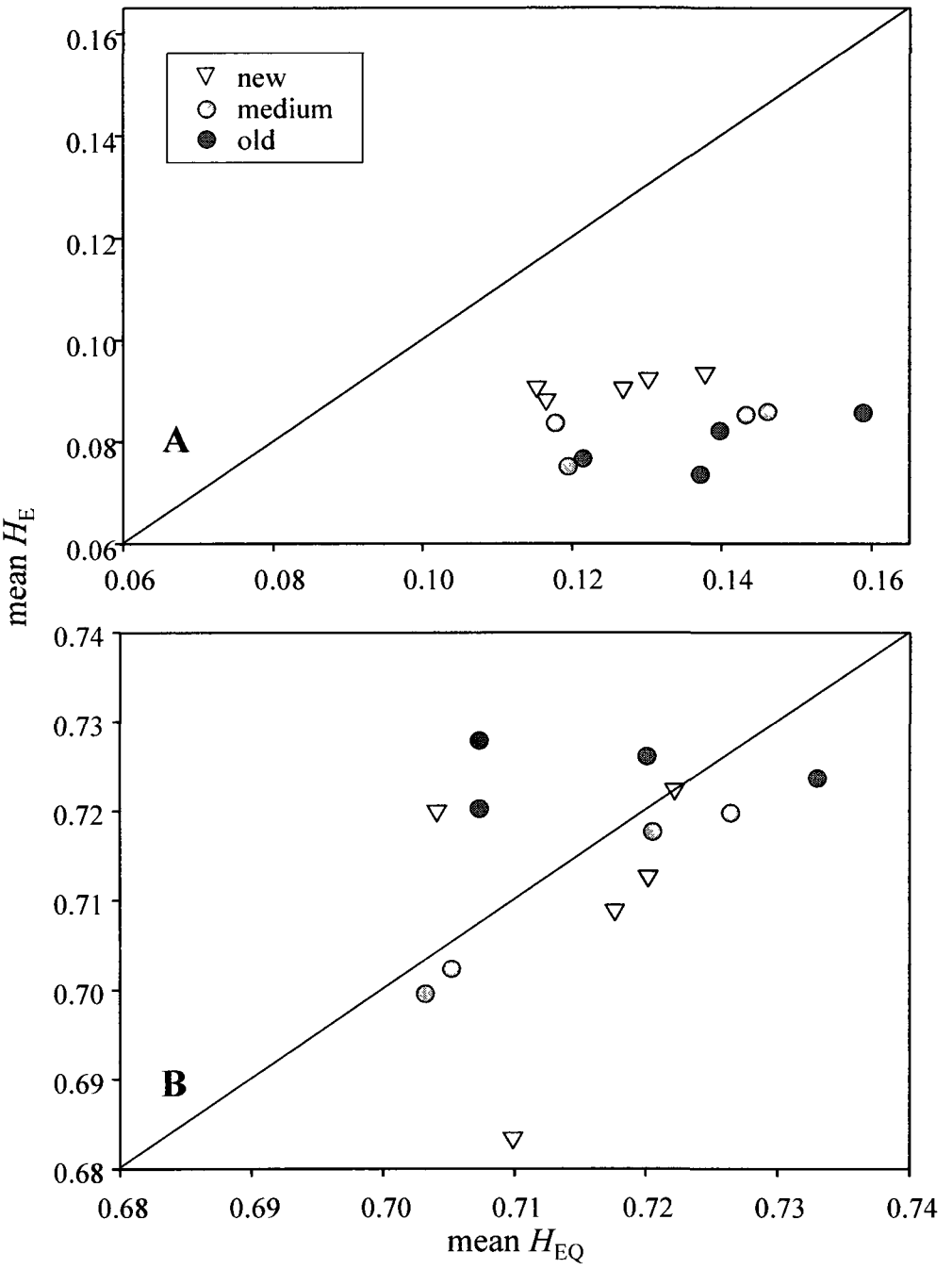


Figure 1.17. Equilibrium heterozygosity (H_{EQ}) and expected heterozygosity (H_E) for odd-year pink salmon from three stream-age categories from the BOTTLENECK program. **(A)** 36 variable allozyme loci with IAM, and **(B)** 9 microsatellite loci with 95% one-step mutations with TPM. The line represents values where H_{EQ} equals H_E.

Chapter 2

The colonization mechanism of even-year pink salmon populations in Glacier Bay, Alaska, based on genetic data[1]

[1] Prepared with the format requirements of the journal Transactions of the American Fisheries Society.

Abstract

Pink salmon *Oncorhynchus gorbuscha* from the even-year broodline colonized many of the watersheds of Glacier Bay, Alaska, following the end of the Neoglacial period in the early 18th century. Streams in the lower Bay were populated first, and colonization proceeded up the Bay during the last 200 years. The population genetic structure based on allozyme and microsatellite allele frequencies and mtDNA haplotype frequencies indicated that colonization was episodic and that streams were probably populated primarily by colonists from nearby populations. Within Glacier Bay, this mechanism of colonization created small-scale homogeneity among populations that shared glacial histories and larger scale heterogeneity among populations that were colonized at different times. For the medium-aged populations in lower Glacier Bay, colonists came from sources outside the Bay. The most recently formed populations appear to have been colonized from older donor sources outside the Bay, as well as from populations within the Bay. Slightly higher linkage disequilibrium, lower relatedness, and higher allelic richness of the youngest population indicated that colonists came from several source populations including its nearest neighbor and older, more variable populations from farther away. Colonization of the most recently deglaciated streams was associated with large increases in pink salmon abundance in Southeast Alaska in the early 1990s. The number of fish involved in initial colonization was not large, but minimal founder effects precluded very small numbers of fish as well. Once colonization started, it resulted in rapid, large increases in population size. Strong heterogeneity among the medium-aged populations as well as a decrease in allele and haplotype richness suggests that recurrent gene flow is minimal. Genetic divergence among some northern Southeast Alaska populations indicates that heterogeneity may persist within the confines of Glacier Bay for many generations, especially for populations that grow large rapidly after colonization due to high survival of offspring rather than gene flow.

Introduction

Pink salmon *Oncorhynchus gorbuscha*, with their two-year life history, have broodlines that spawn in the even and odd calendar years. The two broodlines can be interbred, but they have been reproductively isolated for so long that crosses exhibit outbreeding depression (Gharrett et al. 1999). Thus, the two broodlines serve as replicate experiments for analysis of the colonization process. Broodline dominance varies across the geographic range, and some areas have predominately or exclusively a single broodline (Heard 1991); but in Southeast Alaska, both are abundant and occupy the same streams. The even-year broodline has had consistently lower escapement, about one-quarter that of the odd-year broodline in the Icy Strait area, including Glacier Bay, since at least the 1960s (Heinl and Geiger 2005). The same four questions that were asked about colonization by the odd-year broodline (Kondzela 2010) will be addressed for the even-year broodline in this chapter.

Were source populations from locations nearby or farther away?

There are two time frames to consider when we ask about the sources of colonists: (1) what were the sources of the original colonists to Glacier Bay, that is, the fish that seeded the older Glacier Bay populations in the lower part of the bay about 150 years ago?; and (2) what were the sources of recent colonization events within the last 60 years? If the source populations were from nearby locations, the colonists would genetically resemble those populations more than distant populations, i.e., the allele or haplotype frequency distribution of colonists would be more similar to those of nearby populations (e.g., Hawkins et al. 2002), and F_{ST} values would be low. If the donors were from populations farther away, then colonists would have allele frequencies that differ from fish of older neighboring populations; and the new populations would be heterogeneous in comparison to nearby populations, which would result in higher F_{ST} values. Descriptive analyses, such as population trees and principal component analysis, followed by explicit homogeneity tests that examine spatial and temporal relationships, will be used to

determine whether colonists in Glacier Bay are similar to nearby or more distant populations.

Did single or multiple source populations contribute colonists?

If colonists came from multiple, genetically different population sources, linkage disequilibrium should exist at the time of colonization (Gharrett and Zhivotovsky 2003). Linkage disequilibrium does not disappear in a single generation, unlike the Wahlund effect, and will persist for several generations (Hedrick 2005). A composite measure (correlation coefficients) of linkage disequilibrium will be examined. There would also likely be higher allele richness and heterozygosity, as well as lower average relatedness and larger variance of average relatedness (Ritland 2000) among fish in new populations for several generations, than if new populations were derived from a single source population.

Did new populations start with small or large numbers of fish?

If a small number of fish colonized a location, we would expect founder effects with reduced genetic variation (e.g., heterozygosity, allele richness, and haplotype diversity would all be lower), small effective population size, and higher linkage disequilibrium (Hedrick 2005); however, if a large number of fish were involved, new populations would be genetically similar to older populations, have large effective population sizes, and exhibit lower linkage disequilibrium. Also, due to a stochastic sampling effect, allele frequencies in newer populations established by few colonists could result in significant heterogeneity among new populations (e.g., Gharrett and Thomason 1987). A very small number of colonists would result in a deficit of alleles and an excess of heterozygosity in a mutation-drift steady state situation; this bottleneck signal erodes with a half life of about $4N_e$ generations (Kimura 1983; Cornuet and Luikart 1996).

Did colonization occur as a one-time event or as a recurrent process?

If immigration into new populations was recurrent, then (depending on other factors, such as number and source of colonists) there should be an increase in both genetic diversity (e.g., allele richness) within populations and homogeneity among populations as they advance in age. If colonization was a one-time event, heterogeneity among populations would develop due to drift; and the relatedness of individuals within populations would increase over time, particularly if small numbers of colonists were involved.

Approach

The objective of our study was to examine alternative colonization scenarios that we could expect to distinguish from the genetic signals. We compared mtDNA variation with nuclear variation. Mitochondrial markers are haploid in salmon, have about one-quarter the effective population size of nuclear markers, and should be more sensitive to numbers of colonists and source populations (Nei and Tajima 1981). In addition, because mtDNA is matrilineal, sex-related differences in dispersal may produce differences between mitochondrial and diploid nuclear markers. Clearly, the colonization mechanism would have been somewhere within the contrasting boundaries described above. By addressing those questions, we should be able to narrow the possibilities and possibly generate some other questions to pursue.

Methods

Study Site and Sample Collection

Even-year pink salmon populations were sampled between 1990 and 1994 from many of the same locations that were sampled for the odd-year broodline (Table 2.1; Figure 2.1). Samples were collected and processed as described in Kondzela (2010).

Streams that developed in the wake of ice retreat in Glacier Bay are progressively younger from the mouth to the upper reaches of the Bay. Four locations in the middle-to-lower section of Glacier Bay were sampled: the north head of Berg Bay, the north head of

the north arm of Fingers Bay, the head of Tyndall Cove, and the stream just north of Vivid Lake, to which it connects during high water events. The mouths of these streams were exposed in approximately 1830, 1845, 1875, and 1870 respectively, and these locations, which have been ice free less than 200 years, are referred to as "medium-aged" throughout this study. Pink salmon were scarce or absent in the youngest ice-free locations of Glacier Bay during the first two even years of our study, 1990 and 1992, even though pink salmon were abundant in some of the medium-aged populations (e.g., Tyndall). Sampling the two youngest populations, Wolf Pt. and Nunatak, occurred only in 1994. In that year, several thousand fish were present in Nunatak, whereas only about 150 fish were observed in the entire length of Wolf Pt. Creek from the intertidal zone to the barrier falls approximately 2 km upstream (C. Soiseth, National Park Service [NPS], personal communication). Only 41 fish were sampled from Wolf Pt., and from 15 of them, only skeletal muscle and adipose fin were taken. Populations from four streams adjacent to Glacier Bay—E. Kahtaheena, Dundas, Homeshore, and Spasski—were sampled for comparison with the populations within Glacier Bay. For the two locations that had not been sampled in odd years, fish were collected from the intertidal zone to about 1 km upstream in the Vivid system, and from the intertidal/lower stream reach at the head of the north arm of Dundas Bay.

Data Acquisition

The laboratory protocols for allozyme, microsatellite, and mtDNA markers were the same as those followed for the odd-year pink salmon samples (Kondzela 2010). The same suite of microsatellite loci was analyzed and the same PCR amplification parameters were used (Table 2.2). Even-year pink salmon share some of the same mtDNA haplotypes observed in the odd-year broodline, but enough differences exist that a different suite of mtDNA regions and restriction endonucleases were required to examine the basic haplotype structure of the even-year broodline (Table 2.3). Composite haplotypes of even-year pink salmon were identified from restriction endonuclease digestion of the mtDNA region ND5/ND6 with *Bst*N I, *Bst*U I, and *Hinf* I, and region

Cyt*b*/D-loop with *Msp* I, *Rsa* I, and *Sau*96 I. To further clarify haplotype relationships, including new haplotypes, a subset of samples was amplified for the ND3/ND4 region and digested with *Rsa* I, and for the ND1/ND2 region and digested with *Bst*N I. Samples identified with insertions in the Cyt*b*/D-loop region were examined with double digests (*Ban* II, *Hinf* I, and *Mbo* I) to confirm insertion sizes and narrow down their location in the genome (see Appendix 2.1 for details).

Data Analyses

The data obtained from the allozyme, microsatellite, and mtDNA markers were examined with the same suite of analyses used for the odd-year broodline (Kondzela 2010) to provide a parallel study of the colonization mechanism of the even-year broodline in Glacier Bay, Alaska. Microsatellite genotypes were evaluated for large allele dropout, null alleles, and stuttering to detect potential genotyping errors with the software program MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004). Multiple testing was taken into account by using a sequential Bonferroni correction (Rice 1989) in which the initial Type I error (α) was adjusted by the number of tests within each collection.

Both microsatellite and allozyme data were tested for conformance to Hardy-Weinberg expectations (GENEPOP 3.4; Raymond and Rousset 1995). Isolocus data (Allendorf and Thorgaard 1984) cannot be tested for Hardy-Weinberg frequencies.

Departure from gametic equilibrium was examined with all nine microsatellite loci and allozyme loci that had common allele frequencies less than or equal to 0.95 in at least one collection, or had at least two alleles in every collection. Composite genotypic equilibrium was tested by permutation testing with the program GENETIX 4.05 (http://www.genetix.univ-montp2.fr/genetix/genetix.htm). For each locus-pair in each collection, 10,000 permutations of genotypes were used to test the null hypothesis of no correlation between alleles at two loci.

A reduced median network of pink salmon mtDNA haplotypes was constructed from data observed in our study and in Churikov and Gharrett (2002) by using the program NETWORK 4.1.1.2 (www.fluxus-engineering.com; Bandelt et al. 1995).

Population Structure

The same tree and principal component clustering analyses and rules used for analysis of the odd-year datasets (Kondzela 2010) were also used on the even-year datasets. For the principal component analyses, only the common allele was retained for the bi-allelic microsatellite loci, *Ots208-2* and *Oneμ13*. Testing of homogeneity of allele and haplotype frequencies and estimation of F_{ST} values were completed as described in Kondzela (2010), except that analyses were done with comparisons of three (young, medium, old) and four (the medium-age category was further split in two) stream-age categories.

Measures of Differences among Populations

The effect of stream age (new, medium, old) on genetic diversity measures of the even-year broodline was examined with the same suite of analyses applied to the odd-year dataset (Kondzela 2010). Allele and haplotype richness, heterozygosity, number of private alleles/haplotypes, and residual genotypic linkage disequilibrium were determined in each collection (allele richness was estimated for each locus separately in FSTAT to maximize sample sizes). Relatedness measures were examined in each collection from individual genotypes for which both variable allozyme and microsatellite data were available (Lynch and Ritland 1999; Fernandez and Toro 2006; Kalinowski et al. 2006). Effective population size (N_e) was estimated with linkage disequilibrium and temporal methods (Waples 1989, 1991; Waples et al. 2006; Waples and Do 2008). Each population was tested for recent bottlenecks (BOTTLENECK 1.2.02; Cornuet and Luikart 1996; Piry et al. 1999) with separate analyses for the allozyme and microsatellite data. Thirty-three allozyme loci that were variable in the even-year broodline were assumed to mutate under the infinite allele model (IAM), and all of the microsatellite loci

were assumed to mutate under the two-phase model (TPM; Di Rienzo et al. 1994). Two mutation scenarios were considered to examine the robustness of the TPM: (1) with 95% one-step and 5% multi-step mutations, and the variance among multiple steps of 12 as recommended by Piry et al. (1999); and (2) with 70% one-step and 30% multi-step mutations, and a variance among multiple steps of 30—parameter settings that bring the model closer to the IAM.

Results

Allozyme Loci

Thirty-three allozyme loci were variable in the even-year pink salmon collections with 2–6 alleles per locus (Appendices 2.2 and 2.3). Sixteen loci were monomorphic in all collections. Four isoloci (*sAAT-1,2; GPI-B1,2; sMDH-A1,2; sMDH-B1,2*) exhibited low variability; for subsequent analyses, all variation was assigned to one locus while the other locus was designated monomorphic (Gharrett and Thomason 1987; Waples 1988). Of the variable loci, 19 loci had a common-allele frequency greater than or equal to 0.95, and 14 loci had a common-allele frequency less than 0.95 in at least one collection. The most common allele at each locus was observed in every collection; some additional alleles occurred in single collections, other alleles in multiple collections. A total of 93 alleles were observed at the 33 variable loci; individual collections had 51 to 70 alleles. Data for *sAAT-4* was not obtained for Tyndall 1990 because of poor tissue quality.

Eleven alleles were restricted to populations within Glacier Bay. Of these alleles, five were infrequent and observed only in the youngest collections (*mAAT-1*-80; LDH-B1*135;* and *sSOD-1*175* only in Nunatak; and *G3PDH-1*70* and *PEPA*109* only in Wolf Pt). Five alleles were unique to the medium-aged collections, all at very low frequency as one or two copies in one or two collections (*sAAT-1,2*65; LDH-A1*11* and **90; sMDH-A1,2*105; sMDH-B1,2*56/59*). The *sAH*84* allele was observed in single heterozygotes in the Wolf Pt. and N. Berg collections. Nine alleles unique to the four populations outside Glacier Bay were rare, and most were singletons (*mAAT-1*-112;*

*mAH-1*75, mAH-3*75; sAH*76; GPI-B1,2*63* and **54; sMDH-A1,2*90; MPI*118; TPI-3*95*).

Microsatellite Loci

Most of the microsatellite loci were highly allelic, and a large proportion of alleles were at low frequencies (Appendices 2.2 and 2.4). All nine loci were polymorphic in every collection. Twenty-two alleles were present only in the Glacier Bay populations; twelve as singletons and ten at low frequency (except *One103*297* at a frequency of 0.10 in N. Fingers) in one to four populations. Twenty-four alleles were unique to populations outside Glacier Bay, sixteen as singletons across the four populations, and the remaining at low frequency spread across one or two of the four populations. Interpretation of the microsatellite data did not appear to be affected by large allele dropout or stuttering, and evidence of null alleles existed only at *One102* in two collections, Vivid 1994 (F_{IS} = 0.137, $P < 0.01$) and Tyndall 1992 ($F_{IS} = 0.162$, $P < 0.05$).

mtDNA RFLPs

Restriction endonuclease sites were inferred from single changes in mtDNA fragment patterns (Appendices 2.2 and 2.5). Three of the five major lineages that were previously reported (Churikov and Gharrett 2002) were observed in the even-year pink salmon samples (Figure 2.2; Tables 2.4 and 2.5). Half of the 20 composite mtDNA haplotypes observed in the even-year pink salmon collections were new—all occurred as singletons or at low frequencies. Seven haplotypes were observed only within Glacier Bay: as singletons (p, q, x, y, z) across four collections and in one fish in each of two populations (M*, X). Six haplotypes were observed only outside Glacier Bay: as singletons (BD, r, s, u, v) in either Dundas or Homeshore, and in two fish from Spasski (w). Five fish from four collections had insertions in the Cyt*b*/D-loop regions (Table 2.6; Appendix 2.1). One fish from N. Fingers 1992 had a single 80 bp insert, one fish from N. Berg had two tandem 80 bp inserts, one fish from Tyndall 1990 had a single 260 bp insert, and two fish

from Spasski 1992 had an insertion of about 520 bp, which may represent tandem copies of the 260 bp insert.

Hardy-Weinberg Equilibrium

Of the 152 possible Hardy-Weinberg equilibrium probability tests for allozyme loci, only seven were significant (4.6%), spread across five loci and five collections, none of which remained statistically significant after correction for multiple testing. There were more departures from Hardy-Weinberg equilibrium for microsatellite loci than would be expected by chance (5%); nine of 99 possible tests (9%) were spread across five loci and seven collections. However, no tests were statistically significant after correction for multiple testing.

