INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Genetic Interpretation of Microsatellite Polymorphism in Pacific Salmon:
Case studies in population genetics and kinship analysis

by

Jeffrey B. Olsen

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

1999

Approved by _____
Chairperson of Supervisory Committee

_____

_____

_____

_____

_____

Program Authorized
to Offer Degree FISHERIES

Date 3/12/99

Doctoral Dissertation

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to University Microfilms, 1490 Eisenhower Place, P.O. Box 975, Ann Arbor, MI 48106, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Jeffrey B. Olsen

Date 3/12/99

University of Washington

Abstract

Genetic Interpretation of Microsatellite Polymorphism in Pacific Salmon:
Case studies in population genetics and kinship analysis

by Jeffrey B. Olsen

Chairperson of Supervisory Committee
Professor Paul Bentzen
School of Fisheries

I conducted five independent studies focusing on development and application of microsatellite DNA for intra-specific genetic analyses of Pacific salmon (*Oncorhynchus* spp.). I developed techniques for multiplex analysis of microsatellite loci for chinook, coho, sockeye and pink salmon using a semi-automated 4-color fluorescent detection system. I cloned seven novel microsatellite loci from pink salmon that also amplify in other species of Pacific salmon. I examined genetic variation at five microsatellite loci in 12 odd-year populations and one even-year population of pink salmon from six geographic regions of North America. The results suggest that the statistic θ is a better estimator of intra-lineage (odd-year x odd-year) population structure, whereas ρ_{ST} is best suited for estimating inter-lineage (odd-year x even-year) population structure. The difference in performance of θ and ρ_{ST} for estimating intra-lineage and inter-lineage population structure suggests high migration rates and short divergence times are dominant influences on genetic population structure in odd-year pink salmon. I compared microsatellites and allozymes for estimating population origin of individual pink salmon in a supportive breeding program. Both marker types reveal similarly low estimates of relative genetic differentiation ($\theta = 0.02$) between two fall-run and summer-run populations but microsatellites provided a more accurate estimate of population origin. The difference in assignment success was best predicted by the statistic δ, which estimates cumulative allele frequency differences among populations. Finally, I used

computer simulation and actual microsatellite data from a known chinook salmon pedigree to determine how number of loci, heterozygosity, parent population size and full sibs in the parent population affect pedigree reconstruction. The results indicate that parentage assignment success is highly sensitive to mean locus heterozygosity while moderate increases in parent population size have a minor affect on assignment success. The presence of closely related individuals in the parent population confounded pedigree reconstruction but the use of parent pair-offspring likelihood analysis improved resolution.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Acknowledgments

I am indebted to the members of my graduate committee who gave freely of their time and knowledge and who encouraged and supported my research. Most importantly I thank my graduate advisor, Dr. Paul Bentzen, for his scholarly guidance and willingness to support a student with no background in molecular genetics. I thank Dr. Jim Seeb who encouraged my research, assisted with financial support, and provided insightful advice on successfully navigating graduate school. I thank Dr. Fred Utter for his many helpful comments on earlier drafts of my dissertation. I thank Drs. Chris Foote, Scott Edwards, and Kenneth Sirotnek who helped focus and motivate my research by providing helpful comments and discussion during committee meetings.

I would also like to thank the following colleagues and friends whose assistance and encouragement were an integral part of my success as a graduate student: Dr. Brian Allee, Dr. Craig Busack, Penny Crane, Chris Habicht, Pam Jensen, Rand Little, Dr. Pat O'Reilly, Kristina Ramstad, Rolf Ream, Dr. Lisa Seeb, Dr. James Shaklee, Adrian Spidle, Bruce Suzumoto, Dan Warren, Dr. John Wenburg, Dr. Carol Woody, Sewall Young.

Dedication

I dedicate this dissertation to dad and mom. Your love for one another and your family is an inspiration and the reason why I completed this endeavor.

INTRODUCTION

Genetic markers have been used extensively for management and conservation of Pacific salmon (*Oncorhynchus* spp.) (Allendorf et al. 1987; Carvalho and Hauser 1994; Utter 1994). Protein electrophoresis is the primary method used to detect genetic population structure among discrete spawning groups and estimate stock composition in mixed stocked fisheries (Utter et al. 1987; Seeb et al. 1997). However, three factors have contributed to an interest in applying new DNA based genetic markers. First, in some cases protein coding loci lack sufficient polymorphism to reveal fine scale genetic population structure (Gyllensten and Wilson 1987; Forbes et al. 1994; Tessier et al. 1995). Second, protein electrophoresis commonly requires lethal sampling as well as immediate cold storage (-40 to -80°C) of tissue samples which limits feasibility in threatened or endangered species and imposes logistical difficulties in sampling collection and storage, respectively. Third, the combination of non-lethal sampling and extensive polymorphism detectable with many DNA markers facilitates novel applications such as kinship analysis, the use of DNA profiles as genetic "tags" and marker-assisted selective breeding in aquaculture (Bentzen et al. 1991; Wright and Bentzen 1994; Bentzen et al. 1994; O'Reilly and Wright 1995). In this dissertation I explore the use of a recently revealed class of nuclear DNA marker, microsatellites, for inferring broad- and fine-scale genetic population structure and close kinship in Pacific salmon.

Microsatellites

Microsatellites are highly polymorphic and abundant in eukaryotic genomes surveyed to date (Tautz 1989). They consist of 1-5 base pair (bp) repeating sequences that form arrays <300 bp in length and exhibit high levels of co-dominant allelic variation in repeat number (Wright 1992; O'Reilly and Wright 1995). Polymorphism exhibited by specific microsatellites is readily detected by amplification of the microsatellites through the use of oligonucleotide primers specific to the non-repetitive regions that flank the

repeat array, in combination with the polymerase chain reaction (PCR). Allelic variation is scored by gel electrophoresis of the PCR products, most commonly on denaturing acrylamide gels.

In vitro studies suggest microsatellites mutate via a process of slipped-strand mispairing during DNA replication, resulting in length changes of one or more repeat units (Levinson and Gutman 1987). This process is probably best described by a modification of the single step mutation model called the two phase mutation model (Di Rienzo et al. 1994). The mutation rate can be exceptionally high (e.g. 10^{-3} per generation; Weber and Wong 1993; Ellegren 1995). Nevertheless, the amount of polymorphism can vary widely among loci because of structural differences in the repeat segments that affect mutational stability (Jarne and Lagoda 1996; Petes et al. 1997).

Because polymorphism varies widely among loci, microsatellites have potential for elucidating all levels of intra-specific genetic relationships, and for examining evolutionary forces influencing those relationships (O'Reilly and Wright 1995; Jarne and Lagoda 1996). As a result, the use of microsatellites for studying close kinship and genetic population structure in non-humans is increasing (Goldstein and Pollock 1997). To make best use of microsatellites in different taxa, however, there is important technical development and empirical evaluation that must be conducted (Forbes et al. 1995; Scribner et al. 1996). I address some aspects of these two areas for Pacific Salmon in this dissertation; the results should be helpful for fish geneticists and ecologists planning genetic studies.