Gametic Disequilibrium

For microsatellite and allozyme loci combined, there were 181 significant locus-pair departures from gametic equilibrium out of 3,107 possible comparisons (5.8%), about the proportion expected by chance alone (5%). The departures from equilibrium were spread across all collections and locus-pairs—3.9% of the allozyme-only, 8.1% of the microsatellite-only, and 6.8% of the allozyme-microsatellite locus-pairs. The fraction of locus-pairs out of equilibrium did not differ for collections from different stream ages ($P = 0.99$, Kruskal-Wallis one-way ANOVA on ranks). Only one test remained significant after sequential Bonferroni corrections for multiple testing (*One111* and *Oneμ13*; Vivid 1994).

Population Structure

The neighbor-joining (NJ) tree structures from the allozyme and microsatellite markers were similar (Appendix 2.6); the primary difference was the position of the N. Fingers branch. The topology of the combined microsatellite and allozyme NJ tree was nearly identical to the consensus tree based on 1,000 bootstrap resamplings of the alleles at each collection (Figure 2.3). The strength of the branch order was high for the

medium-aged Glacier Bay collections, Vivid and Tyndall, which clustered together 100% of the time. Another strong grouping included the two youngest populations, Wolf Pt. and Nunatak, with the next youngest populations Vivid and Tyndall (clustered 78% of the time). The 10 continuous maximum-likelihood trees were similar to the NJ tree (Appendix 2.7), although there was some branch swapping among trees, particularly for the medium-aged populations, N. Berg and N. Fingers, in lower Glacier Bay. The position of populations on the mitochondrial NJ tree was less closely associated with geographic locale or time of colonization than the nuclear trees (Figure 2.4), possibly because of smaller samples sizes and fewer haplotypes than nuclear alleles. The Tyndall and Vivid populations clustered together, but the proportion of times that grouping occurred in the consensus tree was not high (32% of the time in the consensus tree). The strongest pairing was in lower Glacier Bay, between the medium-aged populations in N. Berg and N. Fingers (84%). The old populations outside the Bay were dispersed among the Glacier Bay populations, and the youngest populations were most closely aligned with the oldest populations outside Glacier Bay, especially Wolf Pt. and Spasski (77%).

Principal component analysis corroborated the tree analyses. Much of the structure in the trees was mirrored in the separation of populations on the first principal component (Figure 2.5; Appendix 2.8). The youngest populations, Nunatak and Wolf Pt., clustered together in the first four components ($P = 0.33$, pseudo-exact test) and were separated from all other populations on the second component ($P < 10^{-4}$, pseudo-exact test). The four oldest populations (outside Glacier Bay) clustered tightly in the first four components ($P = 0.60$, pseudo-exact test). The medium-aged populations, Tyndall and Vivid, consistently separated from other populations in the first four components ($P < 10^{-4}$, pseudo-exact test). The other two medium-aged populations, N. Fingers and N. Berg, aligned more closely with the oldest populations outside Glacier Bay. Alleles from both allozyme and microsatellite loci contributed to the greatest component loadings of the principal components, but most of the contribution came from the microsatellite loci (Appendix 2.9).

Homogeneity

Tyndall was the only population that was sampled more than once; the two collections from 1990 and 1992 were homogeneous for frequencies of all three marker types. Within stream-age categories, the two youngest populations and four oldest populations were homogeneous for all three marker types, and F_{ST} values were near zero (Table 2.7). In contrast, the medium-aged populations were significantly heterogeneous and had much larger F_{ST} values. The heterogeneity of populations in the medium-aged category was consistent with the structure in the NJ trees and PCA, and could be explained primarily by allele and haplotype frequency differences (e.g., alleles *CK-C1*94*, *PEPLT*83* and **110*, and haplotypes A*, AA*, and H*) between the Tyndall/Vivid group in the mid-upper Bay and the N. Fingers/N. Berg group in the lower Bay (Appendix 2.3; Table 2.5). The presence or absence of alleles and haplotypes also contributed to the heterogeneity of the medium-aged populations, e.g., *sAH*84* and **86,* and *sMDH-B2*95* and **105* were present only in the two oldest populations in lower Glacier Bay; and five mtDNA haplotypes were observed only in either Tyndall/Vivid or N. Fingers/N. Berg. The strong heterogeneity within the medium-aged category, led us to further split this group into two categories, "medium-young" and "medium-old". Significant heterogeneity was observed within these two age categories for the nuclear loci, but not the mtDNA haplotypes (Table 2.8).

Across the three stream-age categories, F_{ST} values were slightly larger for the nuclear loci (0.0065 and 0.0055 for allozyme and microsatellite loci, respectively) as compared to the mtDNA (0.0037, Table 2.7). Although the F_{ST} values were relatively small, heterogeneity occurred due to differences in both the mtDNA haplotype frequencies and the nuclear allele frequencies ($P = 0.0322$ and < 0.0001, respectively).

The population pairwise mitochondrial F_{ST} values were about four times higher than the nuclear F_{ST} values (Figure 2.6; Appendix 2.10). Across stream-age categories, the mitochondrial and nuclear pairwise F_{ST} values had identical patterns (Figure 2.6); the highest F_{ST} values were between populations in the medium-aged category ($P = 0.094$ and 0.002, respectively; left sides Figure 2.6). The mean pairwise F_{ST} values of

populations from different stream-age categories (right sides Figure 2.6) was higher for the "medium-old" and "new-medium" comparisons than the "new-old" comparisons ($P < 0.001$ and $P = 0.013$, respectively). These differences could be attributed to the relatively high F_{ST} values between the medium-aged population in Tyndall and both the youngest populations, as well as the older populations outside Glacier Bay (Appendix 2.10). Several populations contributed to the higher mtDNA F_{ST} values in the medium-old pairwise comparisons.

To place the Glacier Bay study in context with populations from a larger geographic scale, nine additional populations from northern Southeast Alaska (Figure 2.7; Appendix 2.11) were compared with 30 allozyme loci for which there was compatible data (S. Wildes, NOAA, unpublished data). The NJ tree and PCA (Figure 2.8) had the same general pattern of population structure evident in the previous NJ trees and PCA. The populations within Glacier Bay exhibited more divergence than populations outside Glacier Bay.

Unlike the four oldest populations outside Glacier Bay, the nine additional northern Southeast Alaska populations were heterogeneous ($P < 0.0001$, pseudo-exact tests). The genetic similarity of the old populations outside Glacier Bay to the lower Glacier Bay populations suggests that they are candidate source populations for colonizing Glacier Bay, and more likely to be donor sources than some of the other northern Southeast Alaska populations.

The allozyme population pairwise F_{ST} values for the northern Southeast Alaska populations were higher than those observed among the oldest populations used for the Glacier Bay study, but they were still low, with a mean value of 0.005 (left side Figure 2.9). The Chilkoot population was associated with the highest F_{ST} values among the pairwise comparisons with the other northern Southeast Alaska populations. This population also contributed to the highest F_{ST} values in pairwise comparisons with the Glacier Bay study populations ("med-NSE" and "new-NSE" on the right side of Figure 2.9). Likewise, the Tyndall population, which drove the higher F_{ST} values in pairwise

comparisons within the Glacier Bay study, was associated with the highest F_{ST} values in pairwise population comparisons of the "med-NSE" category (right side of Figure 2.9).

Testing of Population Structure

The correlation coefficients of linkage disequilibrium for the 136 locus-pairs for which there was variation in all collections were higher in the two youngest collections in upper Glacier Bay than in collections from older populations within and outside Glacier Bay ($P = 0.003$; Figure 2.10).

Genetic Diversity.—The allele richness of allozyme loci increased across the three stream-age categories, but the relationship was not significant ($P = 0.123$, *sAAT-4* not included; Figure 2.11A). With the exception of N. Fingers, which had the highest number of allozyme alleles in any population, the populations outside Glacier Bay had higher allele richness than the populations within Glacier Bay. On average, two more allozyme alleles were observed across forty loci in the older populations outside Glacier Bay, but the difference was not significant ($P = 0.073$, Mann-Whitney rank sum test, *sAAT-4* not included). More microsatellite alleles per locus were observed in the older populations ($P = 0.032$; Figure 2.11B). Fifteen more microsatellite alleles were observed across nine loci outside Glacier Bay than within the Bay ($P = 0.024$, Mann-Whitney rank sum test). The allele richness of microsatellite loci in the youngest population, Wolf Pt., was higher than in all the other Glacier Bay populations except N. Berg, the oldest population within Glacier Bay, and similar to the older populations outside the Bay. Among Wolf Pt., N. Berg, and the older populations outside the Bay there were twenty more microsatellite alleles across the nine loci than observed in the other populations within the Bay ($P = 0.002$; Mann-Whitney rank sum test). Heterozygosity was higher in the medium-aged populations than the young or old populations for the allozyme loci ($P = 0.036$, *sAAT-4* not included), but lower for the microsatellite loci, although not statistically significant (Figure 2.12). Haplotype richness and diversity were lowest in the medium-aged populations ($P = 0.032$ and 0.007, respectively; Figure 2.13). The number of private alleles and haplotypes was higher in the old populations outside Glacier Bay,

but across the three stream-age categories, the differences were not significant (Figure 2.14). In a comparison of the five youngest and five oldest populations, the higher number of private alleles and haplotypes in the oldest populations was nearly significant ($P = 0.052$, rank sum test).

Relatedness.—The mean Lynch and Ritland (1999) relatedness (r_{xy}) values were less than zero for all collections. Although the mean values of relatedness indicated no relatedness among the fish within collections, the variance of relatedness measures provides different insights into the structure of populations. Observed variance of r_{xy} differed across the three stream-age categories ($P = 0.009$; Figure 2.15; Table 2.9), and was highest in the medium-aged populations and lowest in the oldest populations. There were no differences across the three stream-age categories in permuted variance of r_{xy} ($P = 0.193$). The observed variance of r_{xy} exceeded the permuted variance across the three stream-age categories ($P = 0.020$), possibly indicative of more independent groups of related fish than expected under random mating, particularly in the medium-aged populations. As with some of the other genetic measures, the values of variance of relatedness for the youngest and oldest populations within Glacier Bay (Wolf Pt. and N. Berg, respectively) were similar to those of the populations outside the Bay. Those two populations, combined with the populations outside the Bay, had lower values of relatedness variance than the other Glacier Bay populations ($P = 0.004$ and 0.017, observed and permuted values, respectively).

The mean maximum likelihood estimates of relatedness (Kalinowski et al. 2006) were low (0.030–0.044; Table 2.9) and did not differ across the three stream-age categories (Appendix 2.11). However, the set of five oldest populations (N. Berg and outside populations) and the youngest population, Wolf Pt., had lower relatedness ($P = 0.004$; Appendix 2.11) and a smaller proportion of related individuals ($P < 0.001$, chi-square) than the other populations within Glacier Bay. Within collections, most of the fish were unrelated to each other based on their genotypes; however, the maximum-likelihood method determined that many, typically two-thirds or more, of the individuals within a collection had genotypes that were consistent with a half- or full-sib relationship to at

least one other individual in the collection. If the potential sibling relationships are real, microsatellite genotypes from additional loci would clarify this result. The number of unique parents (standardized by collection size) estimated from the genealogy for individuals within each collection (Fernandez and Toro 2006) was higher in the oldest populations, which corroborates the lower relatedness estimates with the maximum-likelihood method, but the difference across stream-age categories was not significant (P = 0.381; Table 2.9, Figure 2.16).

Effective Population Size.—All of the effective population size estimates (N_e) were fairly large (more than 100 fish; Table 2.10). A negative N_e estimate indicates that the disequilibrium can be explained by sampling error, and the population cannot be distinguished from an infinitely large population (Waples 1991). The 95% confidence intervals were broad and included a minimum of 31–111 fish and a maximum of infinity. The oldest populations tended to have larger N_e, as did one of the medium-aged populations, Tyndall, and the youngest population, Wolf Pt. The small sample size from Wolf Pt. may explain the negative N_e estimate for this population. Only one population was sampled in multiple years, Tyndall in 1990 and 1992, and the N_e estimate was much smaller with the temporal method (215 fish) than the linkage disequilibrium method (818 fish in 1990 and -355 fish in 1992), although the confidence intervals of the estimates for the two methods overlapped. Only one generation of drift was measured, which is the minimum time frame that can be used for the temporal method. For the linkage disequilibrium method, the estimation of N_e is influenced by the disequilibrium created with population mixing, which will bias the estimate of N_e to values lower than expected under a migration-drift steady state.

Bottlenecks.—There was no evidence of a bottleneck in any collection from the comparison of heterozygosity with that expected under mutation-drift equilibrium based on the allozyme loci (no significant Wilcoxon test and no mode-shift of allele frequencies; Figure 2.17A). Instead, the allozyme markers exhibited a slight heterozygosity deficiency—10 of the 11 collections had lower heterozygosity than expected under equilibrium conditions.

There was also no evidence of a bottleneck with the microsatellite loci, although the heterozygosity expected under mutation-drift equilibrium depends on the mutation model tested. There was no sign of a bottleneck with the TPM with 95% one-step and 5% multiple-step mutations (Figure 2.17B); about half the populations had slightly higher and the other half slightly lower than expected heterozygosity under mutation-drift equilibrium. However, when the proportion of one-step mutations was lowered to 70%, bringing the model closer to the IAM, all but one population had an excess of heterozygosity as compared to that expected under mutation-drift equilibrium (not shown). Thus, the TPM model with more one-step mutations is conservative when testing for the presence of a bottleneck, i.e., an excess of heterozygosity under these model conditions increases the likelihood that it represents a real bottleneck. Nevertheless, no Wilcoxon tests were significant and there was no mode-shift of allele frequencies under either TPM mutation scenario.

Discussion

Many Glacier Bay streams are now populated by pink salmon following retreat of ice that began in the early 1700s. The mechanism(s) by which pink salmon colonize newly available stream habitat are poorly understood; however, genetic information can provide some clues about the process. The genetic signals expected under different colonization mechanisms (Table 1.1 in Kondzela 2010) are useful in interpretation of the analyses used to answer the questions posed in this study for the even-year broodline and addressed below.

Were source populations from locations nearby or farther away?

The genetic structure of the populations within and outside Glacier Bay does not offer a simple answer to this question. N. Berg, the oldest population within Glacier Bay and nearest to the entrance of the Bay, was genetically similar to the nearby populations outside the Bay based on the nuclear markers (Figures 2.3, 2.5; Appendices 2.6–2.8, except on PC3); this relationship was less clear with the mtDNA haplotypes (Figure 2.4).

The N. Fingers population, the next oldest population up the Bay, was distinct in some analyses (e.g., Figure 2.5) and clustered with populations in different ways in other analyses (e.g., Figures 2.3, 2.4; Appendix 2.6), but in general this population was genetically intermediate between the older populations and the younger populations further up the Bay. The number of fish present in N. Fingers during sample collection was small compared to other medium-aged populations and this population had the smallest N_e estimate (Table 2.10), which may have contributed to the apparent divergence of this population. Whereas the streams in the lower Bay appear to have been colonized by donors from nearby sources, the source of colonists for other streams in Glacier Bay is less certain. The remaining medium-aged populations, Tyndall and Vivid, clustered together in every tree and in the principal component analysis (PCA) and were genetically distant from all other populations as evident from the high pairwise F_{ST} values (Appendix 2.10). The time of pink salmon colonization and the origin of colonists of these two systems is unknown; the retreat of ice from the mouth was estimated to be within a few years of each other (ca. 1870 and 1875), which increases the likelihood that they experienced colonization at approximately the same time, but exactly when and from where is unknown. Most of the divergence of Tyndall and Vivid from other populations is explained by allele and haplotype frequency differences, rather than the presence or absence of specific alleles; this reduces the probability of colonization from more divergent sources farther away.

Wolf Pt. and Nunatak creeks also shared a similar deglaciation history. Small numbers of even-year pink salmon (125 fish) were first observed in Nunatak in 1992 (G. Streveler and C. Soiseth, NPS, unpublished data), followed by approximately 2,500 fish in 1994. No fish were observed in Wolf Pt. in 1992 and only about 150 fish in 1994. Given the genetic similarity of Wolf Pt. and Nunatak, and the large number of fish in Nunatak in 1994, it is plausible that they share a similar colonization history or that Nunatak was the primary source of colonists to its nearest neighbor, Wolf Pt. Based on population pairwise F_{ST} values, it appears that colonization of the two youngest populations was more likely to be from populations in the lower Bay or even outside the

Bay than from Tyndall or Vivid (Figure 2.6; Appendix 2.10). However, in both the PCA and NJ trees, the two youngest populations tended to lie intermediate to the oldest populations and the medium-aged populations, with the strongest affinity to Tyndall and Vivid (Figure 2.3B). Other possible explanations include (1) possible donor source populations within Glacier Bay that were not sampled, e.g., Beartrack and Bartlett rivers on the lower east side of Glacier Bay, and (2) a mix of colonists from both the medium-aged and older populations.

Although there was heterogeneity among the three stream-age categories, allele- and haplotype-frequency differences between populations were not large. The presence or absence of several alleles and haplotypes contributed to the heterogeneity observed among the even-year pink salmon populations. Some mtDNA haplotypes and nuclear alleles were observed only within Glacier Bay and others only outside the Bay, generally at low frequencies (Table 2.5, Appendices 2.2–2.4). The origin of alleles and haplotypes restricted to the Glacier populations is unknown, but given their low frequency they may exist in the older populations outside Glacier Bay at low frequencies as well, or in populations farther away. The older populations used in this study were genetically similar to some of the other northern Southeast Alaska populations (Figure 2.8); however, the older populations were also more similar to the lower Glacier Bay populations than some northern Southeast Alaska populations (Figure 2.9) and thus more likely candidate donor sources of colonists than such northern Southeast Alaska populations as Chilkoot, Taku, or Auke Creek.

Did single or multiple source populations contribute colonists?

It appears that few source populations contributed colonists to the youngest populations; some genetic signals were clear on this and others were ambiguous. Some signals were consistent with either a single donor source population (or genetically similar fish from multiple locations) or a few genetically divergent donor populations. The lower microsatellite allele richness of populations within Glacier Bay compared with that of populations outside the Bay is consistent with a single donor source (Figure

2.11B). However, at the population level, two exceptions stand out: the youngest population, Wolf Pt., and the oldest population within Glacier Bay, N. Berg, had levels of microsatellite allele richness and mtDNA haplotype richness as high as that in populations outside Glacier Bay. The relatively high number of microsatellite alleles and mtDNA haplotypes in Wolf Pt. indicates multiple contributing source populations in the early formation of this population, or contribution from older populations which have more alleles and haplotypes. There was also higher allozyme allele richness in the oldest populations (Figure 2.11A), but, in contrast to the microsatellite loci, Wolf Pt. had one of the lowest measures of allozyme allele richness, which supports the concept of contribution from few sources during early population formation. Another sign that relatively few populations contributed colonists to the formation of the youngest populations in Glacier Bay was the higher number of private alleles and haplotypes in the older populations (Figure 2.14), although this result was statistically ambiguous (P = 0.052). The lower number of private alleles and haplotypes within Glacier Bay was due primarily to the two medium-aged populations, Vivid and Tyndall. As with the microsatellite loci and mtDNA haplotypes, the youngest populations had a higher number of unique markers similar to that in some of the older populations outside Glacier Bay.

Heterozygosity did not provide a strong signal about the number of source populations. It did not differ across stream-age categories for the microsatellite loci, and was only minimally higher in the medium-aged populations for the allozyme loci (Figure 2.12). Haplotype diversity, a haploid measure of heterozygosity, had a pattern nearly identical to that of the microsatellite loci across the three stream-age categories (Figure 2.13B). Given that the heterozygosity of microsatellite loci and haplotype diversity in the youngest populations was not higher than that in the oldest populations, we can rule out the possibility that many genetically divergent sources contributed colonists to the youngest populations.

Linkage disequilibrium, measured as a correlation coefficient of alleles between loci (R_{ij}), was higher in the youngest populations (Figure 2.10) and can be interpreted as a signal of a bottleneck, population mixing, or immigration from populations with residual

linkage disequilibrium. At the population level, Wolf Pt., the youngest population in this study, had the highest R_{ij} value. The higher R_{ij} value could be a founder effect, but this seems unlikely, given the high allele and haplotype richness which leaves the possibilities of (1) colonization from either a donor source that had residual linkage disequilibrium, e.g., Nunatak, to which it was genetically most similar, or (2) contribution from colonists of multiple donor sources. The allele frequencies support the concept that most of the colonists to Wolf Pt. were from Nunatak; the mtDNA haplotype frequencies suggest that Wolf Pt. is more closely related to the populations outside the Bay. Perhaps the colonists were primarily from Nunatak with smaller contributions from outside populations, which would explain the higher linkage disequilibrium and allele richness, and lower relatedness in the youngest population.