Technical development

The first two chapters address technical issues and focus on genotyping efficiency and microsatellite development. Collecting microsatellite data is often slow and costly, but methods exist for improving efficiency (e.g. Schwengel et al. 1994). In Chapter one I test co-amplification of microsatellites and develop a semi-automated, 6-locus, multiplexing system for chinook, coho, and sockeye salmon using 4-color fluorescent detection of microsatellites. By multiplexing 6 loci it is possible to process 432

genotypes per day using a semi-automated system such as the Perkin-Elmer Applied Biosystems Inc. (ABI) 373A sequencer/genescanner.

Microsatellite studies are often hindered by a shortage of loci. Fortunately, primer pairs developed in one species can often be used to amplify homologous loci in related species (e.g. FitzSimmons et al. 1995). In Chapter one I describe a screening of 35 microsatellites on five species of Pacific salmon and Atlantic salmon (*Salmo salar*) to determine the extent of cross-species amplification. These microsatellites are derived from six salmonid species. The number of loci that amplify varies by species from 11 (chum salmon, *O. Keta*) to 22 (chinook salmon, *O. tshawytscha*). This success at cross-species amplification indicates a shortage of loci in one species of salmonid need not be a hindrance to conducting a microsatellite study.

In chapter two I describe seven novel microsatellite loci discovered in pink salmon (*O. gorbuscha*). Polymorphism at each locus is shown for one pink salmon population. The number alleles per locus range from 2 (*Ogo*6) to 24 (*Ogo*1c) and average 13. The expected heterozygosity ranges from 0.14 (*Ogo*6) to 0.95 (*Ogo*1c) and averages 0.72. At least some of the microsatellites amplify in the 13 salmonid species tested. These new microsatellites should prove useful for a number of conservation genetic applications in pink salmon and, to a lesser degree, other salmonid species.

Empirical evaluation

In Chapters three through five I conduct case studies to evaluate microsatellites for a range of intra-specific genetic analyses – from a broad-scale study of population genetics in pink salmon to pedigree reconstruction in chinook salmon. I address specific issues regarding study design and data analysis including the following: an empirical evaluation of two statistics describing microsatellite variation between populations; an empirical evaluation of microsatellites and allozymes for determining population identity of individuals; an empirical evaluation of microsatellites for reconstructing pedigrees when the parent population consists of many full sibs.

In Chapter three I examine genetic variation at five microsatellite loci in 12 odd-year populations and one even-year population of pink salmon from six geographic

regions of North America. I compare estimates of a standard index of population structure (θ) based on variance in allele frequency with a new index suggested for microsatellites (ρ_{ST}) based on variance in allele size. The later index is intended to reflect the unique mutational properties of microsatellites described above (Slatkin 1995). The results suggest θ is a better estimator of intralineage (odd-year x odd-year) population structure, whereas ρ_{ST} is best suited for estimating interlineage (odd-year x even-year) population structure. The difference in performance of θ and ρ_{ST} for estimating intralineage and interlineage population structure suggests high migration rates and recent divergence times are dominant influences on genetic population structure in odd-year pink salmon. I also show statistical support for genetic isolation by distance and geographically correlated allele frequency clines, suggesting broad-scale gene flow may be described by a linear stepping stone model. An analysis of molecular variation shows weak but significant regional structuring under two different population grouping schemes. These results generally agree with studies of allozyme variation (e.g. Varnavskaya and Beacham 1992; Shaklee and Varnavskaya 1994), and suggest broad-scale population aggregations of odd-year pink salmon are temporally stable but that differentiation is weak, presumably due to migration.

In Chapter four I evaluate two classes of genetic markers, allozymes and microsatellites, for estimating population identity of pink salmon in a supportive breeding program on the Dungeness River in Washington State. Fall-run pink salmon of the Dungeness River are the target of restoration but overlap in timing with an earlier and more abundant summer-run. Both marker types reveal similarly low estimates of relative genetic differentiation ($\theta = 0.02$) suggesting little variation in allele frequency among populations. However, microsatellites provide a more accurate estimate of population identity. Applying a log-likelihood ratio criterion LLR>1.3, 74.8% of individuals are correctly assigned to population using microsatellites versus 3.1% of individuals using allozymes. The difference in assignment success is best predicted by the statistic δ, which estimates cumulative allele frequency differences among populations (Shriver et

al. 1997). These results suggest genetic markers with many alleles are preferred when populations exhibit little genetic differentiation (such as pink salmon) because δ is more likely to be large due to genetic drift at each allele.

In Chapter five I use both computer simulation and actual microsatellite data from a known chinook salmon pedigree to determine how the number of loci, heterozygosity (H_E), parent population size and full sibs in the parent population affect pedigree reconstruction. The results are of practical importance to fish biologist and ecologists planning to conduct pedigree analysis using microsatellites. For example, the computer simulation study indicates assignment success for a given set of markers is very sensitive to H_E. A decrease in mean H_E of about 0.1 almost doubles the number of loci required to achieve a given level of assignment success. In contrast, moderate differences in parent population size have much less effect on assignment success, especially when mean H_E is high (0.750, 0.875). The simulation results show that when high assignment success is sought (e.g. 95%) there is little decrease in actual assignment success for a given set of gene markers if the parent population size doubles. A comparison of results from simulations and actual microsatellite data indicate pedigree reconstruction is confounded by the presence of full sibs in the parent population. I show that six loci with mean H_E = 0.87 is sufficient to achieve 97% unambiguous assignment in random mating population with no variance in reproductive success. The same loci provide only 67% unambiguous assignments in the chinook population in which each parent has between 2 and 11 full sibs. In this case parentage assignment success can be improved using parent pair-offspring likelihood analysis.

Summary and suggestions for future research

By examining odd-year and even-year pink salmon population structure and parent-offspring relationships in chinook salmon I show that microsatellites can be used to study any level of intra-specific genetic variation in Pacific salmon. There are, however, many factors that must be considered when collecting and interpreting microsatellite data from Pacific salmon, and some of these factors will vary depending

upon the level of genetic relationship (e.g. close kinship, individual/population ID, broad-scale population structure). I discuss three of these factors below.

When describing genetic population structure one must consider the choice of statistic (θ or ρ_{ST}). The results from Chapter three suggest that θ provides a better index of population structure than ρ_{ST}, although the later presumably accounts for the unique mutation mechanism and high rate of mutation of microsatellites. Nevertheless, the odd-year x even-year population comparison suggest ρ_{ST} may be appropriate in some cases. Further empirical evaluation is needed to determine what levels of relationship in what species of Pacific salmon favor the use of one or the other statistic. Also, further theoretical work is needed that incorporate the latest information on microsatellite mutation into statistics like ρ_{ST}.

Determining the population identity of individuals is integral to many salmon restoration programs. The results from Chapter four show that, with microsatellites, it is possible to distinguish individuals from two closely related populations. An important consideration in this type of analysis, however, is the method used to establish a criterion for assignment. The assignments for the Dungeness pink salmon are based on the distribution of genotype frequencies in the baseline population. Future research should compare this likelihood ratio approach to assignments based on confidence intervals generated by statistical sampling methods like bootstrapping.

When conducting parentage analysis one must consider how the complexity of the pedigree will influence parentage assignment sucess. The results from Chapter five show that while moderate differences in the parent population size have minor affect on assignment success, full sibs in the parent population clearly have a large affect on assignment sucess. Nevertheless, the impact of parental full sibs on parentage analysis was only examined in one chinook population. Further study is needed (e.g. computer simulation) to better understand how full sib parents impact parentage assignment success.