Relatedness measures were generally higher in Glacier Bay populations than in older populations outside the Bay, with the highest estimates in the medium-aged populations (Figures 2.15 and 2.16; Table 2.9). However, at the population level, the youngest population, Wolf Pt., and the oldest population, N. Berg, had relatedness estimates similar to those in the older populations outside the Bay (e.g., Appendix 2.11). For the new population at Wolf Pt., the low relatedness estimate and the higher number of parents necessary to explain the genotypes present in the collection was consistent with donors from multiple source populations, or from older populations with low relatedness. The population at Nunatak had relatedness estimates that were more similar to the medium-aged populations in Glacier Bay (except N. Berg) than to the Wolf Pt population. We can speculate that the survival of offspring from the 1992 brood year was high enough to explain (at least partially) the larger population size in 1994, and thus the higher relatedness.

Did new populations start with small or large numbers of fish?

Most of the genetic diversity measures eliminated the likelihood that large numbers of colonists accessed new stream habitat, but colonization by a very small number of fish can also be ruled out because much of the genetic variation in the oldest populations was

also observed in the youngest populations. Some of the genetic signals differed between the two youngest populations, Nunatak and Wolf Pt. The higher linkage disequilibrium in both of these populations suggested that smaller numbers of fish were involved during colonization events. However, the microsatellite allele richness and haplotype richness in Wolf Pt. was similar to that in the older populations outside Glacier Bay, whereas in Nunatak these measures were lower, at a level similar to the medium-aged populations within the Bay. Higher heterozygosity of microsatellite loci in the youngest populations (Figure 2.12) and higher mtDNA haplotype diversity in Wolf Pt. eliminates the possibility that very few fish were involved in colonization, which points toward involvement of a larger number of fish or colonists from older, more variable populations.

The effective population size (N_e) estimates tended to be higher in the populations outside the Bay (Table 2.10), but even the smallest estimates within the Bay exceeded 100 fish and the 95% confidence intervals included a minimum of several dozen fish and a maximum of an infinite number for every population. The N_e estimates corroborate the genetic diversity measures. For example, the smaller point estimate of N_e for Nunatak (123 fish) matches the lower microsatellite allele and mtDNA haplotype richness in this population; the larger N_e estimate for Wolf Pt. corresponds with the higher microsatellite allele and mtDNA haplotype richness observed in this population, although the negative estimate can also be explained by sampling error (Waples 1991) due to the small sample size.

There was no evidence of a bottleneck (or founder effect) in any population for either nuclear marker type when heterozygosity was compared with that expected under mutation-drift equilibrium. A lack of bottlenecks rules out colonization by a small number of fish or prolonged persistence at small numbers. A preponderance of heterozygosity deficiency with the allozyme loci (under IAM) suggested that the populations in our study have experienced population mixing (gene flow) or population expansion (Luikart and Cornuet 1998). The IAM is a conservative model for the detection of heterozygosity deficiency, so that when it is observed with this model it is

more likely to be real, although Chakraborty et al. (1980) reported that heterozygosity deficiency was common for many allozyme datasets.

Did colonization occur as a one-time event or as a recurrent process?

Recurrent immigration (gene flow) into populations after their initial formation appears to be limited. The microsatellite allele richness and heterozygosity, the number of private alleles, and mtDNA haplotype richness and diversity was either lower or no different in the medium-aged populations than in the younger populations (Figures 2.11B, 2.12B, 2.13, and 2.14). The slightly higher heterozygosity of the medium-aged populations for the allozyme loci (Figure 2.12A) was due to fewer invariant loci as compared to the younger populations, which is consistent with a mild founder effect offset by subsequent gene flow. If gene flow subsequent to colonization is low, then gene flow over many generations is necessary to explain the similarities of the old populations outside the Bay (e.g., Figure 2.5 and Table 2.7). The time elapsed since the presumed colonization of the medium-aged populations, roughly 40–60 generations, is insufficient for these populations to have become homogeneous. Heterogeneity was observed among some of the even-year pink salmon populations in northern Southeast Alaska (Figure 2.8) which suggests that, even outside the Bay, gene flow is limited amongst some populations that have existed for possibly hundreds, if not thousands of generations.

Relatedness measures also support the idea that recurrent immigration into Glacier Bay is limited. The maximum likelihood relatedness estimate was higher in the medium-aged populations than in the older populations outside the Bay or the youngest population, Wolf Pt. (Appendix 2.11), which indicates that few fish are moving into established populations. On the other hand, the higher variance of relatedness (r_{xy}; Figure 2.15) in the medium-aged populations suggests the presence of several independent groups of related fish (Lynch and Ritland 1999), an event more likely to occur with contribution from multiple, genetically divergent sources or run-timing structure. The old populations outside the Bay had the lowest levels of relatedness (e.g.,

Figure 2.16), possibly due to larger population sizes (large N_e) and at least some gene flow.

Linkage disequilibrium was not considered to address this question because it is influenced by the number of sources and number of colonists (Table 1.1 in Kondzela 2010), but the lower linkage disequilibrium with stream age (Figure 2.10) indicates that recurrent gene flow from divergent sources into the medium-aged populations in Glacier Bay is limited. Contrary to the other medium-aged populations, N. Fingers had one of the highest R_{ij} values, which could reflect the population's small size, both census and N_e (Hedrick 2005). The generally lower linkage disequilibrium and the strong heterogeneity among the medium-aged populations indicate that gene flow subsequent to colonization did not occur at high levels.

Synthesis

Pink salmon colonization within Glacier Bay was likely episodic and the timing of deglaciation resulted in some shared colonization events that promoted small scale homogeneity (e.g., Nunatak and Wolf Pt. or Tyndall and Vivid) and larger scale heterogeneity (lower Glacier Bay and Tyndall/Vivid). Colonists appear to have been from nearest source populations in some circumstances—populations in the lower Bay were most similar to populations outside the Bay—but for other populations within Glacier Bay, the donor sources were less obvious. The youngest populations, Nunatak and Wolf Pt., were genetically similar to each other and intermediate to the oldest populations and the medium-aged populations, the latter of which suggests that multiple donor sources, both nearby and farther away, contributed colonists. If we sampled colonists early in the colonization process in Wolf Pt., then the genetic signals reflect those of the source populations. For example, even though the number of fish in Wolf Pt. was small and our sample size was small, some measures of genetic diversity (e.g., microsatellite allele richness) were as high as that observed in the older populations.

The source of Vivid and Tyndall colonists is unclear. Their divergence from other populations may represent isolation related to their final coastal migration pattern or to

their distance from other populations. For example, Tyndall spawners must turn left at Geikie Inlet, then left again at Tyndall Cove and there are few pink salmon populations in the west arm of Glacier Bay near Vivid. There were a number of alleles and haplotypes observed only within Glacier Bay, but nearly all were at low frequencies, so there is little evidence that the donor sources were from more distant, genetically diverse populations than those we sampled outside Glacier Bay. There were several potential donor sources within Glacier Bay that we were unable to sample, e.g., Bartlett River and Beartrack River on the lower east side of the Bay. Samples from these locations may have provided more information about the population structure within the Bay and a better understanding of how the youngest streams were colonized.

Streams in Glacier Bay were colonized by donors from only a few populations. The higher linkage disequilibrium in the two youngest populations indicates multiple donor sources, but the increase in disequilibrium was not large and does not support the concept that large numbers of genetically diverse populations were involved in contributing colonists. The allele and haplotype richness and number of private alleles and haplotypes in the older populations was higher than that observed in the younger populations, which suggests that the contribution of genetic material into Glacier Bay was limited to a few donor sources. As expected with this mechanism of colonization, heterozygosity of microsatellite loci and haplotype diversity in the youngest populations were no higher than in the oldest populations. The lower relatedness of fish in the youngest population, Wolf Pt., could be a sign that donors from a mix of populations colonized this stream, but it can also be explained by contribution from primarily older populations, which had lower measures of relatedness.

Differences in genetic diversity among all the populations in this study were not large, and founder effects were minimal, which implies that more than a few fish were involved in the initial colonization process, but not so many that all populations were homogeneous. Effective population size estimates were smaller within Glacier Bay than outside the Bay, but they exceeded 100 fish in every population. Once fish inhabited a stream, populations had the capacity to expand rapidly. For example, large numbers of

fish (~2,500) in Nunatak in 1994 followed the 125 fish observed in 1992. If many of the fish present in nearby Wolf Pt. in 1994 were colonists from Nunatak, it would explain the genetic similarity of these two systems (e.g., Figure 2.5).

The exchange of relatively few fish is required to arrest divergence of populations based on neutral markers (Hartl and Clark 1997). So, the divergence of Vivid and Tyndall from the other populations and the lower genetic diversity measures in the medium-aged populations indicates that gene flow subsequent to colonization is limited at least for some populations, although exchange of fish between drainages with similar allele frequencies cannot be detected with genetic data. Recurrent gene flow may be so limited within the time frame of colonization in Glacier Bay that not enough time has elapsed for the populations to have reached a stable state between gene flow and drift. The increased relatedness of medium-aged populations (except N. Berg) reinforces the concept of limited gene flow subsequent to initial colonization. N. Berg was most similar to the older populations outside Glacier Bay and differed in a number of ways from the other medium-aged populations, which suggests that this system, deglaciated before the other locations in this study, was colonized at an earlier time. One observation that is hard to reconcile with limited gene flow after colonization is the dramatic increase in numbers of pink salmon in the youngest populations. After our last sampling period in 1994, several thousand fish were counted in Wolf Pt. in 1998 (K. Monaghan, University of Birmingham, personal communication), much like the observed increase in Nunatak between 1992 and 1994. High survival of offspring from early colonists may explain much of this increase, because escapement in other northern Southeast Alaska populations was also very high during this time period (Der Hovanisian and Geiger 2005).

At the same time that new populations were developing in younger habitat and many populations in northern Southeast Alaska were experiencing all-time high escapement, some populations within Glacier Bay remained small. For example, the even-year population in N. Fingers is small based on our limited census data, which was corroborated by the small N_e (113 fish) estimate. This population had the second highest

mean linkage disequilibrium correlation coefficient (0.133), a sign that it is more influenced by drift and gene flow than larger populations would be.

In future investigations, it would be interesting to resample the populations within Glacier Bay to see how the population structure has changed over time. How has Wolf Pt. changed genetically with the large increase in population size since our study? Will the populations become more homogeneous, as the pattern of structure outside the Bay suggests will occur, or will some populations remain distinct as observed with Chilkoot and Taku river populations in northern Southeast Alaska? If population divergence diminishes within the Bay from gene flow, at what rate will the change occur? Will future periods of high pink salmon escapement in northern Southeast Alaska alter population structure within Glacier Bay: what will happen to streams not yet inhabited, and what changes will occur to the measures of genetic diversity such as relatedness, allele richness, and linkage disequilibrium of existing populations?

References

Allendorf, F. W., and G. H. Thorgaard. 1984. Tetraploidy and the evolution of salmonid fishes. Pages 1-53 *in* B. Turner, editor. Evolutionary genetics of fishes. Plenum Press, New York.

Bandelt, H. J., P. Forster, B. C. Sykes, and M. B. Richards. 1995. Mitochondrial portraits of human populations. Genetics 141:743-753.

Cavalli-Sforza, L. L., and A. W. F. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. Evolution 21:550-570.

Chakraborty, R., P. A. Fuerst, and M. Nei. 1980. Statistical studies on protein polymorphism in natural populations. III. distribution of allele frequencies and the number of alleles per locus. Genetics 94:1039-1063.

Churikov, D., and A. J. Gharrett. 2002. Comparative phylogeography of the two pink salmon broodlines: an analysis based on a mitochondrial DNA genealogy. Molecular Ecology 11:1077-1101.

Cornuet, J. M., and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001-2014.

Der Hovanisian, J. A., and H. J. Geiger. 2005. Stock status and escapement goals for salmon in southeast Alaska 2005. Alaska Department of Fish and Game, Special Publication No. 05-22, Douglas.

Di Rienzo, A., A. C. Peterson, J. C. Garza, A. M Valdes, M. Slatkin, and N. B. Freimer. 1994. Mutational processes of simple-sequence repeat loci in human populations. Proceedings of the National Academy of Sciences USA 91:3166-3170.

Estoup, A., P. Presa, F. Krieg, D. Vaiman, and R. Guyomard. 1993. $(CT)_n$ and $(GT)_n$ microsatellites: a new class of genetic markers for *Salmo trutta* L. (brown trout). Heredity 71:488-496.

Fernandez, J., and A. Toro. 2006. A new method to estimate relatedness from molecular markers. Molecular Ecology 15:1657-1667.

Gharrett, A. J., W. W. Smoker, R. R. Reisenbichler, and S. G. Taylor. 1999. Outbreeding depression in hybrids between odd- and even-broodyear pink salmon. Aquaculture 173:117-129.

Gharrett, A. J., and M. A. Thomason. 1987. Genetic changes in pink salmon (*Oncorhynchus gorbuscha*) following their introduction into the Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 44:787-792.

Gharrett, A. J., and L. A. Zhivotovsky. 2003. Migration. Pages 141-174 *in* E. M. Hallerman, editor. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland.

Greig, C., D. P. Jacobson, and M. A. Banks. 2003. New tetranucleotide microsatellites for fine-scale discrimination among endangered chinook salmon (*Oncorhynchus tshawytscha*). Molecular Ecology Notes 3:376-379.

Hartl, D. L., and Clark A. G. 1997. Principles of population genetics, 3rd edition. Sinauer Associates, Inc., Sunderland, Massachusetts.

Hawkins, S. L., N. V. Varnavskaya, E. A. Matzak, V. V. Efremov, C. M. Guthrie III, R. L. Wilmot, H. Mayama, F. Yamazaki, and A. J. Gharrett. 2002. Population structure of odd-broodline Asian pink salmon and its contrast to the even-broodline structure. Journal of Fish Biology 60:370-388.

Heard, W. R. 1991. Life history of pink salmon (*Oncorhynchus gorbuscha*). Pages 119-230 *in* C. Groot and L. Margolis, editors. Pacific salmon life histories. University of British Columbia Press, Vancouver.

Hedrick, P. W. 2005. Genetics of populations, 3rd edition. Jones and Bartlett Publishers, Boston.

Heinl, S. C., and H. J. Geiger. 2005. Chapter 4: Pink salmon stock status and escapement goals in Southeast Alaska and Yakutat. Pages 153-189 *in* J. A. Der Hovanisian, and H. J. Geiger, editors. Stock status and escapement goals for salmon in Southeast Alaska 2005. Alaska Department of Fish and Game, Special Publication No. 05-22, Douglas.

Hoffman, S. H. 1982. Northern southeastern Alaska pink salmon (*Oncorhynchus gorbuscha*) tagging investigations, 1977-1980. Alaska Department of Fish and Game, Informational Leaflet No. 196, Juneau.

Johnson, J., and K. Klein. 2009. Catalog of waters important for spawning, rearing, or migration of anadromous fishes—Southeastern region, effective June 1, 2009. Alaska Department of Fish and Game, Special Publication No. 09-04, Anchorage.

Kalinowski, S. T., A. P. Wagner, and M. L. Taper. 2006. ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. Molecular Ecology Notes 6:576-579.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, England.

Kondzela, C. M. 2010. The colonization mechanism of pink salmon populations in Glacier Bay, Alaska, based on genetic data. Doctoral dissertation. University of Alaska Fairbanks.

Luikart, G., and J. M. Cornuet. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology 12:288-237.

Lynch, M., and K. Ritland. 1999. Estimation of pairwise relatedness with molecular markers. Genetics 152:1753-1766.

Nei, M., and F. Tajima. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97:145-163.

Olsen, J. B., P. Bentzen, M. A. Banks, J. B. Shaklee and S. Young. 2000. Microsatellites reveal population identity of individual pink salmon to allow supportive breeding of a population at risk of extinction. Transactions of the American Fisheries Society 129:232-242.

Piry, S., G. Luikart, and J-M. Cornuet. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity 90:502-503.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): population genetics 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.

Ritland, K. 2000. Marker-inferred relatedness as a tool for detecting heritability in nature. Molecular Ecology 9:1195-1204.

Scribner, K. T., J. R. Gust, and R. L. Fields. 1996. Isolation and characterization of novel salmon microsatellite loci: cross-species amplification and population genetic applications. Canadian Journal of Fisheries and Aquatic Sciences 53:833-841.

Smith, C. T., B. F. Koop, and R. J. Nelson. 1998. Isolation and characterization of coho salmon (*Oncorhynchus kisutch*) microsatellites and their use in other salmonids. Molecular Ecology 7:1614-1616.

Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535-538.

Waples, R. S. 1988. Estimation of allele frequencies at isoloci. Genetics 118:371-384.

Waples, R. S. 1989. A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics 121:379-391.

Waples, R. S. 1991. Genetic methods for estimating the effective population size of cetacean populations. Pages 279-300 *in* A. R. Hoelzel, editor. Genetic ecology of whales and dolphins. Report of the International Whaling Commission, Special Issue 13, Cambridge, UK.

Waples, R. S., and C. Do. 2008. LDNE: a program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8:753-756.

Waples, R. S., M. Masuda, and J. Pella. 2006. SALMONNb: a program for computing cohort-specific effective population sizes (N_b) in Pacific salmon and other semelparous species using the temporal method. Molecular Ecology Notes 7:21-24.

Weir, B. S., and C. C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358-1370.

Table 2.1. Location of sample collections of even-year pink salmon within and near Glacier Bay, Alaska. Water body number (Johnson and Klein 2009) is given below the location name. Years before present (BP). The name of the nearest geographic feature or the colloquial name was used for watersheds without official names. Abbreviated names used in figures in parentheses.

Location	Latitude	Longitude	*n*	Collection date	Stream age[a]
Youngest populations, upper Glacier Bay, systems ice-free less than 60 years					
Wolf Point (WO) 114-77-11000	59.00	-136.16	41	30 Aug, 4 Sep 1994	1947
Nunatak (NU) 114-77-10350	58.98	-136.10	100	30, 31 Aug 1994	1935
Medium-aged populations, mid-lower Glacier Bay, systems ice-free less than 120 years					
Vivid (VI) 114-75-10960	58.85	-136.49	100	3 Sep 1994	ca. 1870
Tyndall (TY) 114-73-10150	58.59	-136.36	99 100	4 Sep 1990 25 Aug 1992	ca. 1875
Fingers Bay (NF) north head of north arm 114-72-10170	58.60	-136.23	100	22, 27 Aug 1992	ca. 1845
N. Berg Bay (NB) 114-71-10320	58.52	-136.24	100	22, 23, 27 Aug 1992	ca. 1830
Oldest populations, outside Glacier Bay, systems ice-free during Neoglaciation					
E. Kahtaheena (EK) 114-23-10240	58.41	-135.57	100	8 Sep 1990	>13,000 yr BP
Dundas Bay (DU) head of north arm 114-60-10660	58.45	-136.52	99	7 Sep 1994	>13,000 yr BP
Homeshore (HM) 114-25-10100	58.31	-135.37	100	31 Aug 1992	>13,000 yr BP
Spasski (SP) 114-27-10300	58.09	-135.28	100	9 Aug 1992	>13,000 yr BP

[a]Estimated date stream mouth exposed from deglaciation (G. Steveler and C. Soiseth, National Park Service, personal communication).

Table 2.2. Microsatellite loci, repeat type, annealing temperature, number of alleles (n_a), allele size range, unbiased expected heterozygosity (H_E), F_{ST}, and primer source used for even-year pink salmon collected within and near Glacier Bay, Alaska.

Locus	Type	Temperature °C	n_a	Size range	H_E	F_{ST}[a]	Source[b]
One102	tetra-	63	29	191 - 311	0.921	0.015	1
One103	tetra-	58	53	157 - 385	0.952	0.008	1
One109	tetra-	63	20	130 - 210	0.900	0.011	1
One111	tetra-	58	22	189 - 273	0.922	0.012	1
One μ13	di-	50	2	141 - 147	0.211	0.009	2
Ots208-1	tetra-	50	23	119 - 207	0.909	0.013	3
Ots208-2	di-	50	2	291 - 293	0.443	0.005	3
μSat60	di-	59	7	105 - 121	0.233	0.011	4
Oki10	di-, tetra-	63	43	139 - 349	0.935	0.008	5
Over all loci						0.010	

[a] Weir and Cockerham (1984).

[b] 1 = Olsen et al. (2000)
2 = Scribner et al. (1996)
3 = Greig et al. (2003)
4 = Estoup et al. (1993)
5 = Smith et al. (1998).

Table 2.3. Mitochondrial DNA restriction sites examined in even-year pink salmon as shown in Figure 2.2.