CHAPTER 1: SEMI-AUTOMATED MULTILOCUS GENOTYPING OF PACIFIC SALMON USING MICROSATELLITES

Introduction

Microsatellites have begun to be applied in fisheries and aquacultural contexts, and display particular promise in high resolution population and kinship studies (Bentzen et al. 1991; Wright and Bentzen 1994; Nielsen et al. 1994; McConnell et al. 1995a,b; O'Reilly and Wright 1995; Tessier et al, 1995). One advantage of microsatellites as genetic markers stems from the fact that particular microsatellite loci and their flanking sequences are conserved in related species. This means that primers developed for one species can frequently (although not invariably) be used to amplify polymorphic microsatellites in related species. This attribute offers particular benefits for the application of microsatellites in salmonid fishes, for which the sequences of >50 microsatellite primer pairs have been reported from cloning efforts in several species (see below).

Despite the positive attributes of microsatellites described above, the application of these markers may be limited by a number of operational difficulties. These include development costs, sample throughput constraints, and allele scoring difficulties. I address these issues below for Pacific salmon.

As noted above, the availability of numerous primer sequences is an advantage for the application of microsatellites to salmonids. This advantage, however, is offset to some extent by uncertainty about which microsatellite primers are most suitable for particular species/application issues. Here the availability of many primer sequences presents a dilemma of choice. It is relatively costly to purchase and test a large fraction of the available primers prior to any particular application, but failure to do so can lead to subsequent inefficiencies. A guide summarizing the performance of a wide range of primers would therefore be of general utility.

Sample throughput is another issue that closely affects the cost/benefit ratio of microsatellite application. Most microsatellite analyses have involved typing a single locus at a time. However, major gains in throughput as well as savings in consumables and labor are possible through combined analyses of multiple loci, a process known as multiplexing (Urquhart et al. 1995, O'Reilly et al. 1996). Multiplexing is the combination of allele fragments from more than one locus from an individual in a single lane of an electrophoretic gel. This can be achieved by co-amplification of multiple loci in the same PCR, post PCR mixing of amplified loci or both. Multiplexing using conventional labeling techniques is constrained by the wide allelic range of many salmonid microsatellites (O'Reilly et al. 1996). The use of multicolor fluorescent detection systems can improve throughput by permitting co-amplification of overlapping loci (Ziegle et al. 1992; Paetkau et al. 1995; Urquhart et al. 1995).

Some microsatellite loci, particularly those with dinucleotide repeats, produce ladders of bands following amplification via PCR. These "stutter bands", believed to result from slipped strand mispairing during DNA synthesis, make allele scoring difficult and may lead to mis-identification of allelic states using conventional scoring methods (Levinson and Gutman 1987; Hauge and Litt 1993; O'Reilly and Wright 1995). Automated genotyping systems using in-lane sizing standards and fluorescent detection have been shown to increase accurate detection of alleles (Ziegle et al. 1992; Schwengel et al. 1994).

This study has three objectives: First, I assess interspecific priming of various microsatellite primer pairs in Pacific salmon. I do this by completing a broad multispecies screening of 35 salmonid microsatellite primer pairs. These primers, developed from six different salmonid taxa, are tested in five species of *Oncorhynchus* and *Salmo salar*. Second, using the screening results, I demonstrate semi-automated multilocus genotyping by multiplexing six microsatellite loci in sockeye, coho and chinook salmon and scoring their alleles with a 4-color fluorescent detection system. Finally, I provide an indication of the allelic variability of microsatellites in these salmon. The potential application of multilocus microsatellite genotyping to studies of

Oncorhynchus spp. population structure as well as kinship analysis are discussed here and in a companion paper by Wenburg et al. (1996).

Methods

Microsatellite screening

Microsatellite loci were amplified by the PCR using recently developed salmonid primer pairs (Appendix A). PCRs were performed in a Perkin Elmer 9600 thermocycler. Two individuals were screened from the following species: chinook (*O. tshawytscha*), coho (*O. kisutch*), sockeye (*O. nerka*), pink (*O. gorbuscha*), chum (*O. keta*) and Atlantic salmon (*Salmo salar*). The DNA for microsatellite screening was extracted from a variety of tissues including fin, heart, and liver using the phenol/chloroform method (Hoelzel and Green 1992). Heart and liver tissue were stored frozen at -70°C whereas fin tissue was preserved in 100% ethanol. Microsatellite primers were synthesized in the Marine Molecular Biotechnology Laboratory (MMBL) on a Beckman Oligo 1000 DNA synthesizer. PCR was typically carried out in a 10 µL volume (10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM each dNTP, 0.4 units *Taq* polymerase, 0.3 µM primer, and 100-250 ng DNA template). DNA amplifications involved the following profile: 1 cycle of (94°C (2min)) + 7 cycles of (94°C (1min) + X°C (30sec) + 72°C (15sec)) + 18 cycles of (94°C (30sec) + X°C (30sec) + 72°C (15sec)), where X was an annealing temperature that varied among microsatellites. To account for possible base mismatches during inter-specific priming, the annealing temperature was set at 3-5°C below the primer melting temperature calculated from the nucleotide composition.

Results of each PCR were assessed using one of two fluorescent detection systems. In most instances I used a Molecular Dynamics FluorImager™ 575 to detect fluorescently stained microsatellite alleles. Typically, 5 µL of each PCR product and 1 µL loading buffer (15% w/v ficoll 400, 0.06% w/v bromophenol blue, 0.06% w/v xylene cyanol, 30 mM EDTA) was loaded on a 20 cm, 6% non-denaturing polyacrylamide gel and electrophoresed for approximately 2 h at 150 V. At least two lanes of each gel

contained 3 µL of Superladder-low 20 x 100 base pair (GenSura laboratories Inc.) size standard for estimating microsatellite allele length. Following electrophoresis the contents of each gel was stained with a 1:10,000 solution of SYBR™ Green 1 nucleic acid gel stain (Molecular Probes Inc.) and 1X Tris borate EDTA (TBE) buffer for 30 min and scanned on the FluorImager at a PMT voltage of 500-600.

Some microsatellites were detected using a Perkin Elmer-Applied Biosystems Inc. (ABI) 373A DNA Sequencer in GeneScan™ mode (ABI 1993). The gel preparation and analysis procedures for this system are discussed below. The ABI 373A produces a linear display (electropherogram) of emission intensity for each of four fluorescent labels. The PCR results were evaluated by the electropherogram when using the ABI 373A.

Following electrophoresis, two numerical codes were recorded for each species/locus combination, indicating the PCR product quality and approximate allele size. A quality code of 1-5 was assigned following a protocol modified from Pepin et al. (1995). The quality codes were defined as follows: (1) amplification of one or two bands and no stutter, (2) amplification of one or two bands and some stutter, (3) multiple bands and no smearing, (4) multiple bands and smearing, (5) no amplification at all. The allele sizes were coded as follow: (1) 60-120 b, (2) 120-180 b, (3) 180-240 b, (4) 240-300 b, (5) >300 b. Loci with alleles near the end of a given range were assigned two size codes.

Multiplex development

Multiplex systems were developed for chinook, coho and sockeye salmon using the ABI 373A to detect alleles. The ABI 373A uses an argon laser similar to that in the FluorImager but is capable of detecting and distinguishing the emission spectra of four fluorescent DNA labels (Ziegle et al. 1992). In this system, each dye label is depicted as a different color. Up to three loci with overlapping alleles may be multiplexed by labeling a single primer for each locus with a different color. The fourth color is reserved for the internal lane (sizing) standard. One of three colors was assigned to each microsatellite locus based on its observed allelic range previously detected with the

FluorImager 575. Labeled primers were synthesized at the MMBL as described above or purchased from Keystone Laboratories (961 Hamilton Ave., Menlo Park, CA. 94025).

I attempted co-amplification of various combinations ("triplex sets") of three microsatellite primer pairs for each of the three species. When formulating the species-specific triplex sets I chose from among those loci initially screened. Only loci with quality codes of 1 or 2 and similar primer annealing temperatures (+/- 2°C) were considered for grouping. For each species two triplex sets were combined before electrophoresis. Thus, no more than three loci from among any group of six were allowed to share the same allelic range. These criteria reduced the number of possible primer combinations for each species.

Co-amplification of each species-specific triplex was attempted in four individuals. The DNA extraction and PCR were carried out as above except that each PCR cocktail contained a 0.3 µM concentration of each of six primers. The PCR profile described above was used. If the annealing temperature differed for the three primer pairs, then the lowest was used for the initial PCR.

Samples from each PCR were electrophoresed on a 6% denaturing polyacrylamide gel to determine the quality of co-amplification. Approximately 1.0 µL of each PCR was combined with 3.15 µL formamide, 0.60 µL 50 mM EDTA and 0.25 µL (1.0 fmol) Perkin-Elmer GS350 internal size standard. All samples were denatured at 95°C for approximately 3 min, chilled on ice, and then loaded on the gel. Each gel was run for approximately 8 h at 25 W. Following the gel run, data were analyzed using the local Southern sizing algorithm in the GeneScan 672 analysis software, ver. 1.1 (ABI 1993). The electropherogram for each color label was used to determine which of the three loci in each triplex amplified. Those triplexes in which one or more of the loci did not amplify were not tested further. Those groups in which all loci amplified were optimized by adjusting individual primer concentrations to equalize signal intensity as depicted by peak height on the electropherogram (Urquhart et al. 1995; O'Reilly et al. 1996). Two groups of three loci were chosen for each of the three species based on

expected levels of polymorphism and quality of co-amplification. Finally, multiple screenings of the same individuals and loci during multiplex development provided a test of the repeatability of this system.

Automated multilocus genotyping with microsatellites

I genotyped two populations from each of the three species using the species specific multiplex suite of six microsatellites. Twenty five individuals from each population were screened. The populations included: Dungeness R., Puget Sound, Washington (Chinook 1), Stolle Meadow, south fork Salmon R., Idaho (Chinook 2), Togiak R., Western Alaska (Coho 1), Big Beef Ck., Hood Canal, Washington (Coho 2), Nikolai Ck., Kenai Peninsula, Alaska (Sockeye 1), and Okanagan R., Columbia River, Washington (Sockeye 2).

The DNA was extracted using a rapid, simplified cell lysis protocol modified from Hoelzel and Green (1992). Approximately 5-10 mg of tissue was placed in 100 µL of cell lysis buffer consisting of 40 mM TrisHCL at pH 9.5, 50 mM EDTA, and 0.5% TWEEN 20. One µL of proteinase K ($10 mg \cdot mL^{-1}$) was added to each sample before incubation at 37°C for approximately 12 h. The samples were then heated to 95°C for 15 min, centrifuged for 10 min at 17,000 x *g* and frozen at -20°C until needed for PCR. The DNA from coho and Nikolai Creek sockeye populations were obtained from fin tissue preserved in 100% ethanol. The DNA from the Okanagan River population was obtained from liver or heart tissue frozen at -70°C. Chinook DNA prepared using the phenol/chloroform extraction protocol was provided by outside sources.

The PCR profile described above was used for each triplex and population; only the annealing temperature was adjusted. Between 0.5 and 1.0 µL of DNA was used from the stock prepared by the lysis method and 100-250 ng was used from phenol/chloroform extracted DNA. Following PCR, 1 µL from each of the two triplexes was added to a 0.5 mL microtube and combined with 3.15 µL formamide, 0.60 µL 50 mM EDTA, and 0.25 µL(1.0 fmol) Perkin Elmer GS350 Tamra size standard. All samples were denatured at 95°C for approximately 3 min, chilled on ice, and then loaded on the gel. Each

individual sample was electrophoresed on the ABI 373A and analyzed automatically with the GeneScan 672 software ver. 1.1 (ABI 1993) as described above. Scoring of allele sizes for each locus and tabulation of data for importing into statistical software was performed with Genotyper software version 1.1 (ABI 1994)

Statistical Analysis

Conformity to Hardy-Weinberg equilibrium (HWE) was tested to evaluate within and between population genetic variation for each species using the algorithms of Louis and Dempster (1987) and Guo and Thompson (1992). A chi square analysis of independence between populations and allelic composition was computed according to Raymond and Rousset (1994). Computations were performed using GENEPOP ver. 1.1 (Raymond and Rousset 1995). Statistical significance levels (α) for the HWE and chi-square analyses were determined using sequential Bonferroni adjustments for simultaneous tests (Rice 1989).

Allelic variation and information content of each locus and multiplex suite for each species was measured by combining the data from each population. The number of alleles and allelic range were computed. The polymorphic information content (PIC) was computed for each species/locus combination according to Botstein et al. (1980). Finally, the matching probability (P_M) was computed for each locus and multiplex suite for each species using the most common genotypes (Edwards et al. 1992)

Results

Microsatellite screening

Table 1.1 summarizes the PCR annealing temperature, quality of the PCR product, and the estimated allelic range for each species/locus combination. The quality of some amplifications, particularly those receiving a grade of 3, may be improved by increasing the annealing temperature. The source species for each primer pair was included as a positive control with the exception of the three brown trout-derived loci. The positive control received a score of 3-5 in eight instances (positive control data for *Sfo*8, 12, 18, and 23 not shown). This occurred three times in steelhead (see Table 1 in

Wenburg et al. 1996) and brook trout (*Salvelinus fontinalis*), and twice in Atlantic salmon. In each case product quality was no better in the other species. As expected, the highest percentage of quality grades 1 and 2 occurred in those species from which some of the microsatellites were developed.

Table 1.2 summarizes PCR product quality for each species. The loci receiving a quality grade of 1 or 2 are summarized for each species as a percent of total loci scored. This value ranged from 31% (chum salmon) to 63% (chinook salmon). Only pink and chum salmon had fewer than 50% of the loci scoring 1 or 2. Between 57% and 74% of the loci given a quality grade of 1 or 2 also had an allelic range of 1 or 2 (60-180 b). In chinook, coho, sockeye, chum and Atlantic salmon, this value was greater than 65%. The small number of loci yielding amplification products greater than 200 b constrained the number of locus combinations for multiplexing.