Site	mtDNA region	Restriction endonuclease	Position of binary code[a]
1	ND1/ND2	*Bst* N I	1
2	ND1/ND2	*Bst* N I	3
9	ND3/ND4	*Rsa* I	1
10	ND3/ND4	*Rsa* I	2
18	Cyt*b* /D-loop	*Sau* 96 I	1
19	Cyt*b* /D-loop	*Sau* 96 I	2
20	Cyt*b* /D-loop	*Sau* 96 I	1,2
21	Cyt*b* /D-loop	*Sau* 96 I	4
22	Cyt*b* /D-loop	insert +80 bp	1
23	Cyt*b* /D-loop	insert +160 bp	1
24	Cyt*b* /D-loop	insert + 260 bp	1
25	Cyt*b* /D-loop	insert + 520 bp	1
26	ND5/ND6	*Bst* N I	1
27	ND5/ND6	*Bst* U I	1
28	ND5/ND5	*Bst* U I	2
29	ND5/ND6	*Bst* U I	1,2
30	ND5/ND6	*Hin* f I	1
31	ND5/ND6	*Hin* f I	3
32	ND5/ND6	*Hin* f I	4
33	Cyt*b* /D-loop	*Msp* I	1
34	Cyt*b* /D-loop	*Msp* I	2
35	Cyt*b* /D-loop	*Msp* I	1,2
36	Cyt*b* /D-loop	*Msp* I	3
37	Cyt*b* /D-loop	*Msp* I	4
38	Cyt*b* /D-loop	*Rsa* I	1

[a]See Appendix 2.4.

Table 2.4. Even-year pink salmon composite mtDNA haplotypes derived from haplotypes of specific mtDNA regions cut with restriction endonucleases (see Appendix 2.4). Composite haplotypes denoted in upper case letters match Churikov and Gharrett (2002); those in lowercase letters indicate new haplotypes. Letters in parentheses denote presumed haplotypes.

Composite	Region	ND5/ND6			Cyt*b*/D-loop				ND3/ND4	ND1/ND2
haplotype[a]	Enzyme	*Bst* N I	*Bst* U I	*Hin* f I	*Msp* I	*Rsa* I	*Sau* 96 I	insert	*Rsa* I	*Bst* N I
A*		a	a	a	a	a	a	none	(a)	a
C*		a	a	a	a	a	b	none	(a)	(a)
D*		a	a	a	a	a	c	none	(a)	(a)
H*		a	a	b	a	a	a	none	(a)	(a)
M*		a	a	a	b	a	b	none	(a)	(a)
V*		b	a	a	a	a	a	none	a	(a)
X		b	a	a	a	a	a	none	b	(a)
AA*		a	b	a	a	a	a	none	(a)	b
BD		a	b	a	a	a	b	none	(a)	(b)
o		a	a	a	a	a	f	none	a	(a)
p		a	a	a	c	a	a	none	a	(a)
q		a	a	a	g	a	a	none	a	(a)
r		a	a	b	e	a	a	none	(a)	(a)
s		a	a	e	b	a	b	none	(a)	(a)
u		a	d	a	a	a	a	none	(a)	a
v		b	a	a	e	a	a	none	a	(a)
w		a	a	a	a	a	a	+520 bp	(a)	(a)
x		a	b	a	a	a	a	+260 bp	(a)	(a)
y		b	a	a	a	a	a	+80 bp	(a)	(a)
z		b	a	a	a	a	a	+160 bp	(a)	(a)

[a] A* = A,a,b
C* = C,BV
D* = D,BB
H* = H,I,AQ,AX,AZ
M* = M,N,AP,BJ,BK,BZ
V* = V,AV
AA* = AA,AB,AD,AE,AF,AG,AH,AI,AJ,AL,AT,AW,BA,BC,BO,BP,BS,BT,CA,CD,CE,e.

Table 2.5. Number of even-year pink salmon samples (n), composite mtDNA haplotype frequencies, number of haplotypes per collection (n_c), haplotype richness (h_{rich}), and haplotype diversity (h). Upper case letters match Churikov and Gharrett (2002)[a]; lowercase letters indicate new haplotypes.

	n	A*	C*	D*	H*	M*	V*	X	AA*	BD	o	p	q	r	s	u
Youngest populations, upper Glacier Bay																
Wolf Pt. 1994	40	0.450	0.025	0.050	0.175	0.025	0.100	0	0.125	0	0.050	0	0	0	0	0
Nunatak 1994	40	0.400	0	0	0.250	0	0.175	0	0.125	0	0	0.025	0.025	0	0	0
Medium-aged populations, mid-lower Glacier Bay																
Vivid 1994	40	0.400	0	0	0.350	0.025	0.150	0.025	0.050	0	0	0	0	0	0	0
Tyndall 1990	40	0.425	0	0	0.450	0	0	0.025	0.075	0	0	0	0	0	0	0
Tyndall 1992	40	0.375	0.025	0	0.425	0	0.075	0	0.100	0	0	0	0	0	0	0
N. Fingers 1992	40	0.625	0	0	0.125	0	0.050	0	0.175	0	0	0	0	0	0	0
N. Berg 1992	40	0.550	0.025	0	0.100	0	0.075	0	0.225	0	0	0	0	0	0	0
Oldest populations, outside Glacier Bay																
Dundas 1994	39	0.359	0	0.051	0.205	0	0.128	0	0.179	0.026	0	0	0	0.026	0.026	0
E. Kahtaheena 1990	40	0.375	0.050	0.025	0.325	0	0.050	0	0.175	0	0	0	0	0	0	0
Homeshore 1992	40	0.450	0	0	0.175	0	0.175	0	0.150	0	0	0	0	0	0	0.025
Spasski 1992	40	0.300	0.050	0.025	0.225	0	0.150	0	0.150	0	0.050	0	0	0	0	0

[a] A* = A,a,b H* = H,I,AQ,AX,AZ
C* = C,BV M* = M,N,AP,BJ,BK,BZ
D* = D, BB V* = V,AV
AA* = AA,AB,AD,AE,AF,AG,AH,AI,AJ,AL,AT,AW,BA,BC,BO,BP,BS,BT,CA,CD,CE,e.

Table 2 5 Continued

	v	w	x	y	z	n_c	h_{rich}	h
Youngest populations, upper Glacier Bay								
Wolf Pt 1994	0	0	0	0	0	8	7 950	0 7538
Nunatak 1994	0	0	0	0	0	6	5 950	0 7487
Medium-aged populations mid-lower Glacier Bay								
Vivid 1994	0	0	0	0	0	6	5 950	0 7090
Tyndall 1990	0	0	0 025	0	0	5	4 950	0 6256
Tyndall 1992	0	0	0	0	0	5	4 975	0 6795
N Fingers 1992	0	0	0	0 025	0	5	4 975	0 5744
N Berg 1992	0	0	0	0	0 025	6	5 950	0 6462
Oldest populations outside Glacier Bay								
Dundas 1994	0	0	0	0	0	8	8 000	0 7962
E Kahtaheena 1990	0	0	0	0	0	6	5 975	0 7359
Homeshore 1992	0 025	0	0	0	0	6	5 950	0 7308
Spasski 1992	0	0 050	0	0	0	8	7 975	0 8269

Table 2.6. The restriction endonuclease fragments within which the Cyt*b* /D-loop region insertions are located or the size of the extra fragment observed for specific samples. Fragments are listed in Appendix 2.5; the position of double-digest fragments from *Ban* II, *Mbo* I, and *Hin* f I are shown in Appendix 2.1.

	Restriction endonuclease					
Sample	*Rsa* I	*Msp* I	*Sau* 96 I	*Ban* II	*Mbo* I	*Hin* f I
Tyndall 1990 #10	560	extra ~250 bp	1630	1665	1350	extra 260 bp
Spasski 1992 #11,12	560	extra ~500 bp	1630	1665	1350	extra 260 bp[a]
N. Berg 1992 #30	extra ~80 bp[b]	235	265	1110	580	705
N. Fingers 1992 #31	extra ~80 bp	235	265	1110	580	705

[a] Two copies of the 260 bp insert.

[b] Two copies of the 80 bp insert.

Table 2.7. Fixation indices and probability of pseudo-exact tests used to test homogeneity of even-year pink salmon collections within and near Glacier Bay, Alaska, for three stream-age categories. Based on nuclear locus allele frequencies and mtDNA haplotype frequencies. Probability values for the nuclear loci were pooled across loci with Fisher's method.

	Allozyme		Microsatellite		mtDNA	
Location	F_{ST}	P	F_{ST}	P	F_{ST}	P
Youngest populations, upper Glacier Bay						
Wolf Pt. 1994						
Nunatak 1994						
Among New	-0.0004	0.5298	0.0004	0.1653	-0.0110	0.4432
Medium-aged populations, mid-lower Glacier Bay						
Vivid 1994						
Tyndall 1990	0.0047[a]	0.7162	-0.0004	0.2324	-0.0167	0.4528
Tyndall 1992						
N. Fingers 1992						
N. Berg 1992						
Among Medium[b]	0.0146	<0.0001	0.0176	<0.0001	0.0567	<0.0001
Oldest populations, outside Glacier Bay						
Dundas 1994						
E. Kahtaheena 1990						
Homeshore 1992						
Spasski 1992						
Among Old	0.0009	0.7151	0.0005	0.2737	-0.0063	0.4871
Among Stream Age[c]	0.0065	<0.0001	0.0055	<0.0001	0.0037	0.0322

[a] *sAAT-4* locus removed--data unavailable for Tyndall 1990.

[b] Multi-year collections pooled within locations.

[c] Populations pooled within stream-age.

Table 2.8. Fixation indices and probability of pseudo-exact tests used to test homogeneity of even-year pink salmon collections within and near Glacier Bay, Alaska, for four stream-age categories. Based on nuclear locus allele frequencies and mtDNA haplotype frequencies. Probability values for the nuclear loci were pooled across loci with Fisher's method.

	Allozyme		Microsatellite		mtDNA	
Location	F_{ST}	P	F_{ST}	P	F_{ST}	P
Youngest populations, upper Glacier Bay						
Wolf Pt. 1994						
Nunatak 1994						
Among New	-0.0004	0.5298	0.0004	0.1653	-0.0110	0.4432
Medium-young populations, mid-upper Glacier Bay						
Vivid 1994						
Tyndall 1990	0.0047[a]	0.7162	-0.0004	0.2324	-0.0167	0.4528
Tyndall 1992						
Among Medium-Young[b]	0.0138	<0.0001	0.0055	<0.0001	-0.0017	0.2321
Medium-old populations, mid-lower Glacier Bay						
N. Fingers 1992						
N. Berg 1992						
Among Medium-Old	0.0098	<0.0001	0.0118	<0.0001	-0.0160	0.8701
Among Medium[b, c]	0.0101	<0.0001	0.0183	<0.0001	0.0892	<0.0001
Oldest populations, outside Glacier Bay						
Dundas 1994						
E. Kahtaheena 1990						
Homeshore 1992						
Spasski 1992						
Among Old	0.0009	0.7151	0.0005	0.2737	-0.0063	0.4871
Among Stream Age[c]	0.0065	<0.0001	0.0055	<0.0001	0.0037	0.0322

[a] *sAAT-4* locus removed--data unavailable for Tyndall 1990.

[b] Multi-year collections pooled within locations.

[c] Populations pooled within stream-age.

Table 2.9. Pairwise relatedness coefficients (r_{xy}) of even-year pink salmon from moment and maximum-likelihood (ML) estimators. The observed variance of r_{xy} is ranked with the permuted variance of r_{xy} from the allele resampling procedure. Number (n) of fish whose genotypes were consistent with being either unrelated or related (half- or full-sib) to at least one other fish within respective collections. The number of parents required to produce the genotypes present within each collection based on genealogies reconstructed from the molecular marker data.

	Lynch and Ritland (1999)				Kalinowski et al. (2006)			Fernandez and Toro (2006)		
		Obs.	Perm.	% Rank		*n* individuals		Collection	*n*	Parent:collection
Location	Mean r_{xy} [a]	var. r_{xy}	var. r_{xy}	obs. var.	ML mean r_{xy} [b]	Unrelated	Related	size	parents	ratio
Youngest populations, upper Glacier Bay										
Wolf Pt. 1994	-0.0253	0.0058	0.0051	0.9685	0.0302	16	25	41	30	0.73
Nunatak 1994	-0.0205	0.0076	0.0057	> highest	0.0390	10	40	50	37	0.74
Medium-aged populations, mid-lower Glacier Bay										
Vivid 1994	-0.0205	0.0083	0.0060	> highest	0.0369	6	44	50	37	0.74
Tyndall 1990	-0.0206	0.0070	0.0064	0.9593	0.0420	11	39	50	39	0.78
Tyndall 1992	-0.0206	0.0074	0.0066	0.9859	0.0386	9	41	50	34	0.68
N. Fingers 1992	-0.0205	0.0095	0.0057	> highest	0.0442	10	40	50	38	0.76
N. Berg 1992	-0.0204	0.0061	0.0049	> highest	0.0327	12	38	50	34	0.68
Oldest populations, outside Glacier Bay										
Dundas 1994	-0.0209	0.0049	0.0046	0.9185	0.0319	19	31	49	38	0.78
E. Kahtaheena 1990	-0.0205	0.0055	0.0052	0.9115	0.0328	20	30	50	38	0.76
Homeshore 1992	-0.0204	0.0052	0.0045	0.9964	0.0336	8	42	50	40	0.80
Spasski 1992	-0.0216	0.0048	0.0058	0.1192	0.0331	17	33	49	36	0.73

[a] Median and mode were less than zero for all collections.

[b] Median and mode were zero for all collections.

Table 2.10. Estimates of effective population size (N_e) based on nine microsatellite loci and the variable allozyme loci for even-year pink salmon collections within and near Glacier Bay, Alaska. Estimates represent the N_e of the previous parental brood-year. 95% confidence intervals are given in parentheses.

	SALMONNb[a]	LDNE[b]
Location	Temporal	Linkage disequilibrium
Youngest populations, upper Glacier Bay		
Wolf Pt. 1994		-65 (111 - ∞)
Nunatak 1994		123 (34 - ∞)
Medium-aged populations, mid-lower Glacier Bay		
Vivid 1994		116 (31 - ∞)
Tyndall 1990		818 (47 - ∞)
Tyndall 1992	215 (49 - ∞)	-355 (54 - ∞)
N. Fingers 1992		113 (32 - ∞)
N. Berg 1992		327 (37 - ∞)
Oldest populations, outside Glacier Bay		
Dundas 1994		636 (41 - ∞)
E. Kahtaheena 1990		-2191 (46 - ∞)
Homeshore 1992		307 (42 - ∞)
Spasski 1992		-601 (52 - ∞)

[a] Waples et al. (2006).

[b] Waples and Do (2008).

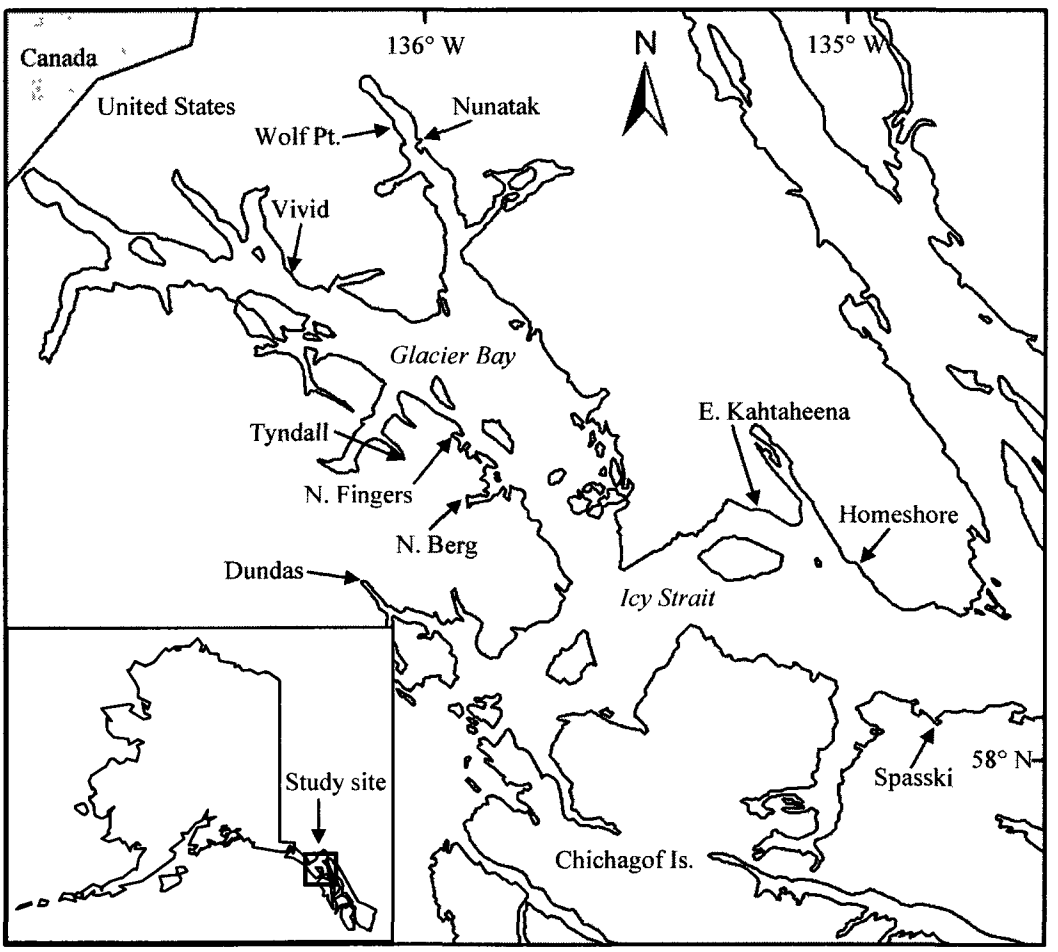


Figure 2.1. Location of even-year pink salmon collections within and near Glacier Bay, Alaska.

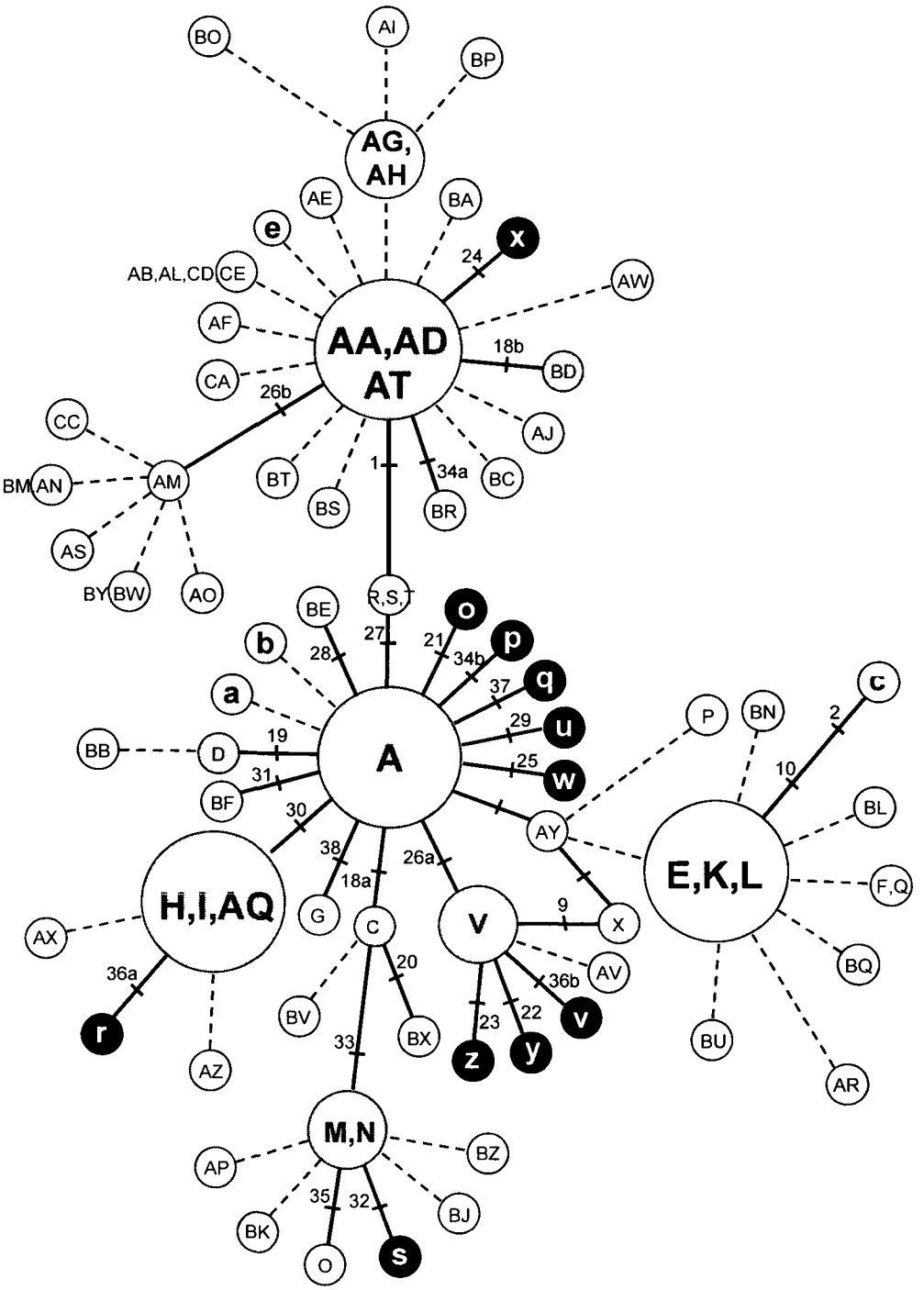


Figure 2.2. Restriction endonuclease sites examined in even-year pink salmon. Observed haplotypes in gray and black (new) circles. Dashed lines indicate sites not examined. Details in Tables 2.3 and 2.4.

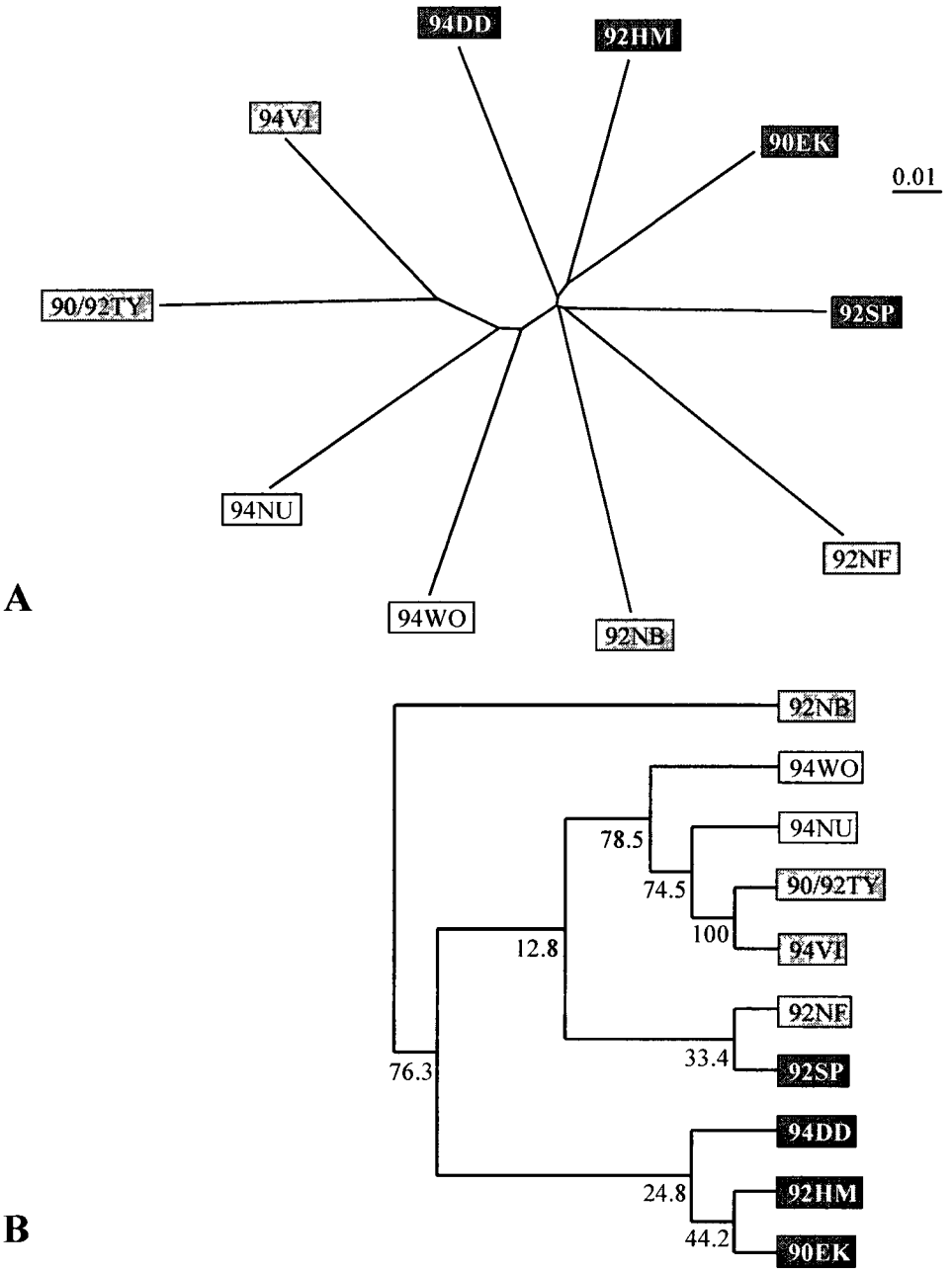


Figure 2.3. Neighbor-joining trees from nuclear allele frequencies for even-year pink salmon populations within and near Glacier Bay, Alaska. Trees made with chord distances (Cavalli-Sforza and Edwards 1967) from 9 microsatellite and 49 allozyme loci: **(A)** neighbor-joining tree, and **(B)** consensus neighbor-joining tree from 1000 bootstrap re-samplings of allele frequencies. Numbers at nodes indicate percentage the group occurred among the trees.

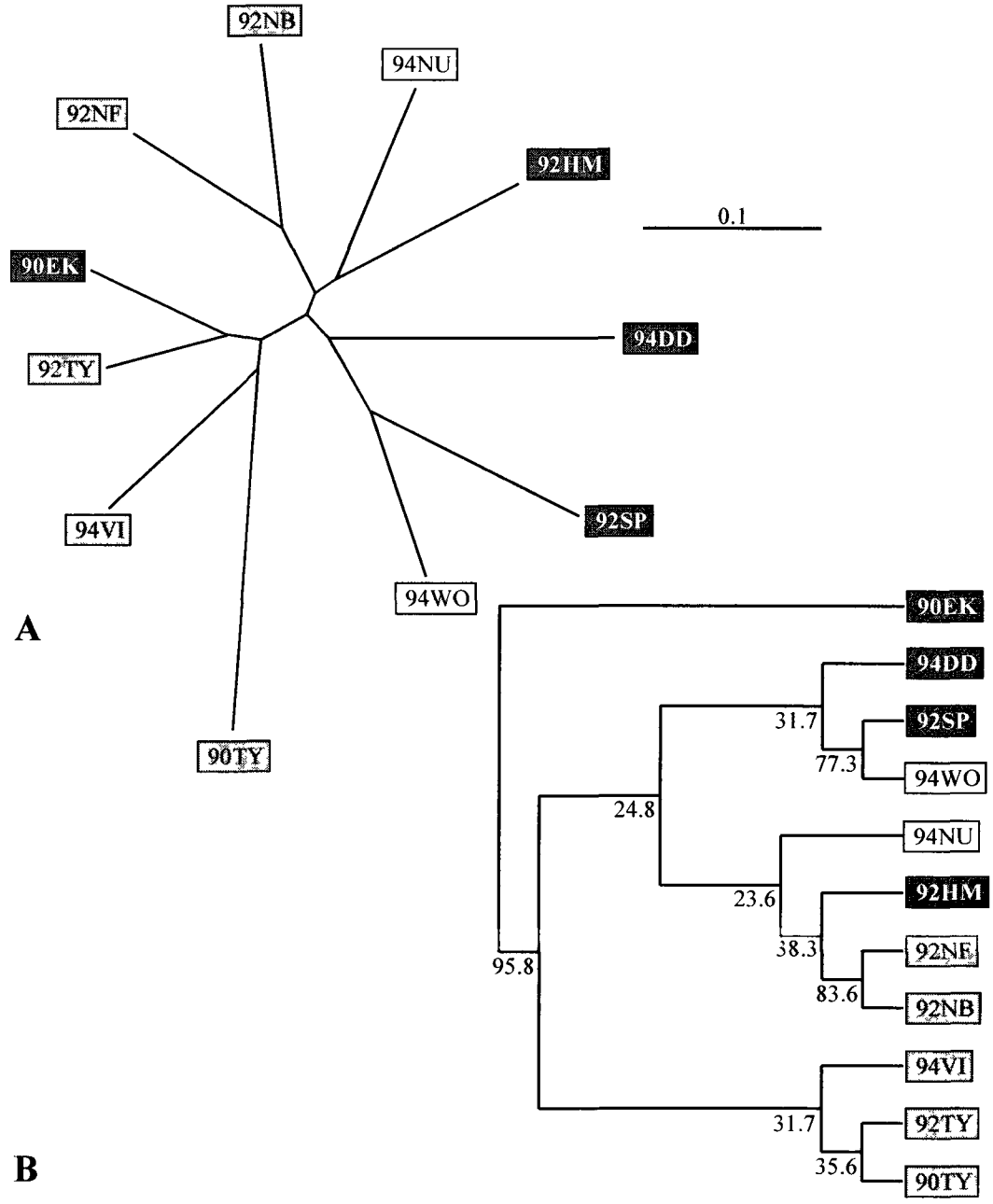


Figure 2.4. Neighbor-joining trees from mtDNA haplotype frequencies for even-year pink salmon populations within and near Glacier Bay, Alaska. Trees made with chord distances (Cavalli-Sforza and Edwards 1967): **(A)** neighbor-joining tree, and **(B)** consensus neighbor-joining tree from 1000 bootstrap re-samplings of haplotype frequencies. Numbers at nodes indicate percentage the group occurred among the trees.

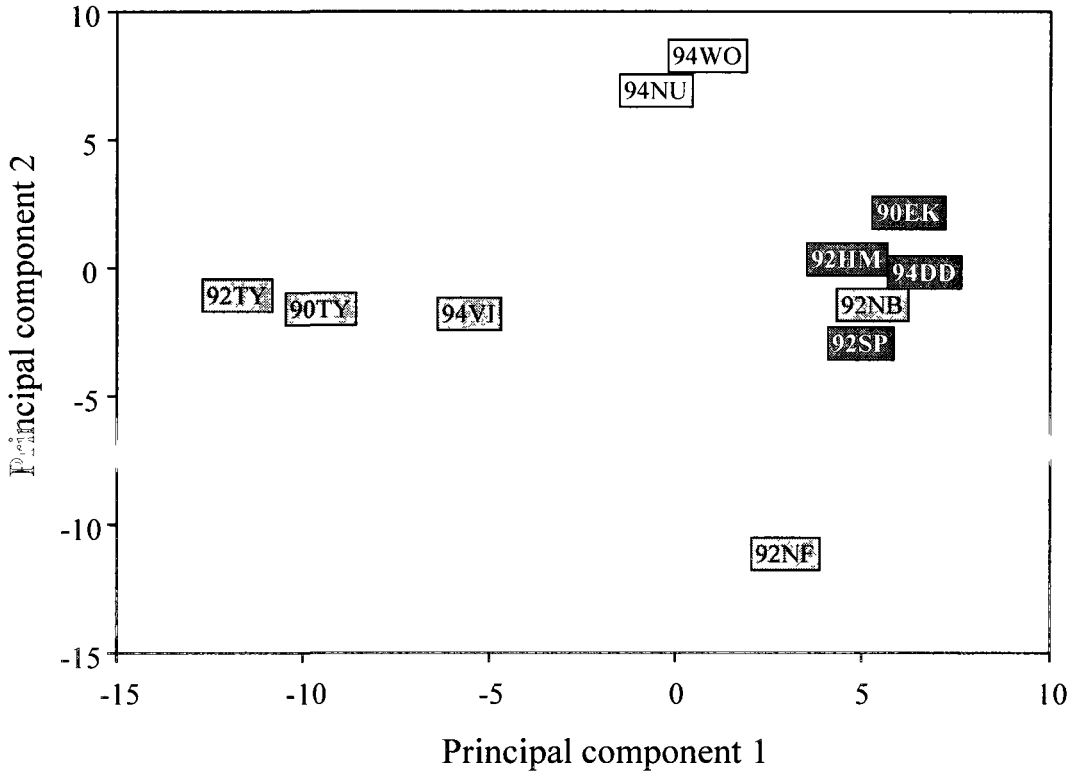


Figure 2.5. The first two principal components based on 37 allozyme and 144 microsatellite allele frequencies for the even-year pink salmon collections within and near Glacier Bay, Alaska.

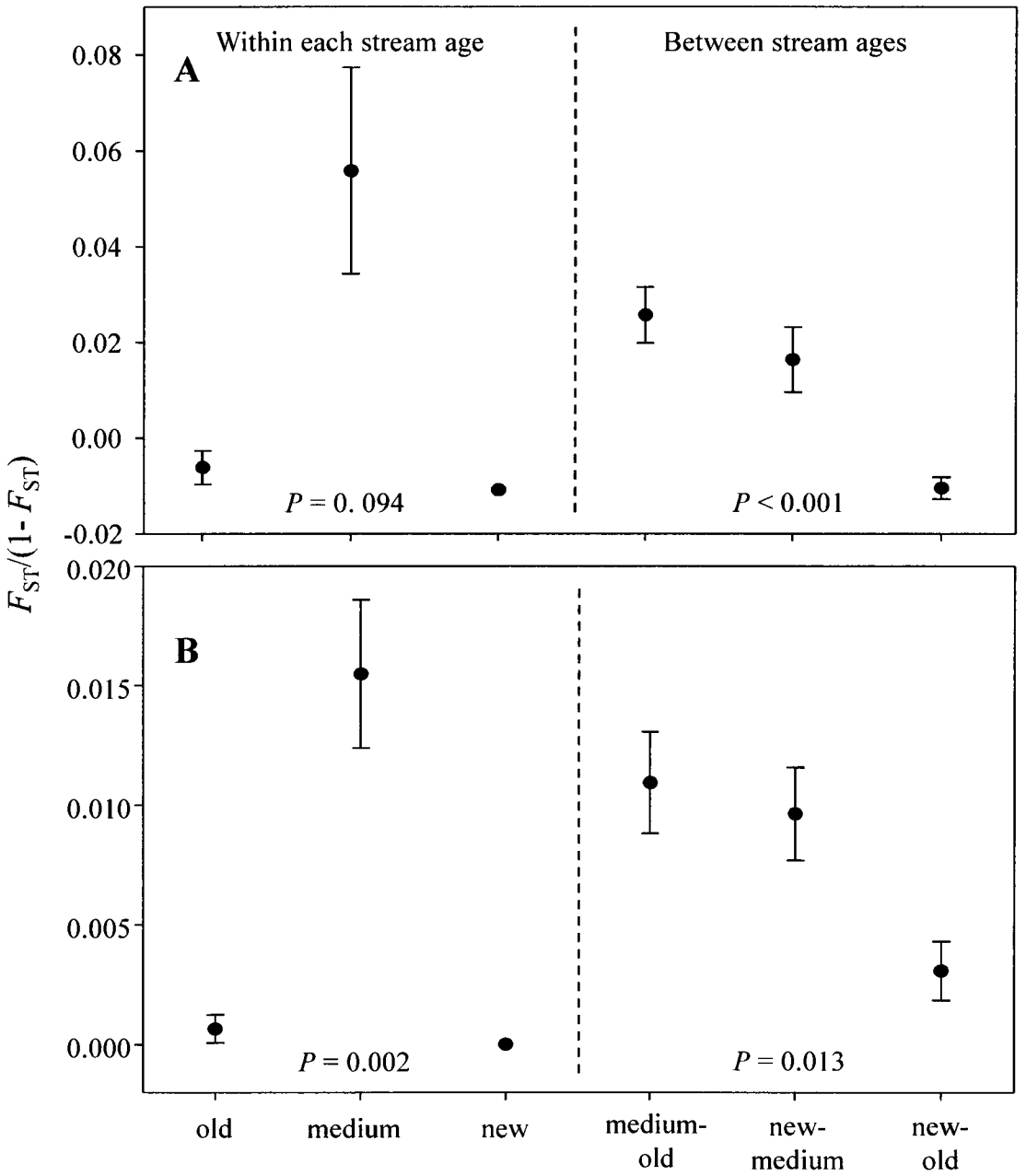


Figure 2.6. Pairwise fixation indices for even-year pink salmon populations from three stream-age categories. Mean and SE for **(A)** mtDNA and **(B)** microsatellite and allozyme loci. Multiple year collections pooled within location. Within stream-age categories on the left; between stream-age categories on the right. *P*-values from Kruskal-Wallis ANOVA on ranks.

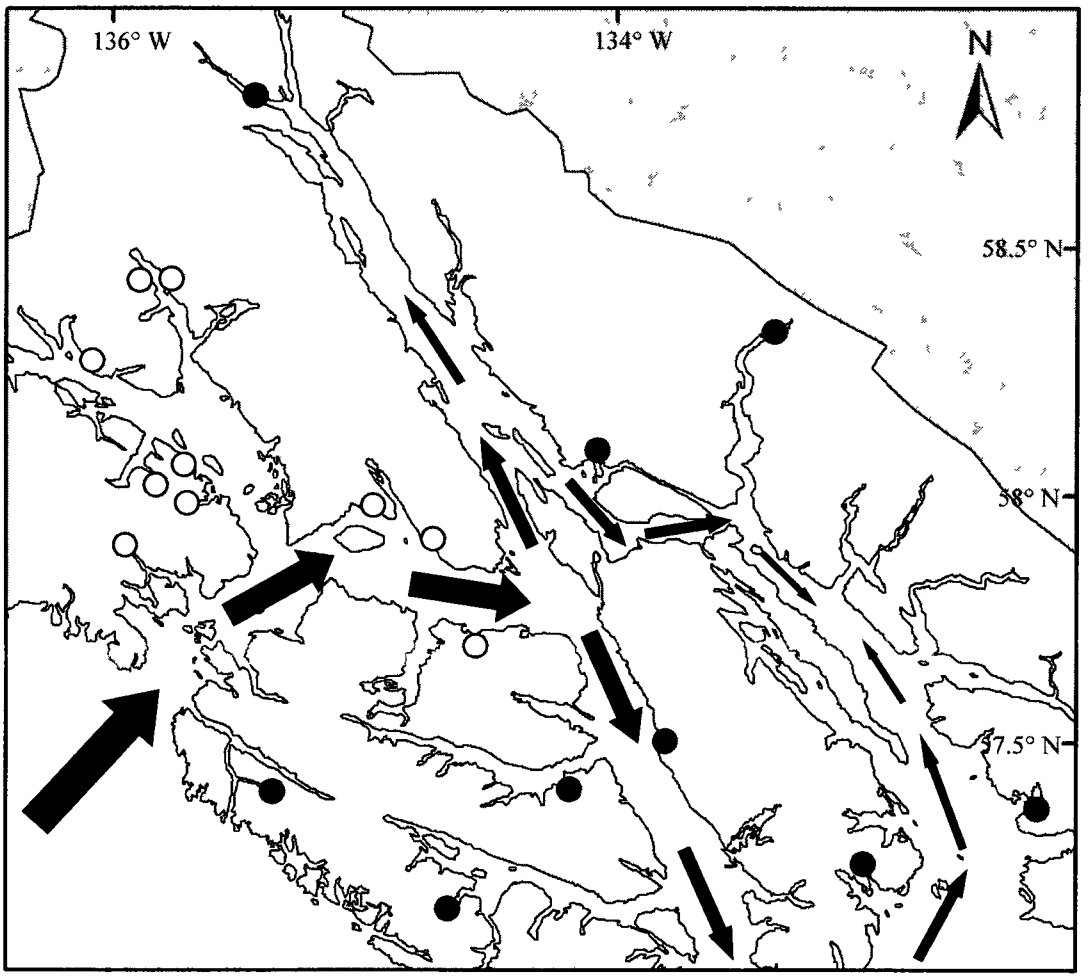


Figure 2.7. Location of even-year pink salmon collections: (1) within and near Glacier Bay, Alaska (open circles), and (2) from nine additional populations from northern Southeast Alaska for comparison (closed circles). Arrows indicate major migration routes for salmon returning to coastal waters (Hoffman 1982).