Multiplex development

Multiplex systems comprised of suites of six loci were developed for coho, chinook and sockeye salmon using nineteen of the 35 microsatellite loci (Table 1.1). In a related project, Wenburg et al. (1996) developed multiplex systems for cutthroat (*Oncorhynchus clarki clarki*) and steelhead (*Oncorhynchus mykiss*). My goal of six-locus multiplex suites containing two triplex PCR combinations was based on allele size data from Table 1.1 which suggested I use each of the three color labels only twice to avoid potential allelic overlap. Further, my initial attempts at co-amplifying microsatellites in sockeye failed frequently when using more than three primer pairs.

Twenty-eight triplex combinations were tested, including thirteen in chinook, nine in coho and six in sockeye. At least two loci co-amplified in twenty six of the triplex combinations while three loci co-amplified in eleven combinations. Triplexing was successful in three of thirteen chinook combinations, five of nine coho combinations and three of six sockeye combinations. Two loci, *One*μ10 and *Ssa*14, did not amplify in the presence of other primer pairs and one locus, *Ssa*171, appeared limited to one or two alleles across taxa.

The species-specific triplex sets used in the population survey included a total of nine loci (Table 1.3). Optimization of each triplex reaction was achieved by adjusting individual primer concentrations, annealing temperature, and amount of template DNA. Table 1.3 shows the reaction conditions and fluorescent label assignments used for each of the six triplex combinations. Primer concentrations ranged from 0.05 μM to 0.5 μM per reaction. Between 0.5 μL and 1.0 μL of lysis prepared DNA was used.

Multiple independent screenings of the same individual and locus during multiplex development demonstrated the high resolving power and repeatability of fluorescent-based semi-automated genotyping. Two important observations were made. First, I achieved high allele scoring precision when PCR conditions were constant for most species and locus combinations. For example, the standard error of estimated allele size at four loci for three independent screenings (PCR and GeneScan) of four heterozygote individuals ranged from 0.03 to 0.06 for coho, chinook, sockeye and cutthroat (Table 1.4). Allele size estimates differed by no more than 0.10 b for the three runs, assuring reproducible scoring of heterozygote individuals with alleles that differ by as little as 2 b (see for example the coho, chinook and cutthroat in Table 1.4). By contrast, I found that changing PCR conditions for some species and locus combinations caused a shift of 1 b in scored allele size. For example, reducing the PCR primer concentration from 0.15 μM to 0.10 μM for *Omy*325 in coho simplified the product stutter pattern and resulted in a 1 b increase in the most abundant fragment. Increasing the PCR annealing temperature from 52°C to 57°C for *One*μ11 in sockeye eliminated a single base stutter and resulted in a 1 b increase in apparent allele size.

Multilocus genotyping

Six-locus multiplex surveys of 50 individuals each of chinook, coho, and sockeye revealed extensive genetic variation (Table 1.5). The average number of alleles per locus was 13, 13 and 9 for chinook, coho and sockeye, although the number of alleles at individual loci was quite variable, ranging from 2 (*One*μ11 in coho and *One*μ1 in sockeye) to 31 (*Omy*325 in coho). The allelic size range of individual loci was also quite

variable (2-80 b, mean = 30 b). About 75% of the amplified microsatellite products did not exceed 200 b. Of the six loci used in chinook salmon, two with the same fluorescent label (*Ots*1 and *Ssa*85) were found to have overlapping allelic ranges. This overlap was due to a rare occurrence of a 147 b allele at the *Ots*1 locus and was verified by separate amplification and electrophoresis of each locus. The peak profile of each locus appears to produce a unique amplification signature that may be useful for distinguishing rare instances of overlap. For further explanation of signature peak profile see Figure 2 in Wenburg et al. (1996)

The relative utility of each locus as a genetic marker for fine scale genetic studies was assessed by estimating two index parameters. First, PIC provides an index of the informativeness of a locus for use in gene mapping as well as kinship and parentage studies. Values greater than 0.5 are considered highly informative (Botstein et al. 1980). Of the 18 species/locus combinations surveyed, 14 had PIC values greater than 0.5 and nine had values greater than 0.7. The probability of a match (P_M) (finding two unrelated individuals with the same genotype) was conservatively estimated for each species/locus combination and each multiplex group using the most common genotype. P_M values for most loci fell below 0.1. Combined P_M values were computed for each multiplex set and were 4.0×10^{-10}, 7.2×10^{-8} and 3.2×10^{-7} for chinook, coho and sockeye, respectively.

Table 1.6 summarizes the observed and expected percent heterozygosity for each population and locus combination. Here also, the differences in variability among loci are evident. Heterozygosities ranged from zero (*One*µ11 in Togiak River coho) to 0.96 (*Ssa*85 in Dungeness River chinook). No instances of significant departure from Hardy-Weinberg equilibrium were found (initial $\alpha = 0.025$).

Table 1.7 summarizes the chi-square pairwise test of independence between populations and allele frequencies. Significant differences were found for all loci except *One*µ1 (sockeye), *One*µ11 and *Ots*4 (coho) and *Ots*1 (chinook) (initial $\alpha = 0.008$). The probability of independence at locus *Ots*4 (coho) was marginally insignificant ($P = 0.026$, $\alpha = 0.025$).

Discussion

Inter-specific priming of microsatellites

My screening of microsatellites in Pacific salmon demonstrates that sequence conservation in priming regions often permit inter-specific exchange of primers. These results are supported by previous findings (e.g. McConnell et al 1995; Morris et al. 1996), suggesting microsatellite based genetic studies of Pacific salmon are possible using existing primers. The high cost and time associated with creating microsatellite libraries and designing primers need not be a constraint. Further, my data should be helpful in directing researchers to useful species/primer combinations, reducing development costs associated with primer testing.

By exchanging primers across species of salmon I am assuming they amplify homologous loci. Evidence supporting this assumption has been found in salmonid fishes, mammals and sea turtles (Morris et al. 1996, Forbes et al. 1995, Fitzsimmons et al. 1995, Pepin et al. 1995). However, final verification will require sequencing of the PCR product. This is of particular importance when conducting phylogenetic surveys across taxa (Forbes et al. 1995, Estoup et al. 1995).

In some instances the degree of complementarity between a primer and the microsatellite flanking sequence may vary among alleles within a species. In extreme cases some alleles will not amplify. The presence of "null" alleles can be inferred through population level screening and testing for departures from Hardy-Weinberg equilibrium (Callen et al. 1993). Using this approach I did not see evidence of null alleles in the six loci screened in coho, chinook and sockeye. Likewise, Wenburg et al. (1996) did not see evidence of null alleles in the loci screened in steelhead and cutthroat. However, in an earlier study, I did observe evidence of one or more null alleles at microsatellite *Ssa*293 in sockeye (Bentzen and Olsen, unpublished data).

Microsatellite multiplexing and data processing

I have demonstrated the potential for rapid throughput of Pacific salmon microsatellites using one ABI 373A automated sequencer/genescanner and analysis

software. Using the six locus multiplex systems described here and those in Wenburg et al. (1996) it is possible to process 432 genotypes per day (6 loci x 72 individuals) running two GeneScan gels. I anticipate increasing this rate by multiplexing additional microsatellites. The scarcity of loci greater than 200-300 b in my current multiplex systems should allow for added loci. I feel a realistic short term goal is nine loci (648 genotypes per day), and as more data is gathered on the allelic range of microsatellites in Pacific salmon it may be possible to exceed 10 loci per lane. In some cases microsatellite flanking sequences may be used to re-position primers and "customize" loci to maximize vertical gel space on the ABI 373A.