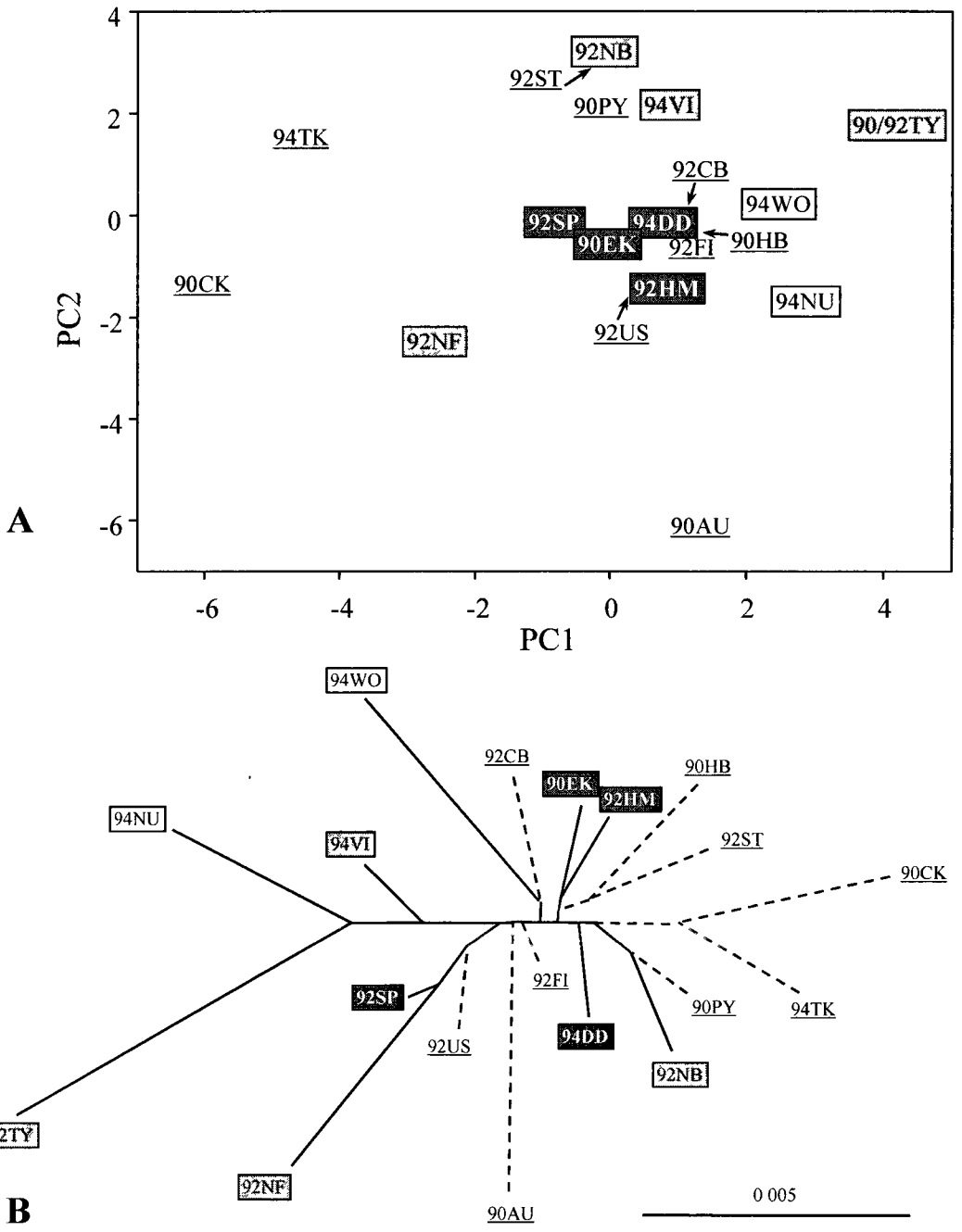


Figure 2.8. Population structure analyses for even-year pink salmon populations within and near Glacier Bay, Alaska, and nine additional populations (underlined/dashed lines) from northern Southeast Alaska based on 30 allozyme loci. **(A)** principal component analysis, and **(B)** neighbor-joining tree from chord distances (Cavalli-Sforza and Edwards, 1967).

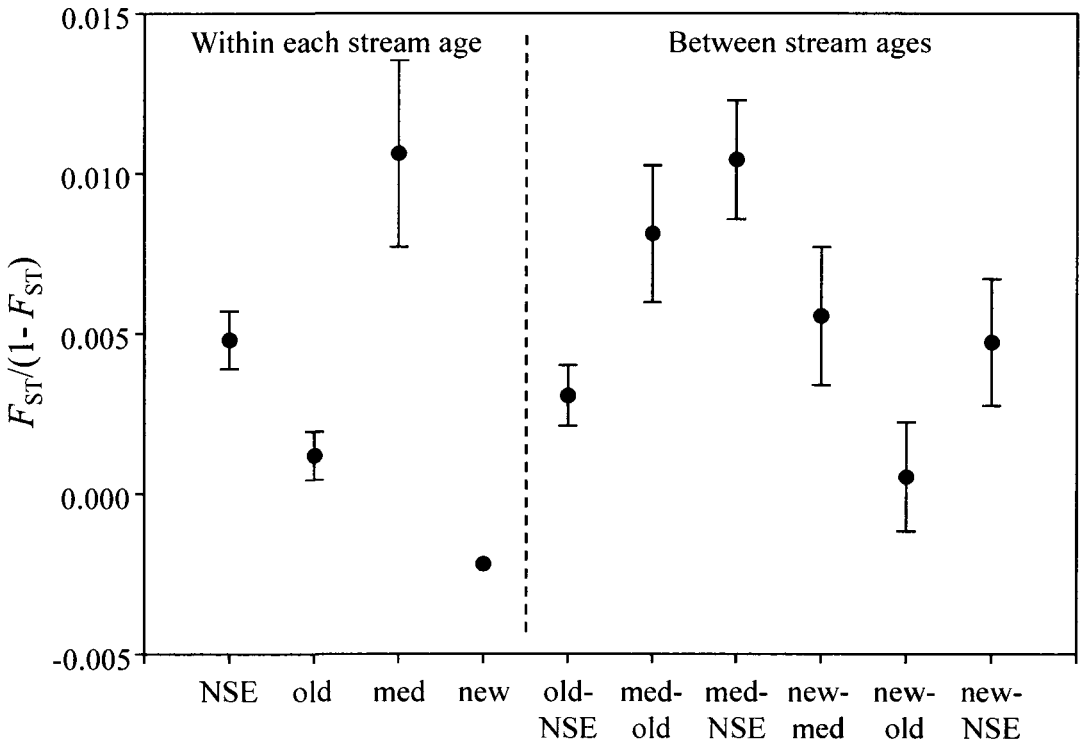


Figure 2.9. Pairwise fixation indices for even-year pink salmon populations within and near Glacier Bay, and nine populations from northern Southeast Alaska from 23 variable allozyme loci. Mean and SE for three stream-age categories—old, medium (med), new; multiple year collections pooled within location. Within stream-age categories on the left; between stream-age categories on the right.

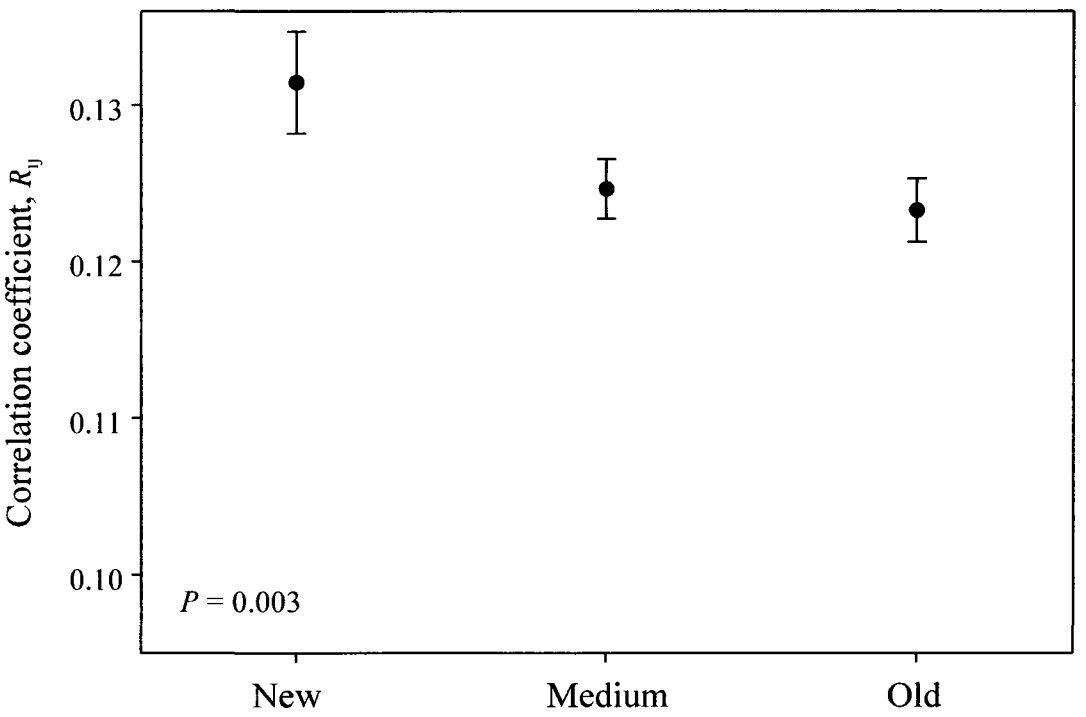


Figure 2.10. Linkage disequilibrium correlation coefficients for 136 locus-pairs for even-year pink salmon collections from three stream-age categories (mean, SE). *P*-values from Kruskal-Wallis ANOVA on ranks.

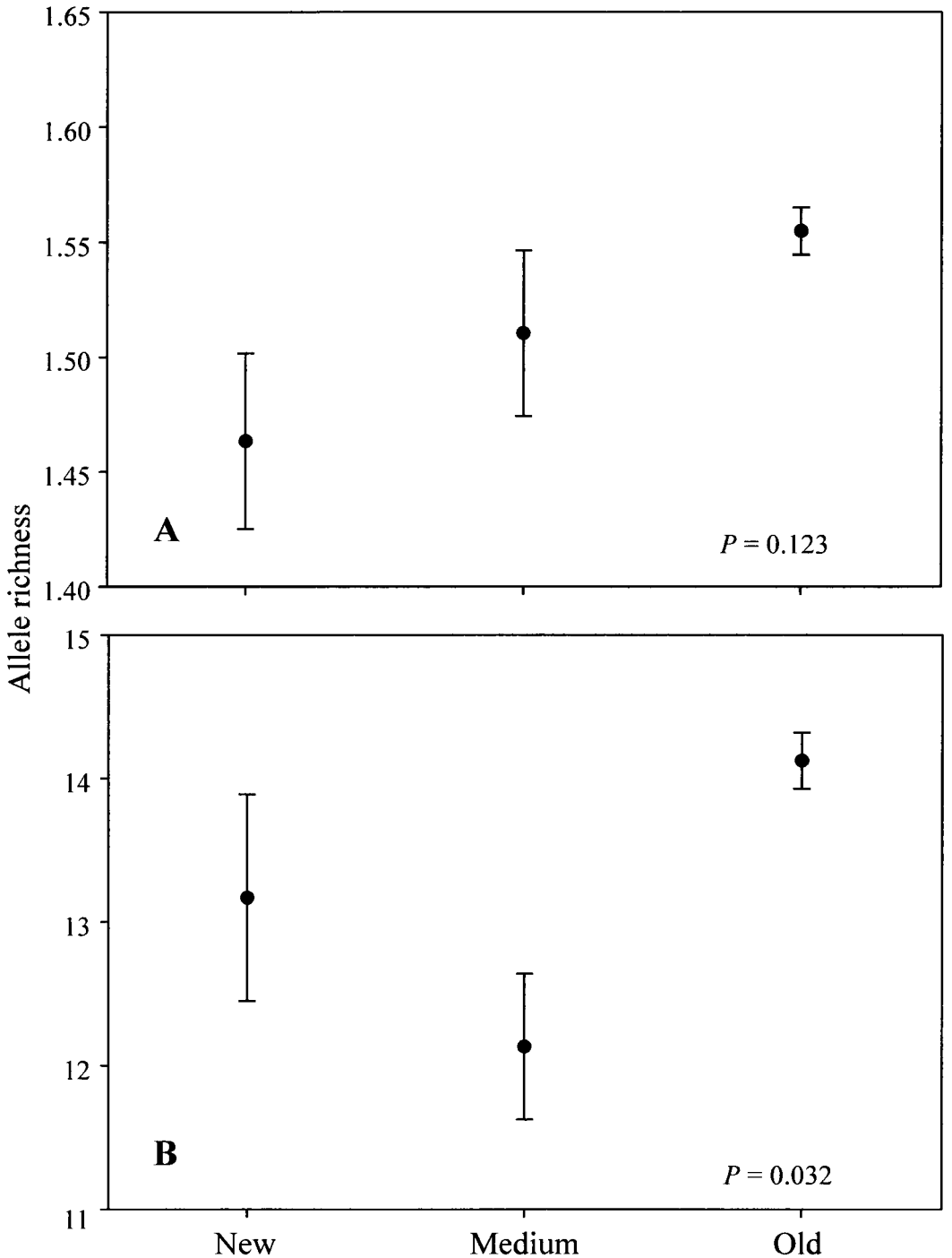


Figure 2.11. Allele richness per locus for even-year pink salmon collections from three stream-age categories. Mean and SE for **(A)** 40 allozyme loci (*sAAT-4* not included) and **(B)** 9 microsatellite loci. *P*-values from Kruskal-Wallis ANOVA on ranks.

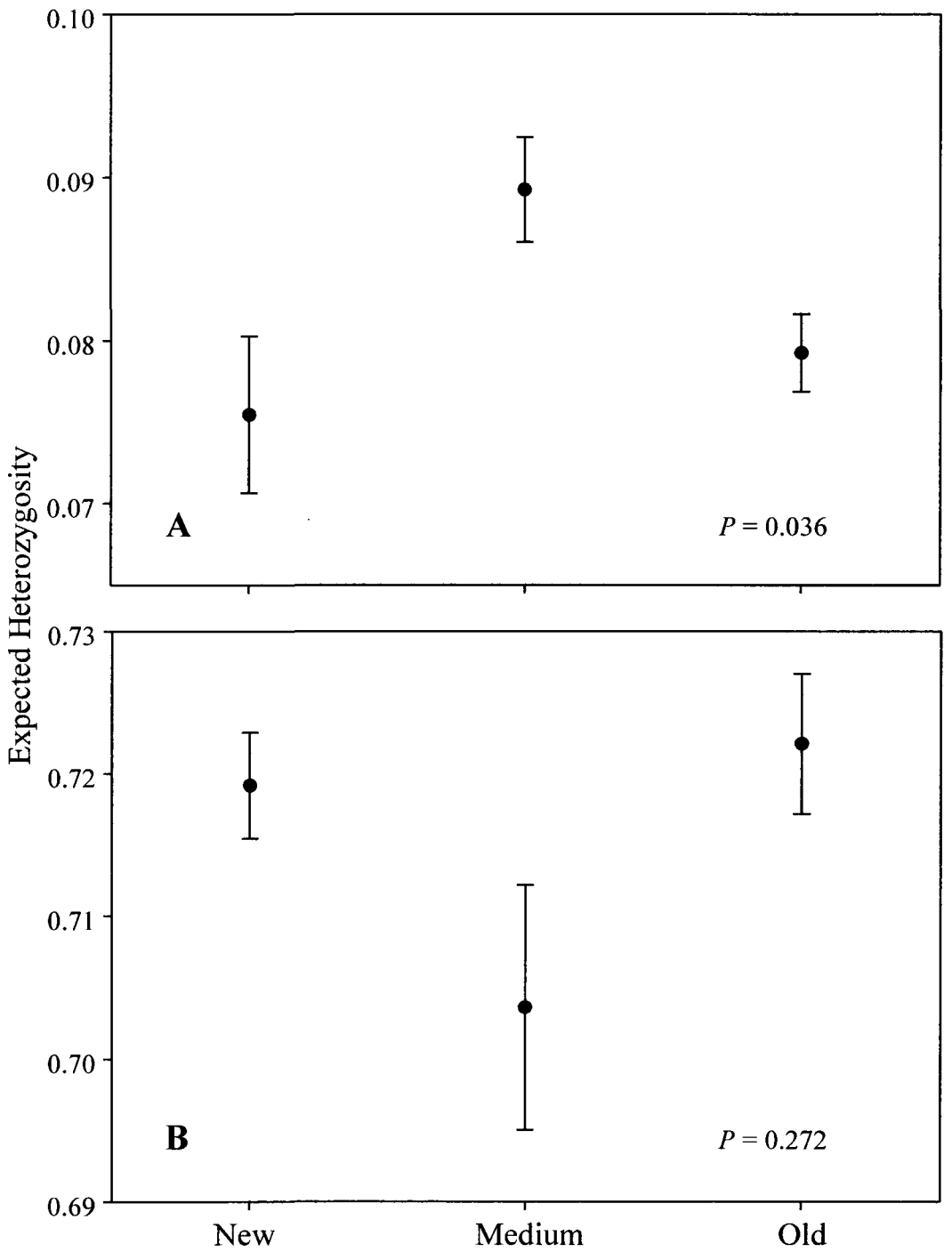


Figure 2.12. Expected heterozygosity for even-year pink salmon collections from three stream-age categories. Mean and SE for **(A)** 40 allozyme loci (*sAAT-4* not included), and **(B)** 9 microsatellite loci. *P*-values from Kruskal-Wallis ANOVA on ranks.

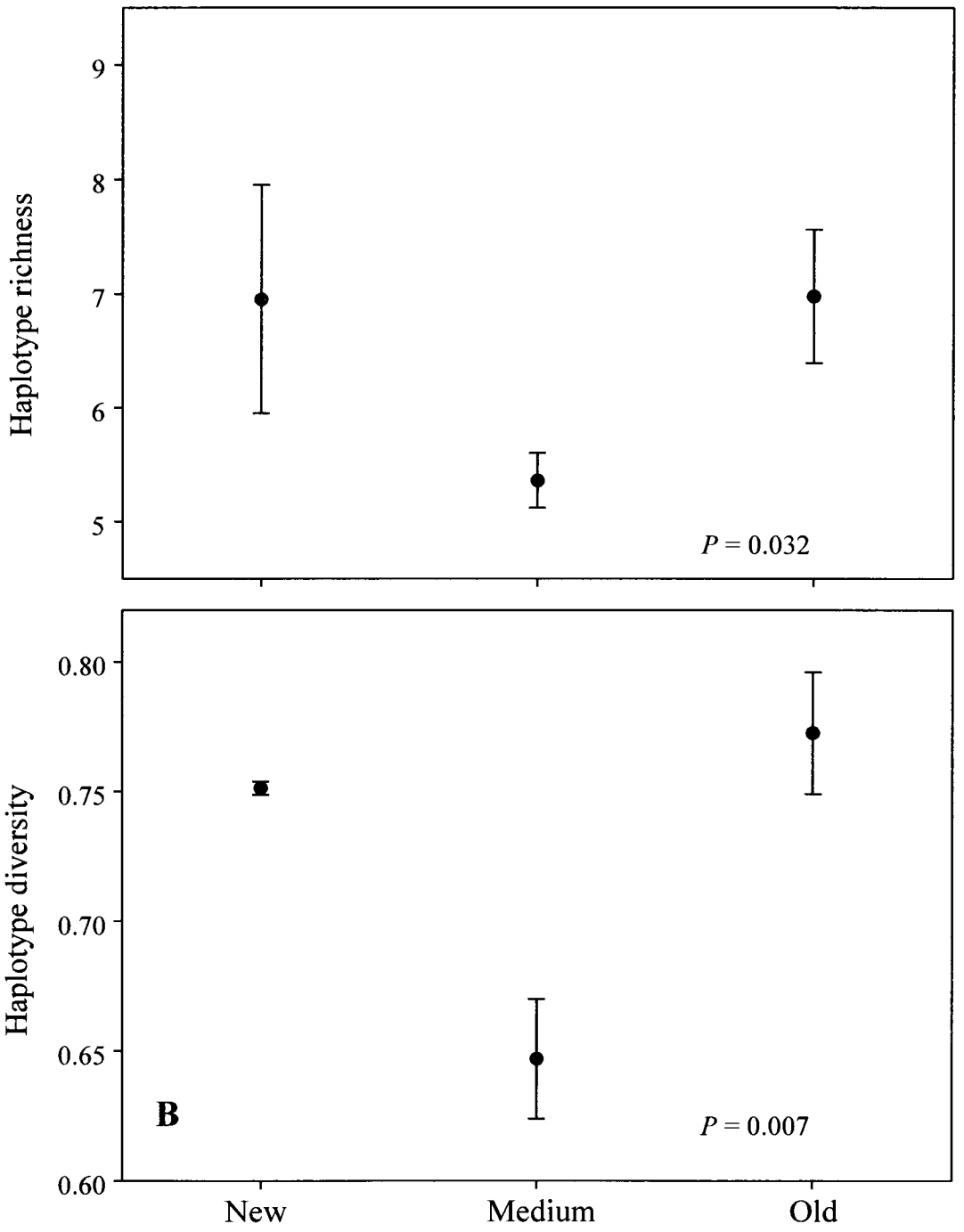


Figure 2.13. Measures of mtDNA variation for even-year pink salmon collections from three stream-age categories. Mean and SE for **(A)** number of composite haplotypes and **(B)** haplotype diversity. *P*-values from Kruskal-Wallis ANOVA on ranks.

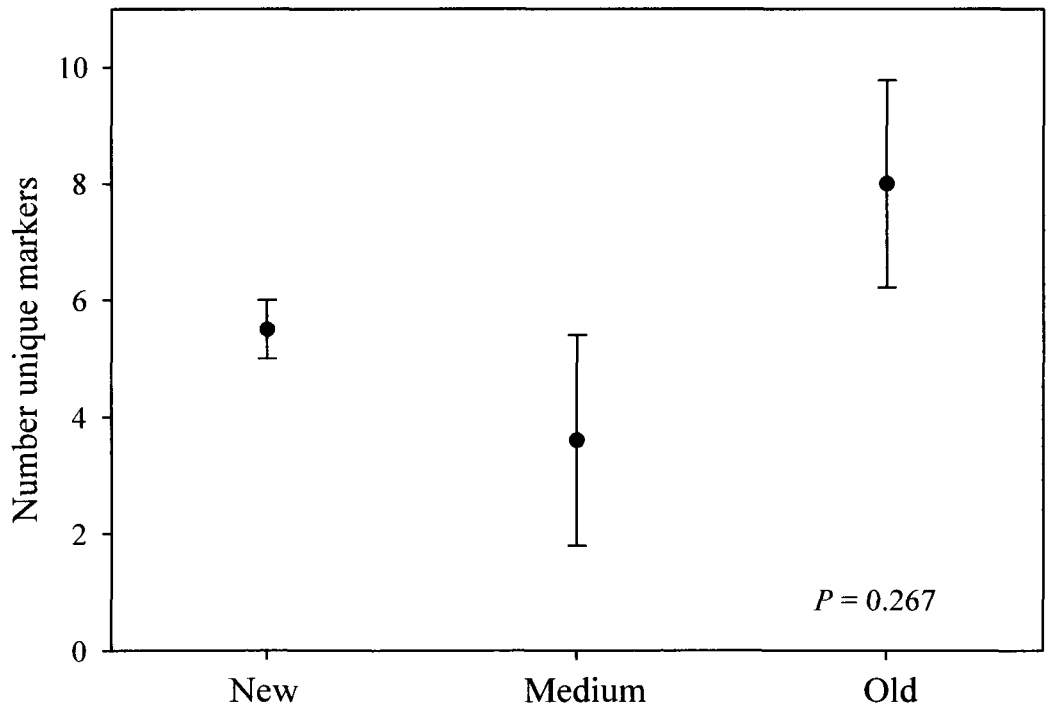


Figure 2.14. Number of private alleles and composite haplotypes in the microsatellite, variable allozyme, and mtDNA datasets for even-year pink salmon collections. Mean and SE from three stream-age categories; *P*-value from Kruskal-Wallis ANOVA on ranks.

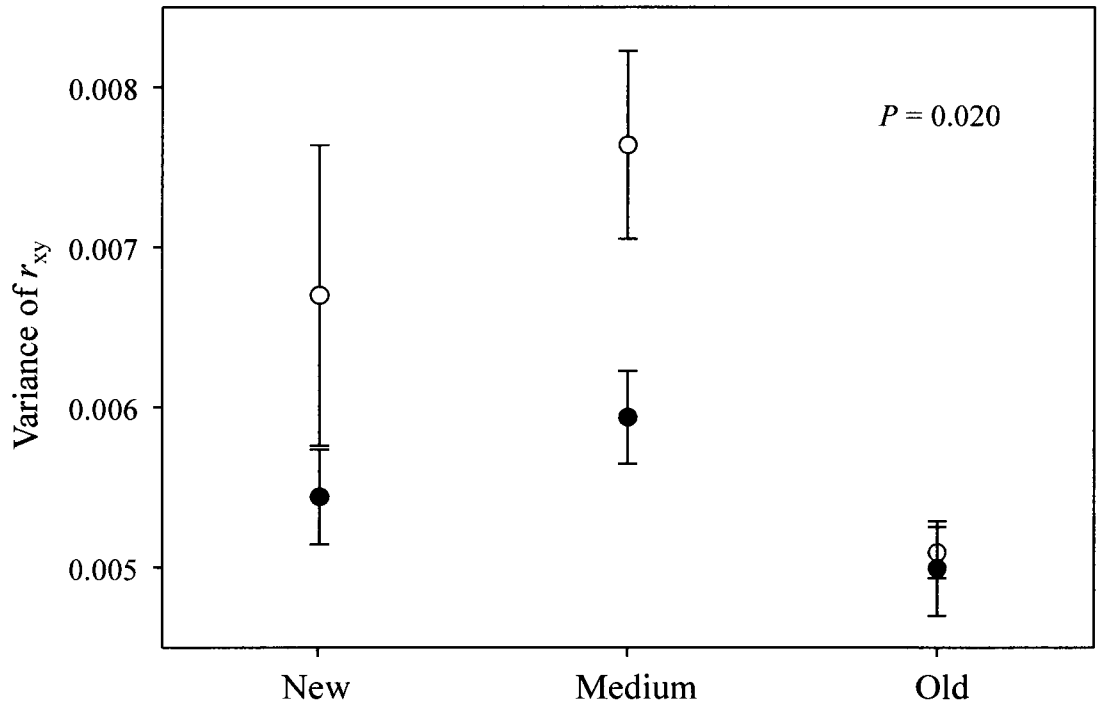


Figure 2.15. Lynch and Ritland (1999) relatedness estimates (r_{xy}) for even-year pink salmon collections from three stream-age categories. The mean observed variances (open circles) differed across stream-ages ($P = 0.009$); the permuted variances (solid circles) did not ($P = 0.193$). Observed variance was greater than the permuted variance ($P = 0.020$, ANOVA).

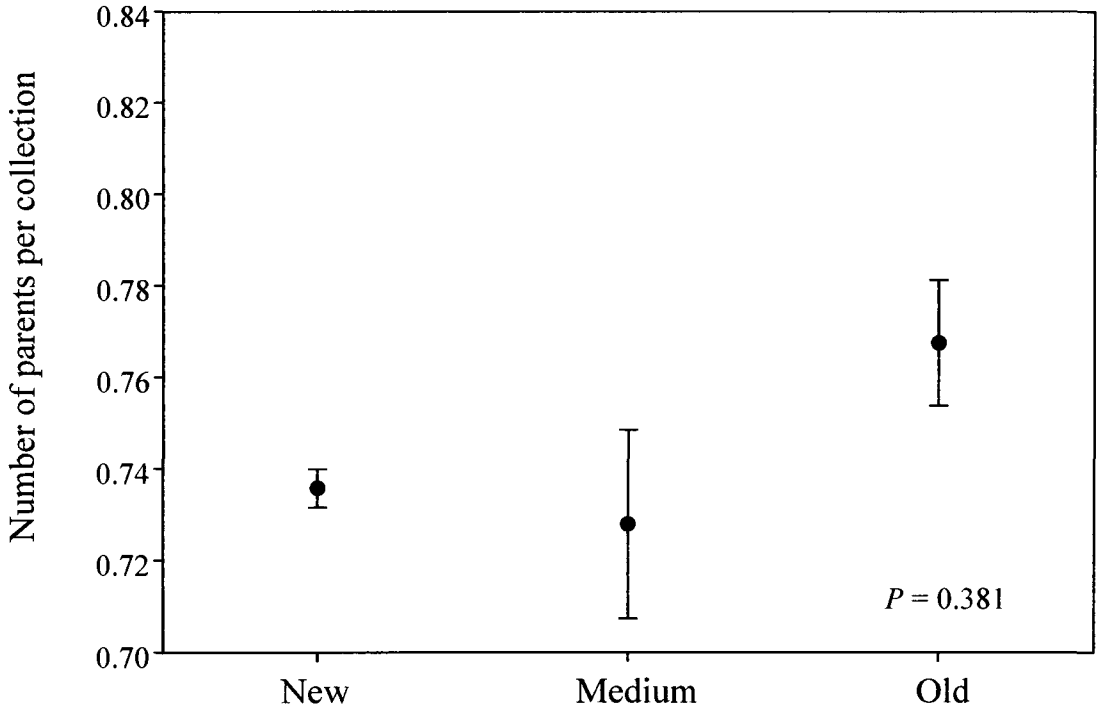


Figure 2.16. Estimated number of parents per collection for the even-year pink salmon from three stream-age categories (mean, SE). Standardized by collection sample size; from the genealogy that generated the highest correlation between the pedigree and molecular co-ancestries (Fernandez and Toro, 2006). *sAAT-4* not included for Tyndall 1990. *P*-value from Kruskal-Wallis one-way ANOVA on ranks.

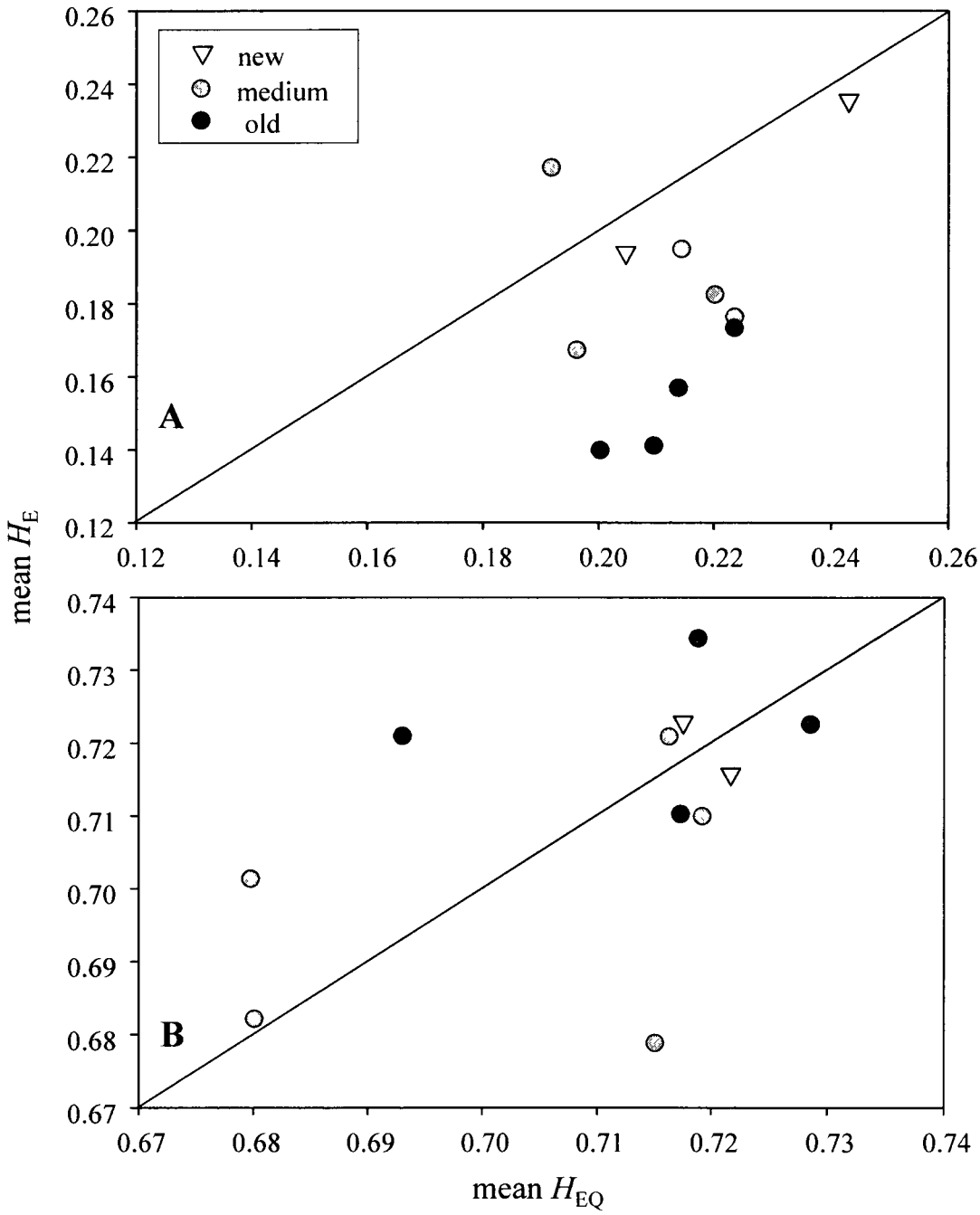


Figure 2.17. Equilibrium heterozygosity (H_{EQ}) and expected heterozygosity (H_E) for even-year pink salmon from three stream-age categories from the BOTTLENECK program. **(A)** 33 variable allozyme loci with IAM, and **(B)** 9 microsatellite loci with 95% one-step mutations with TPM. The line represents values where H_{EQ} equals H_E.

General Conclusion

The objective of this dissertation research was to use genetic analyses to deduce probable modes of pink salmon colonization in Glacier Bay, Alaska. The analyses were organized to address a series of questions (see General Introduction) for each broodline that should elucidate mechanisms for the questions that were asked.

Glacier Bay lies in the central region of the distribution of pink salmon in North America, where both even- and odd-year broodlines are present and serve as replicate experiments for the colonization process. The mechanisms of colonization for the two broodlines of pink salmon in Glacier Bay were similar in some respects, and different in others. Given the recent common ancestry of the two broodlines (Churikov and Gharrett 2002), it is not surprising that many aspects of colonization were shared. On the other hand, some differences were not unexpected because of stochastic changes in genetic (e.g., neutral markers; Hawkins et al. 2002), life history (e.g., run time; Gharrett et al. 2001), and biological (e.g., meristics; Beacham et al. 1988) characteristics that have developed as the two broodlines evolve separately.

Comparison of even- and odd-year pink salmon broodlines

The abundance of fish in the youngest populations differed between the broodlines. Our sampling occurred during years of large returns for both broodlines when few salmon returned in the youngest populations in Glacier Bay in the even years even though they were present in large numbers in the nearby, medium-aged populations. The higher abundance observed in the youngest populations in the odd years tracked the higher escapement of the odd-year broodline throughout the NSE Alaska region. (Der Hovanisian and Geiger 2005).

Genetic markers.—The same common allele at every allozyme locus was shared by both broodlines; however, there were a number of alleles present in only one or the other broodline. There was major overlap in the size range of microsatellite alleles, although most of the microsatellite loci were more allelic and the size range of alleles was typically larger in the even-year broodline (Tables 1.4 and 2.2). With the exception of

*One*μ*13* and *One111*, with the latter having a truncated allele-size distribution in the odd year, heterozygosity by locus was slightly lower in the even year. By locus and across all loci, the F_{ST} values across all populations were 2–3 times greater in the even-year broodline.

The mtDNA F_{ST} values were roughly four times that of the nuclear markers, which was expected, given their haploid matrilineal nature. Interpretation of genetic signals from the mtDNA did not differ substantially from the nuclear genetic markers, which suggests that the gene flow rates of male and female colonists were similar. A stronger genetic signal might have been expected with this marker, e.g., a more pronounced difference in haplotype richness than allele richness across the stream ages, but their haploid inheritance and smaller sample sizes than those used to obtain nuclear data (40 fish vs. 50–100 fish) reduced the power to detect population differences.

The greatest difference in the genetic markers between the broodlines occurred with the mtDNA haplotypes. The even- and odd-year broodlines shared only three mtDNA haplotypes (A*, M*, and AA*), with the remaining haplotypes restricted to one or the other broodline. As previously described (Churikov and Gharrett 2002), the A* haplotype was one of the most abundant haplotypes in the even-year broodline and at low frequency or absent in the odd-year broodline; the AA* and E* haplotypes were the most common in the odd-year broodline. Twenty composite haplotypes were observed in the even year, half of which were previously unreported; thirteen composite haplotypes were observed in the odd year, one-third of which were previously unreported. The greater number of haplotypes in the even year contrasted with observations in a previous survey of Southeast Alaska populations (Churikov and Gharrett 2002).

The differences in haplotypes and alleles between the two broodlines, as well as the higher F_{ST} values in the even year, may be due in part to differences in the ages of the broodlines, which relate to past historical influences beyond the short time frame of contemporary population structure. For example, even-year population structure based on a suite of allozyme loci is much stronger than in the odd year among areas of Asia (Hawkins et al. 2002). Variation in the mtDNA genome suggests that the even-year

broodline may be the ancestral lineage from which the odd-year lineage was derived (Churikov and Gharrett 2002).

Source location?—In general, the allele and haplotype frequencies of populations within Glacier Bay were most consistent with colonists that derived from nearby sources; however, there were some differences between the broodlines. Descriptive analyses (NJ trees and PCA), homogeneity testing, and estimates of genetic diversity (F_{ST}) support this conclusion. The oldest populations outside Glacier Bay within both broodlines clustered most closely. Of the medium-aged populations, the two oldest in lower Glacier Bay (N. Berg and N. Fingers) were most closely aligned with the oldest populations outside the Bay. The other medium-aged populations farther up the Bay had allele/haplotype frequencies either intermediate to the oldest populations in the lower bay and the youngest populations in the Bay (odd year), or somewhat distinct from all other populations (even year). The youngest populations in the upper Bay were most similar to the medium-aged populations within the Bay (odd year), or intermediate to the medium-aged and oldest populations (even year).

Two differences between the broodlines were the larger genetic distances overall within the even year and increased homogeneity of populations outside Glacier Bay within the odd year. In addition, allozyme data from a larger set of NSE Alaska populations indicated that the odd-year populations within Glacier Bay were heterogeneous, and the representative NSE Alaska populations were not; whereas in the even-year broodline, heterogeneity was observed in the populations within the Bay and in NSE Alaska. The range of population pairwise F_{ST} values of nuclear markers varied considerably depending on which populations were compared, but were on average about two times higher for the even-year data than the odd-year data. The increased divergence in the even year may be a result of more recent colonization of the even-year broodline in Southeast Alaska. However, a similar pattern of increased divergence in the even-year broodline is also present in Asian pink salmon populations (Noll et al. 2001; Hawkins et al. 2002), which hints at older inherent differences between the two broodlines (e.g., Churikov and Gharrett 2002).

For both brood years, the adult spawners returned during the "middle" run-time of Southeast Alaska (Royce 1962; Sheridan 1962; Mathisen 1994), which is prevalent in most populations nearby Glacier Bay (Spasski has an earlier run time). Although the run time in Glacier Bay best matches the middle run, there were minor differences between the two broodlines; peak abundance occurred later in the even years.

Colonization of new habitat by immigrants from nearby sources is concordant with the probing behavior of pink salmon, which is strongly associated with streams nearby the final spawning site as fish approach reproductive maturity (Maselko et al. 1999). It is also consistent with the preponderance of straying by post-spawners to nearby natal streams in Southeast Alaska (Thedinga et al. 2000).

Another factor that may explain the propensity of colonists to come from nearby populations is the less stable environment of watersheds in Glacier Bay, which may promote straying (Quinn and Tallman 1987). Pink salmon spawners subject to frequent, severe geologic events have weaker population structure, suggestive of higher levels of straying (Gharrett et al. 1988). The rapidly evolving physical and biological features of Glacier Bay have generated dynamic stream environments (Sidle and Milner 1989; Milner 1997; Engstrom et al. 2000; Milner et al. 2000). As an example, large changes in the hydrological characteristics of Wolf Pt. were observed during the study; the stream was notably warmer and less turbid after 1992 as the result of both substantial shrinkage of remnant ice in the lake above the barrier falls and the increased role of rain and snowmelt (Milner et al. 2007). Before successional change results in a developed riparian zone, stream flow is not well buffered and floods can scour streambeds. Perhaps the inherent instability of new stream habitat also promotes straying of fish within the Bay, especially among the young streams, and explains in part the rapid salmonid colonization in Glacier Bay (Milner and Bailey 1989).

High abundance of fish in nearby streams may also be an element of successful colonization in certain circumstances. Thrower (1988) hypothesized that the final spawning site is dependent on the suitability of spawning habitat and that stressful conditions may induce straying. One such condition could be overcrowding; pink salmon

populations have large variation in escapement with runs that can vary by an order of magnitude across short time periods (Der Hovanisian and Geiger 2005). Crowding during high abundance may explain the establishment of populations in the youngest systems during years of very high escapement in Southeast Alaska. The colonization of pink salmon in the youngest systems in Glacier Bay coincided with very high historical escapement in the 1990s that followed the increasingly larger escapements of pink salmon to Southeast Alaska in the early 1980s. Years of high escapement could increase the likelihood that pink salmon expand their geographic range, and once a suitable, open site is discovered, imprinting by a single generation fixes the stream for use by subsequent generations (Cury 1994).