This rate of production results in a large volume of data; hence, efficient analysis requires precision in allele scoring and rapid data processing. I have shown that reproducible allele scoring is possible for dinucleotide microsatellites when heterozygotes may differ by as little as two bases. Using the Genotyper™ software (ABI 1994), I have developed species specific template files that allow rapid allele scoring and data synthesis into formats easily imported by statistical software such as GENEPOP (Raymond and Rousset 1995).

I found that 1 b shifts in apparent allele size for some microsatellites may occur when changing the PCR temperature profile and reaction mixture. Brownstein et al (1996) demonstrated that variable adenylation, influenced by the PCR profile and nucleotide sequence at the 5' end of the reverse primer, is one mechanism that can control single base shifts in PCR product size. Brownstein et al. (1996) also showed that the degree of adenylation can be controlled by designing 5' "tails" for reverse primers and altering the PCR profile. My results suggest that adherence to a single PCR temperature profile and reaction mixture will help assure reproducible results when using primers prone to adenylation. However, modification and testing of adenylation prone salmonid microsatellite primers, following the protocol of Brownstein et al. (1996), will likely result in significant reduction in single base stutter by promoting or inhibiting adenylation.

Microsatellite based genetic studies of Pacific Salmon

The use of microsatellites for genetic studies of Pacific salmon is in its infancy. My results provide some of the first indications of the level of polymorphism of these markers in three species of Pacific salmon. The results are encouraging and in general agree with the range of microsatellite variation reported for other fishes and in particular Atlantic salmon. The high level of polymorphism found in the multiplexed loci suggest assessment of fine scale population structure of Pacific salmon, unresolvable by other genetic markers, may be possible with microsatellites. In general, the PIC values reported here are similar to those reported for microsatellites in mammals (Ostrander et al., 1993; Pepin et al. 1995). Further, the probability of match for the individual loci and multiplex suites are similar to those reported by Urquhart et al. (1995) for human microsatellites and O'Reilly et al. (1996) for Atlantic salmon microsatellites. My data suggest that these markers may be employed when high levels of discrimination are needed, as in kinship studies and parentage analysis.

Table 1.1. Salmonid microsatellite screening results. See text for quality (Q) and allele size (S) codes.

Microsat. Locus[a]	Chinook (T)-(Q)-(S)	Coho (T)-(Q)-(S)	Sockeye (T)-(Q)-(S)	Pink (T)-(Q)-(S)	Chum (T)-(Q)-(S)	Atlantic (T)-(Q)-(S)
Fgt1	(56)-2-3,4	(56)-4-3,4	(56)-2-3	(56)-2-2	(56)-3-4	(56)-4-3,4
Omy77	**(50)-2-2**	**(50)-2-2**	**(50)-2-1**	(50)-5	(50)-5	(50)-2-1
Omy78	(55)-4-1	(55)-4-1	(55)-4-1	(55)-4-1	(55)-4-1	(55)-4-1
Omy87	(55)-4-1,2	(55)-4-2,3	(55)-3-2	(55)-3-1,2	(55)-4-3,4	(55)-4-1,2
Omy207	(56)-1-1	**(56)-1-1**	(56)-5	(53)-5	(53)-5	(53)-1-1
Omy293	(55)-5	(55)-5	(55)-5	(55)-5	(55)-5	(55)-5
Omy325	**(58)-2-1**	**(58)-2-1,2**	**(60)-2-2**	(55)-2-3	(55)-3-2,3	(58)-2-2,3
*One*μ1	(58)-5	**(58)-2-2**	**(58)-2-1**	(58)-2-1	(58)-5	(58)-2-1,2
*One*μ2	(58)-5	**(58)-2-2,3**	**(58)-2-3,4**	(57)-5	(57)-2-4	(58)-2-3,4
*One*μ8	**(58)-2-2**	(55)-2-3	**(58)-2 -3**	(55)-2-4	(55)-2-3	(55)-5
*One*μ10	(57)-5	(57)-5	(57)-2-1,2	(57)-5	(57)-5	(57)-5
*One*μ11	**(58)-1-2**	**(58)-1-2**	**(58)-1-2**	(58)-1-2	(58)-1-2	(58)-1-2
*One*μ14	**(58)-2-3**	**(58)-2-3,4**	**(58)-2-2**	(58)-5	(58)-5	(58)-2-3
*Ots*1	**(50)-1-3**	**(50)-1-3**	(50)-5	(50)-2-3	(50)-2-2	(50)-1-3
*Ots*2	**(48)-2-1**	**(48)- 2-1**	(48)-5	(50)-5	(50)-5	(48)-5
*Ots*3	**(50)-2-1**	(50)-5	**(50)-2-1**	(50)-5	(50)-2-1	(50)-1-1
*Ots*4	**(56)-2-2**	**(54)-2-2**	(57)-2-2	(54)-5	(54)-5	(48)-2-1
*Ots*5	(45)-2-2	(45)-5	(45)-5	(45)-5	(45)-2-2	(45)-5
*Ots*6	**(57)-1-3**	**(57)-1-3**	(57)-5	(57)-5	(57)-5	(57)-5
PuPuPy	(53)-5	(53)-5	(52)-4-5	(53)-5	(53)-5	(53)-5
*Sfo*8	**(60)-2-4**	**(60)-2-4**	**(60)-2-4**	(55)-2-4	(55)-2-4	(60)-2-3
*Sfo*12	(50)-5	(50)-5	(50)-5	(50)-5	(50)-5	(50)-5
*Sfo*18	(52)-5	(52)-5	(52)-5	(52)-5	(52)-5	(52)-5
*Sfo*23	(52)-5	(52)-5	(52)-5	(52)-5	(52)-5	(52)-5
*Ssa*4	(57)-2-2	(57)-2-2	(57)-2-2	(57)-3-4	(57)-5	(57)-2-2
*Ssa*14	**(52)-1-2,3**	(52)-1-2	(52)-1-2	(52)-1-2	(52)-1-2	(52)-1-2
*Ssa*85	**(58)-2-2**	(60)-5	**(58)-2-2**	(57)-3-3	(57)-3-3,4	(58)-2-1
*Ssa*171	(56)-1-1	(56)-1-1	**(56)-1-1**	(57)-5	(57)-5	(56)-1-3,4
*Ssa*197	(57)-1-3,4	(57)-1-1	(57)-4-1	(57)-1-2	(57)-1-2	(57)-1-2
*Ssa*202	(58)-5	(58)-5	(58)-5	(58)-5	(58)-5	(58)-5
*Ssa*289	(46)-5	(46)-5	(46)-5	(46)-5	(46)-5	(46)-5
*Ssa*293	(53)-2-1	(53)-2-1	**(53)-2-1**	(53)-2-3	(53)-2-2	(53)-2-1,2
μ*Sat*15	(57)-5	(57)-5	(57)-5	(57)-5	(57)-5	(57)-5
μ*Sat*60	(60)-2-2	(60)-3-1	(60)-2-2	(57)-2-1	(57)-3-3	(60)-2-2
μ*Sat*73	(57)-2-2	(57)-2-2	(57)-3-2	(57)-2-2	(57)-2-1,2	(57)-5

[a] (T) PCR annealing temperature; (Q) quality of the PCR amplification; (S) size range (bases) of the PCR product. The loci used in multiplex development are shown in bold.