Number of sources?—For both broodlines, limited linkage disequilibrium within populations and the marginally higher correlation coefficients in the youngest populations was consistent with mixing of only a small number of populations at the time of colonization. Additionally, as would be expected with few population sources, allelic diversity (allele and haplotype richness, haplotype diversity, and number of private alleles) was slightly lower within Glacier Bay. Heterozygosity did not provide clear signals and was not always consistent with allelic diversity either within or between the broodlines, which is not surprising given the lack of an effect on heterozygosity except from severe bottlenecks (Spencer et al. 2000). Higher relatedness in populations within Glacier Bay also supported a limited number of donor sources.

One of the most obvious differences between the two broodlines was the high allelic diversity of the youngest population, Wolf Pt., in the even year, but not in the odd year. In the even year, Wolf Pt. also had the highest linkage disequilibrium, which can be explained by a smaller number of colonists (the large estimate of N_e was due to large sampling error associated with the small sample size; harmonic mean $n = 33.5$; Waples 2006) or a contribution of fish from several donor sources. Low numbers of fish were observed in 1994, the first even year that pink salmon were observed in Wolf Pt. Considering the small number of fish observed in 1994, the frequencies of the nuclear markers indicated that most of the colonists were likely from nearby Nunatak, the other

young even-year population, but frequencies of the mtDNA haplotypes, as well as the higher allelic diversity of fish in Wolf Pt., indicated that contributions also came from sources farther away, either from the lower Bay or outside Glacier Bay.

Number of colonists?—There was no evidence in either broodline that the number of pink salmon involved in the initial colonization process was either large or small. Rather, the genetic data suggested that the number of colonists lies somewhere between the extreme possibilities. Together, the small number of locus-pairs out of linkage equilibrium and the only slightly smaller correlation coefficients within Glacier Bay populations precluded colonization by a very small number of fish. On the other hand, allelic diversity tended to increase with population age and was higher in the older populations outside Glacier Bay, which eliminated the possibility that a large number of fish was involved in colonization of the youngest streams.

As an index of genetic diversity, allelic diversity (particularly of microsatellite loci) was more sensitive than heterozygosity to the limited founder effects present in Glacier Bay, a pattern that has been previously reported in theoretical (Nei et al. 1975) and empirical studies (Leberg 1992; Spencer et al. 2000; Keller et al. 2001). Heterozygosity of allozyme loci has been shown to poorly reflect loss of genetic diversity in all but the most severe population bottlenecks (Leberg 1992). Rather than decrease as expected in a bottlenecked population, heterozygosity can become higher than before the bottleneck due to stochastic changes in allele frequencies, e.g., infrequent alleles in the pre-bottleneck population that by chance become more abundant in the post-bottleneck population. Such changes in allele frequencies may explain the higher heterozygosity of allozyme loci in the youngest odd-year populations and the medium-aged even-year populations.

The effective population size (N_e) estimates were generally lower in the youngest populations in both broodlines, but in every population N_e exceeded 100 fish and the upper 95% confidence intervals included infinity, which ruled out a very small number of colonists. Likewise, N_e of the younger populations would have been more like the estimates of the older populations if a large number of fish had colonized the streams.

No bottleneck signal was evident in any population of either broodline from the comparison of heterozygosity with that expected under mutation-drift equilibrium. In fact, the heterozygosity deficiency of allozyme loci, most notably in the odd year, was indicative of population mixing or a recent expansion of population size rather than a bottleneck (Cornuet and Luikart 1996; Luikart and Cornuet 1998), although this pattern has been reported for many allozyme datasets (Chakraborty et al. 1980). Depending on the mutation model invoked, the microsatellite loci tended more toward heterozygosity excess, the expected outcome of a bottleneck, but none of the tests were significant at either two-phase model (95% and 70% one-step mutations) and there was no mode shift in allele frequencies.

Gene flow frequency?—There was a trend of increased allelic diversity with stream age in both broodlines, suggestive of gene flow after initial colonization; however, the relationship between stream age and genetic diversity was not always linear. For example, in the even year the lowest diversity occurred in the medium-aged populations, which implies limited gene flow in at least some populations once they become established. As explained above, heterozygosity as a measure of genetic diversity offered contradictory information within and between broodlines, and as a consequence was a less useful measure for addressing the questions posed except to exclude the possibility of severe founder effects.

Based on several measures (Lynch and Ritland 1999; Fernandez and Toro 2006; Kalinowski et al. 2006), the relatedness of individuals was very low in all populations in both broodlines, but it was slightly greater within Glacier Bay populations than in the older populations outside the Bay. The greater relatedness within Glacier Bay also argues against extensive gene flow. One exception to the higher relatedness within Glacier Bay was the comparatively low relatedness in Wolf Pt. in the even year that was similar to the oldest populations outside Glacier Bay. This may have had more to do with the possibility that initial colonizers were sampled than it had to do with higher subsequent gene flow.

Even though odors of conspecifics and kin can be detected by salmonids (e.g., Quinn and Tolson 1986; Griffiths and Armstrong 2000) and could potentially attract strays, colonization of empty habitat may occur at a higher immigration rate than among populated streams, and migration rates may change over time (Waples 1991). In one field study, the rate of straying, as determined by fin marking, did not reflect the true rate of gene flow among populations, and suggested that the reproductive success rate of strays is reduced in systems that are already occupied by locally adapted populations (Tallman and Healey 1994). The relatively large effective population size (N_e) estimates for the Glacier Bay populations imply that few immigrants per generation are required to balance random drift of allele frequencies and arrest development of genetic divergence among populations (Gharrett and Zhivotovsky 2003). Thus, the presence of heterogeneity among some populations in Glacier Bay suggests restricted gene flow in both broodlines subsequent to initial colonization.

One explanation for the rapid increase in abundance in Wolf Pt. (Milner et al. 2007) and Nunatak creeks in years after samples were collected could be high gene flow, but given the heterogeneity of populations within Glacier Bay, a more reasonable explanation may be the presumed high marine survival during the 1990s that led to record high escapements in Southeast Alaska during this time period. Samples from additional generations might have provided insight. For example, if the increase in abundance in the new populations was primarily from high survival of offspring, the relatedness estimates within populations would be similar or even higher than prior to population expansion, whereas relatedness would decrease if immigration from other sources occurred. In addition, high reproductive success of colonizing females due to less competition for spawning habitat (e.g., Anderson et al. 2010) may more than compensate for the potentially lower quality characteristics of recently deglaciated habitat (colder, more turbid, and prone to flooding) and fewer potential mates. Although high marine survival may explain much of the increase in abundance in the youngest populations after sampling was completed for this study, some gene flow cannot be ruled out. The similarity of allele frequencies among the geographically proximate older populations

outside Glacier Bay also suggests that some exchange of immigrants occurs among proximate populations over time.

Summary

Genetic signals generated in this study to examine the colonization process of pink salmon in Glacier Bay were moderate. Genetic diversity and population genetic structure indicated that the colonization mechanisms lay between the extreme contrasting possibilities proposed in Table 1.1. Some genetic signals differed between the two broodlines, but many parallels were evident. Allelic diversity provided a stronger, more consistent signal than heterozygosity; the highly allelic microsatellite loci were the most sensitive marker to the mild founder effects.

Populations in Glacier Bay were more similar to nearby populations than far away sources based on allele/haplotype frequencies and estimates of genetic diversity (F_{ST}), although in the even-year there appeared to be some contribution to the youngest populations from older populations outside or in lower Glacier Bay. The moderate genetic signals observed in Glacier Bay match other studies of pink salmon, which show low-to-moderate population genetic structure at various spatial and temporal scales (e.g., Beacham et al. 1988; Gharrett et al. 1988, 2001; Seeb et al. 1999; Noll et al. 2001; Hawkins et al. 2002). The limited linkage disequilibrium, higher relatedness, and lower allelic diversity within Glacier Bay populations indicated that few genetically divergent donor sources contributed colonists. The genetic data also indicated that a moderate number of fish were involved in initial colonization. The genetic diversity measures of the younger populations within the Bay were generally lower than those of the older populations outside the Bay, which eliminates the possibility that large numbers of fish colonized the streams. However, small numbers of fish can also be ruled out because (1) allelic diversity in the youngest populations was not severely reduced; (2) the relatedness of fish within populations within Glacier Bay, although higher than in older populations outside the Bay, was still low; (3) the estimates of N_e for all populations were relatively large (>100 fish); and (4) heterozygosity did not exceed that expected under mutation-

drift equilibrium. The increase in allelic diversity and decline in relatedness with population age, as well as homogeneity among the older populations outside Glacier Bay suggested that some gene flow occurs after initial colonization. But, heterogeneity within Glacier Bay, particularly among the medium-aged populations, also signified that gene flow must be limited among some populations.

Colonization of streams in Glacier Bay appeared to be episodic; and together with limited gene flow among some populations, resulted in significant heterogeneity in both broodlines, particularly among the medium-aged populations. The timing of colonization in the most recently deglaciated streams and the ensuing increase in population size was associated with a decade of record high pink salmon abundance in Southeast Alaska (Heinl and Geiger 2005), likely due to high survival in the marine environment. During the 1990s, populations in Southeast Alaska were above replacement levels, which resulted in population growth that we speculate improved the odds of successful establishment of populations in new habitat. Pink salmon survival from fry to adult is highly variable (Taylor 2008) and conceivably the large increase in population size in the new Glacier Bay populations subsequent to this study resulted from high survival rather than extensive gene flow.

Our focus was on the colonization mechanisms that explain the patterns in one small-scale geographic locale over a short time period (<200 years), which does not preclude other colonization mechanisms at other times and under different environmental and ecological conditions. Knowledge of pink salmon abundance in Southeast Alaska about 150 years ago, during the time that streams in lower Glacier Bay were colonized, would have provided further insight into colonization, but those data are not available. However, large variability in salmon abundance in Alaska prior to the development of commercial fisheries (Finney et al. 2000, 2002) suggests that periods of high pink salmon production as documented in the 1990s likely occurred in NSE Alaska in the more distant past, and we can speculate that episodic colonization tied to periods of high abundance may be typical for this species. Pink salmon have been observed to expand into new territory when large spawning runs lead to overcrowding, and to use habitat that was not

formerly occupied during poor environmental conditions (Vernon 1962; Withler 1982; Blair and Quinn 1991). In combination with probing and straying behavior, periods of high abundance may increase the probability that pink salmon find new, nearby habitat, e.g., as we propose occurred in lower Glacier Bay shortly after deglaciation. Even if only low numbers of spawners stray from natal sites and encounter suitable spawning habitat, imprinting permits the establishment of a population in new environments in a single generation (Cury 1994). Homing is strong even in the progeny of colonists (McDowall 2001), and during periods of good environmental conditions as occurred in the 1990s in NSE Alaska, homing and high survival can facilitate rapid population growth.

References

Anderson, J. H., P. L. Faulds, W. I. Atlas, G. R. Pess, and T. P. Quinn. 2010. Selection on breeding date and body size in colonizing coho salmon, *Onchorhynchus kisutch.* Molecular Ecology 19:2562-2573.

Beacham, T. D., R. E. Wither, C. B. Murray, and L. W. Barner. 1988. Variation in body size, morphology, egg size, and biochemical genetics of pink salmon in British Columbia. Transactions of the American Fisheries Society 117:109-126.

Blair, G. R., and T. P. Quinn. 1991. Homing and spawning site selection by sockeye salmon (*Oncorhynchus nerka*) in Iliamna Lake, Alaska. Canadian Journal of Zoology 69:176-181.

Chakraborty, R., P. A. Fuerst, and M. Nei. 1980. Statistical studies on protein polymorphism in natural populations. III. distribution of allele frequencies and the number of alleles per locus. Genetics 94:1039-1063.

Churikov, D., and A. J. Gharrett. 2002. Comparative phylogeography of the two pink salmon broodlines: an analysis based on a mitochondrial DNA genealogy. Molecular Ecology 11:1077-1101.

Cornuet, J. M., and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001-2014.

Cury, P. 1994. Obstinate nature: an ecology of individuals. Thoughts on reproductive behavior and biodiversity. Canadian Journal of Fisheries and Aquatic Sciences 51: 1664-1673.

Der Hovanisian, J. A., and H. J. Geiger. 2005. Stock status and escapement goals for salmon in southeast Alaska 2005. Alaska Department of Fish and Game, Special Publication No. 05-22, Douglas.

Engstrom, D. R., S. C. Fritz, J. E. Almendinger, and S. Juggins. 2000. Chemical and biological trends during lake evolution in recently deglaciated terrain. Nature 408: 161-166.

Fernandez, J., and A. Toro. 2006. A new method to estimate relatedness from molecular markers. Molecular Ecology 15:1657-1667.

Finney, B. P., I. Gregory-Eaves, M. S. V. Douglas, and J. P. Smol. 2002. Fisheries productivity in the northeastern Pacific Ocean over the past 2,200 years. Nature 416: 729-733.

Finney, B. P., I. Gregory-Eaves, J. Sweetman, M. S. V. Douglas, and J. Smol. 2000. Impacts of climatic change and fishing on Pacific salmon abundance over the past 300 years. Science 290:795-799.

Gharrett, A. J., S. Lane, A. J. McGregor, and S. G. Taylor. 2001. Use of a genetic marker to examine genetic interaction among subpopulations of pink salmon (*Oncorhynchus gorbuscha*). Genetica 111:259-267.

Gharrett, A. J., C. Smoot, A. J. McGregor, and P. B. Holmes. 1988. Genetic relationships of even-year northwestern Alaskan pink salmon. Transactions of the American Fisheries Society 117:536-545.

Gharrett, A. J., and L. A. Zhivotovsky. 2003. Migration. Pages 141-174 *in* E. M. Hallerman, editor. Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Bethesda, Maryland.

Griffiths, S. W., and J. D. Armstrong. 2000. Differential responses of kin and nonkin salmon to patterns of water flow: does recirculation influence aggression? Animal Behavior 59:1019-1023.

Hawkins, S. L., N. V. Varnavskaya, E. A. Matzak, V. V. Efremov, C. M. Guthrie III, R. L. Wilmot, H. Mayama, F. Yamazaki, and A. J. Gharrett. 2002. Population structure of odd-broodline Asian pink salmon and its contrast to the even-broodline structure. Journal of Fish Biology 60:370-388.

Heinl, S. C., and H. J. Geiger. 2005. Chapter 4: Pink salmon stock status and escapement goals in Southeast Alaska and Yakutat. Pages 153-189 *in* J. A. Der Hovanisian and H. J. Geiger, editors. Stock status and escapement goals for salmon in Southeast Alaska 2005. Alaska Department of Fish and Game, Special Publication No. 05-22, Douglas.

Kalinowski, S. T., A. P. Wagner, and M. L. Taper. 2006. ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. Molecular Ecology Notes 6:576-579.

Keller, L. F., K. J. Jeffery, P. Arcese, M. A. Beaumont, W. M. Hochachka, J. N. M. Smith, and M. W. Bruford. 2001. Immigration and the ephemerality of a natural population bottleneck: evidence from molecular markers. Proceedings of the Royal Society of London Series B—Biological Sciences 268:1387-1394.

Leberg, P. L. 1992. Effects of population bottlenecks on genetic diversity as measured by allozyme electrophoresis. Evolution 46:477-494.

Luikart, G., and J. M. Cornuet. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology 12:288-237.

Lynch, M., and K. Ritland. 1999. Estimation of pairwise relatedness with molecular markers. Genetics 152:1753-1766.

Maselko, J. M., A. C. Wertheimer, and J. F. Thedinga. 1999. Estimating probing behavior of pink salmon. Pages 40-48 *in* S. Hawkins, rappoteur. Proceedings of the Northeast Pacific Pink & Chum Salmon Workshop, Juneau, Alaska.

Mathisen, O. A. 1994. Spawning characteristics of the pink salmon (*Oncorhynchus gorbuscha*) in the eastern North Pacific Ocean. Aquaculture and Fisheries Management 25(Supplement 2):147-156.

McDowall, R. M. 2001. Anadromy and homing: two life-history traits with adaptive synergies in salmonid fishes? Fish and Fisheries 2:78-85.

Milner, A. M. 1997. Glacial recession and freshwater ecosystems in coastal Alaska. Pages 303-331 *in* A. M. Milner, and M. W. Oswood, editors. Freshwaters of Alaska: ecological syntheses. Ecological Studies, Volume 119, Springer, New York.

Milner, A. M., and R. G. Bailey. 1989. Salmonid colonization of new streams in Glacier Bay National Park, Alaska. Aquaculture and Fisheries Management 20:179-192.

Milner, A. M., E. E. Knudsen, C. Soiseth, A. L. Robertson, D. Schell, I. T. Phillips, and K. Magnusson. 2000. Colonization and development of stream communities across a 200-year gradient in Glacier Bay National Park, Alaska, U.S.A. Canadian Journal of Fisheries and Aquatic Sciences 57:2319-2335.

Milner, A. M., K. Monaghan, E. A. Flory, A. J. Veal, and A. Robertson. 2007. Ecological development of the Wolf Point Creek watershed; a 25-year colonization record from 1977 to 2001. Pages 3-7 *in* J. F. Piatt, and S. M. Gende, editors. Proceedings of the fourth Glacier Bay science symposium. U.S. Geological Survey Scientific Investigations Report 2007-5047, Reston, Virginia.

Nei, M., T. Maruyama, and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. Evolution 29:1-10.

Noll, C., N. V. Varnavskaya, E. A. Matzak, S. L. Hawkins, V. V. Midanaya, O. N. Katugin, C. Russell, N. M. Kinas, C. M. Guthrie III, H. Mayama, F. Yamazaki, B. P. Finney, and A. J. Gharrett. 2001. Analysis of contemporary genetic structure of even-broodyear populations of Asian and western Alaskan pink salmon, *Oncorhynchus gorbuscha*. Fishery Bulletin 99:123-138.

Quinn, T. P., and R. F. Tallman. 1987. Seasonal environmental predictability and homing in riverine fishes. Environmental Biology of Fishes 18:155-159.

Quinn, T. P., and G. M. Tolson. 1986. Evidence of chemically mediated population recognition in coho salmon (*Oncorhynchus kisutch*). Canadian Journal of Zoology 64: 84-87.

Royce, W. F. 1962. Pink salmon fluctuations in Alaska. Pages 15-33 *in* N. J. Wilimovsky, editor. Symposium on pink salmon. H. R. MacMillan Lectures in Fisheries, Institute of Fisheries, University of British Columbia, Vancouver.

Seeb, J. E., C. Habicht, W. D. Templin, L. W. Seeb, J. B. Shaklee, and F. M. Utter. 1999. Allozyme and mitochondrial DNA variation describe ecologically important genetic structure of even-year pink salmon inhabiting Prince William Sound, Alaska. Ecology of Freshwater Fish 8:122-140.

Sheridan, W. L. 1962. Relation of stream temperatures to timing of pink salmon escapements in Southeast Alaska. Pages 87-102 *in* N. J. Wilimovsky, editor. Symposium on pink salmon. H. R. MacMillan Lectures in Fisheries, Institute of Fisheries, University of British Columbia, Vancouver.

Sidle, R. C., and A. M. Milner. 1989. Stream development in Glacier Bay National Park, Alaska, U.S.A. Arctic and Alpine Research 21:350-363.

Spencer, C. C., J. E. Neigel, and P. L. Leberg. 2000. Experimental evaluation of the usefulness of microsatellite DNA for detecting demographic bottlenecks. Molecular Ecology 9:1517-1528.

Tallman, R. F., and M. C. Healey. 1994. Homing, straying, and gene flow among seasonally separated populations of chum salmon (*Oncorhynchus keta*). Canadian Journal of Fisheries and Aquatic Sciences 51:577-588.

Taylor, S. G. 2008. Auke Creek weir 2007 annual report, operations, fish counts, and historical summaries. Unpublished Report, National Marine Fisheries Service, Auke Bay Fisheries Laboratory, Juneau, Alaska.

Thedinga, J. F., A. C. Wertheimer, R. A. Heintz, J. M. Maselko, and S. D. Rice. 2000. Effects of stock, coded-wire tagging, and transplant on straying of pink salmon (*Oncorhynchus gorbuscha*) in southeastern Alaska. Canadian Journal of Fisheries and Aquatic Sciences 57:2076-2085.

Thrower, F. P. 1988. Migratory behavior of maturing pink salmon in Gastineau Channel, Southeast Alaska. M.S. Thesis, University of Alaska Fairbanks.

Vernon, E. H. 1962. Pink salmon populations of the Fraser River system. Pages 53-58 *in* N. J. Wilimosvky, editor. Symposium on pink salmon. H. R. MacMillan Lectures in Fisheries, Institute of Fisheries, University of British Columbia, Vancouver.

Waples, R. S. 1991. Pacific salmon, *Oncorhynchus* spp., and the definition of "species" under the Endangered Species Act. Marine Fisheries Review 53:11-22.

Waples, R. S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conservation Genetics 7:167-184.

Withler, F. C. 1982. Transplanting Pacific salmon. Canadian Technical Report of Fisheries and Aquatic Sciences 1079, Nanaimo, British Columbia.