Table 1.2. Summary of PCR product quality by species for all microsatellite loci screened.

Species	No. Loci Screened	Score 1-2	% Score 1-2	Score 3-4	No Product
Chinook	35	22	63%	2	11
Coho	35	19	54%	4	12
Sockeye	35	18	51%	5	12
Pink	35	12	34%	4	19
Chum	35	11	31%	6	18
Atlantic	35	18	51%	3	14

Table 1.3. Multiplex set composition: fluorescent label assignments, PCR annealing temperature, and primer concentration for each triplex set.

	Anneal	Microsatellite loci and primer concentration (μM)		
Triplex set	(°C)	6Fam(blue)	Hex(yellow)	Tet(green)
Chinook A	58		*Ssa*85(0.30)	*Omy*325(0.20)/ *One*μ8(0.40)
Chinook B	54	*Ots*4(0.25)/ *One*μ14(0.60)	*Ots*1(0.40)	
Coho A	58	*One*μ2(0.25)	*One*μ1(0.50)	*Omy*325(0.10)
Coho B	54	*One*μ11(0.20)	*Ots*4(0.45)	*Ots*1(0.20)
Sockeye A	58	*One*μ14(0.19)	*One*μ1(0.13)	*One*μ11(0.15)
Sockeye B	58	*One*μ2(0.08)	*Ssa*85(0.15)	*One*μ8(0.16)

Table 1.4. Microsatellite allele scoring variation. Each run for each species and locus combination represents an independent PCR and GeneScan event.

			Allele Size (b)				
Species	Sample	Locus-Allele	Run 1	Run 2	Run 3	Avg.	S.E.
Chinook	C01	*Ots*4-1	146.35	146.34	146.25	146.31	0.06
		*Ots*4-2	148.50	148.44	148.42	148.45	0.04
Coho	T03	*One*μ1-1	171.70	171.75	171.67	171.71	0.04
		*One*μ1-2	173.63	173.62	173.56	173.60	0.04
Sockeye	NM01	*One*μ14-1	135.10	135.13	135.20	135.14	0.05
		*One*μ14-2	146.15	146.17	146.20	146.17	0.03
Cutthroat[a]	VI08	*Omy*77-1	135.11	135.01	135.05	135.06	0.05
		*Omy*77-2	137.01	136.96	137.02	137.00	0.03

[a] Data from Wenburg et al. 1996

Table 1.5. Allelic variability. Number of alleles (A), allele range (R), polymorphic information content (PIC), and match probability (P_M) for the most frequent genotype for each species/locus combination.

Stat[a].	*Omy*325	*One*µ1	*One*µ2	*One*µ8	*One*µ11	*One*µ14	*Ots*1	*Ots*4	*Ssa*85	6 loci
Chin										
A	10			14		23	6	8	17	13
R	89-125			159-191		183-249	147-195	142-162	118-170	
PIC	0.758			0.873		0.906	0.519	0.681	0.885	0.770
P_M	0.020			0.010		0.010	0.116	0.090	0.020	4.2×10^{-10}
Coho										
A	31	9	28		2		5	4		13
R	92-174	168-184	200-262		140-142		187-199	134-140		
PIC	0.934	0.745	0.924		0.057		0.381	0.333		0.562
P_M	0.004	0.048	0.004		0.884		0.336	0.384		8.8×10^{-8}
Sock										
A		2	11	8	3	9			20	9
R		112-114	262-290	194-214	146-156	129-151			129-187	
PIC		0.136	0.767	0.691	0.511	0.675			0.834	0.602
P_M		0.706	0.032	0.068	0.102	0.078			0.026	3.2×10^{-7}

Table 1.6. Observed heterozygosity (H_O), and expected heterozygosity (H_E) for each population/locus combination.

Pop.[a,b]	*Omy*325	*One*μ1	*One*μ2	*One*μ8	*One*μ11	*One*μ14	*Ots*1	*Ots*4	*Ssa*85	Avg./pop
Chin1										
H_O	0.64			0.76		0.84	0.56	0.56	0.96	0.72
H_E	0.75			0.84		0.89	0.60	0.61	0.89	0.75
Chin2										
H_O	0.72			0.68		0.80	0.56	0.60	0.84	0.70
H_E	0.66			0.82		0.81	0.61	0.77	0.82	0.75
Coho1										
H_O	0.92	0.76	0.80		0.00		0.08	0.36		0.49
H_E	0.90	0.70	0.82		0.00		0.08	0.31		0.47
Coho2										
H_O	0.96	0.84	0.92		0.12		0.68	0.36		0.65
H_E	0.96	0.83	0.95		0.12		0.63	0.41		0.65
Sock1										
H_O		0.12	0.76	0.72	0.56	0.72			0.72	0.60
H_E		0.12	0.80	0.75	0.51	0.76			0.73	0.61
Sock2										
H_O		0.20	0.68	0.64	0.36	0.52			0.84	0.54
H_E		0.18	0.74	0.65	0.51	0.56			0.91	0.59

[a] Chin1=Dungeness R., Chin2=Stolle Meadow, Coho1=Togiak R., Coho2=Big Beef Ck., Sock1=Nikolai Ck., Sock2=Okanagan R.

[b] The difference between H_O and H_E was not significant for any locus/population combination.

Table 1.7. Results of the pairwise test of independence between populations and allelic composition.

Population[a,b]	*Omy325*	*Oneμ1*	*Oneμ2*	*Oneμ8*	*Oneμ11*	*Oneμ14*	*Ots1*	*Ots4*	*Ssa85*
Chin1xChin2									
P	<0.001*			<0.001*		<0.001*	0.075	<0.001*	<0.001*
S.E.	0.000			0.000		0.000	0.006	0.000	0.000
Coho1xCoho2									
P	<0.001*	<0.001*	<0.001*		0.246		<0.001*	0.026	
S.E.	0.000	0.000	0.000		0.003		0.000	0.003	
Sock1xSock2									
P		0.720	<0.001*	<0.001*	<0.001*	<0.001*			<0.001*
S.E.		0.003	0.000	0.000	0.000	0.000			0.000

[a] Chin1=Dungeness R., Chin2=Stolle Meadow, Coho1=Togiak R., Coho2=Big Beef Ck., Sock1=Nikolai Ck., Sock2=Okanagan R.

[b] The P-value is the probability of independence between populations and the allele composition. S.E. is the standard error of the P-value estimate. Probability values judged significant following sequential Bonferroni adjustment (initial $\alpha = 0.008$) are indexed with an asterisk (*).

CHAPTER 2: CHARACTERIZATION OF SEVEN MICROSATELLITE LOCI DERIVED FROM PINK SALMON

Introduction

Genetic studies of pink salmon population structure have relied largely on allozymes (e.g. Skaklee et al. 1991) but recent interest in fine-scale resolution of population structure, kinship analysis, studies of reproductive success and gene mapping have contributed to the need for highly polymorphic, neutral, and heritable markers. Microsatellites meet these criteria. To date no microsatellites have been isolated from pink salmon and attempts at cross-species amplification of microsatellite loci have not been as successful in pink salmon as in other *Oncorhynchus* (Olsen et al. 1996). I report here the development of primers for seven polymorphic microsatellite loci in pink salmon.

Methods

DNA from heart tissue was extracted from a single individual, digested with *Mbo*I and size fractionated by agarose gel electrophoresis. DNA fragments between ~ 300 bp and 800 bp were recovered from the gel using phenol-chloroform extraction and ligated into pZErO^{TM} –2.1 vector (Invitrogen Corp., San Diego, CA) cut with *Bam*HI. TOP10F' (Invitrogen Corp., San Diego, CA) competent cells were transformed with ligated vector, plated and allowed to incubate overnight at 37°C. Approximately 3,600 colonies were lifted using nylon membranes and probed with fluorescein-labeled oligonucleotides [$(GT)_{15}$, $(GA)_{15}$] (Amersham Life Science Inc., Arlington Heights, IL.). Approximately 400 positive colonies (presumably with DNA inserts complementary to the labeled probes) were identified following stringency washes and signal amplification using the Vistra signal amplification kit (Amersham Life Science Inc., Arlington Heights, IL.). A Molecular Dynamics FluorImager 575 was used to visualize positive colonies on probed membranes. Sixty-three positive colonies were selected from the initial screen of which

57 were positive following a second screen. DNA from 21 positives with inserts of 150-1500 bp was extracted using the Qiaprep plasmid DNA prep kit (Qiagen Inc., Santa Clarita, CA.) and sequenced using an Applied Biosystems Inc. (ABI) Prism *Taq* DyeDeoxy™ terminator chemistry and an ABI 373A automated sequencer. Microsatellite sequences were identified in 18/21 positive inserts of which 11 contained adequate flanking region to locate primers. Primers were designed using the program Primer 3 (Rozen and Skaletsky 1996) and purchased from Life Technologies Inc. (Grand Island, NY). One sequence contained two linked microsatellites; the first a perfect tetranucleotide repeat [$(GTCT)_{26}$] separated by 70 bp from a compound dinucleotide repeat [$(AT)_5 (GT)_{11}$]. For this sequence three primer pairs were designed, the first two to amplify each microsatellite separately and the third to amplify both together.

Primer pairs were tested for amplification effectiveness in pink salmon and 13 other salmonid species representing four genera. PCR was carried out in 10 µL volumes (10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM each dNTP, 0.5 units *Taq* polymerase (Promega, Madison, WI), 0.3 µM each primer, and 100 ng DNA template) using a Perkin Elmer model 9600 thermo cycler. DNA amplifications involved the following profile: one cycle of 94°C (2 min); seven cycles of 94°C (1 min) + X°C (30 s) + 72°C (15 s); and 18 cycles of 94°C (30 s) + X°C (30 s) + 72°C (15 s) where X was an annealing temperature that varied among primer pairs. A Molecular Dynamics FluorImager 575 was used to detect fluorescently stained microsatellite alleles electrophoresed in a 6% non-denaturing polyacrylamide gel. Primer pairs for 9 sequences yielded high quality amplification product in pink salmon (Table 1).

Results and Discussion

An estimate of microsatellite polymorphism in pink salmon was obtained by screening 20 individuals using the ABI 373A in GeneScan mode (ABI 1993) to detect fluorescein-labeled primers in a 6% denaturing polyacrylamide gel. Scoring of alleles for each locus was performed with Genotyper software, ver. 2.0 (ABI 1996a). The number alleles per locus ranged from 2 (Ogo6) to 24 (Ogo1c) and averaged 13 (Table 1). The

expected heterozygosity ranged from 0.14 (Ogo6) to 0.95 (Ogo1c) and averaged 0.72. A seven base "pigtail" consisting of the sequence GTTTCTT (Brownstein et al. 1996) was added to the 5' end of the non-labeled primer for microsatellite Ogo2 to remove single base "stutter" caused by adenylation.

At least some of the microsatellites amplified in the 13 salmonid species tested (Table 2). The mean number of loci that amplified was highest among the six pink salmon congeners (7.7). Over half the loci amplified in *Salmo* spp. (6.5) and *Salvelinus* spp. (5.8) and three loci amplified in the most distantly related taxon, arctic grayling (*Thymallus arcticus*). These microsatellites should prove useful for a number of conservation genetic applications in pink salmon and, to a lesser degree, the other species examined here.

Table 2.1. Repeat sequences for seven pink salmon microsatellite loci. Allele size range, allele number and the estimate of expected heterozygosity (H_E) were derived from a sample of 20 individuals.

Locus	Repeat sequence of cloned allele	Primer sequence (5'-3') (F>forward, R>reverse)
Ogo1a	$(GTCT)_{26}$	F>GATCTGGGCCTAAGGGAAAC
		R>ACTAGCGGTTGGAGAACCC
Ogo1b	$(AT)_5(GT)_{11}$	F>AGGGTTCTCCAACCGCTAGT
		R>CGCAAGCCCAAACAGATAA
Ogo1c	$(GTCT)_{26}X_{70}(AT)_5(GT)_{11}$	F>CAATCGCTCTCTCGCTACACT
		R>CGCAAGCCCAAACAGATAA
Ogo2	$(GA)_{24}$	F>ACATCGCACACCATAAGCAT
		R>GTTTCTTCGACTGTTTCCTCTGTGTTGAG
Ogo3	$(GT)_{42}$	F>CATGTAAGGAATGCAGTTTAGTGTC
		R>ACGTTAGGAGTGAGGCGGT
Ogo4	$(GT)_{55}$	F>GTCGTCACTGGCATCAGCTA
		R>GAGTGGAGATGCAGCCAAAG
Ogo5	$(GT)_4TT(GT)_2GCAT(GT)_8$	F>GGTTTGACATTTAAGGCGGA
		R>GGGTGTTCAAGCTCACTGCT
Ogo6	$(GT)_6TT(GT)_6TT(GT)_4X_8$	F>CACACGCATACTCCGAAACA
	$(GT)_3TT(GT)_2$	R>AGAGCTGTCATGCACCAGAA
Ogo8	$(GA)_{22}$	F>TCGCAGAGCGATACCAATG
		R>GAGGAAGACCATTGAGGTGG

Table 2.1. continued

Locus	Repeat sequence of cloned allele	Anneal (°C)	Size range (bp)	No. of alleles	H_E	GenBank no.
Ogo1a	$(GTCT)_{26}$	59	183-323	21	0.94	AF007827
Ogo1b	$(AT)_5(GT)_{11}$	60	157-175	6	0.67	""
Ogo1c	$(GTCT)_{26}X_{70}(AT)_5(GT)_{11}$	60	260-406	24	0.95	""
Ogo2	$(GA)_{24}$	58	230-286	8	0.74	AF009794
Ogo3	$(GT)_{42}$	59	299-359	16	0.90	AF009795
Ogo4	$(GT)_{55}$	60	217-261	16	0.89	AF009796
Ogo5	$(GT)_4TT(GT)_2GCAT(GT)_8$	55	184-204	4	0.59	AF009797
Ogo6	$(GT)_6TT(GT)_6TT(GT)_4X_8$	60	205-219	2	0.14	AF009798
Ogo8	$(GT)_3TT(GT)_2$ $(GA)_{22}$	55	96-154	8	0.69	AF009780

Table 2.2. Results of cross-species testing of pink salmon microsatellite primers.

Species	Ogo1a	Ogo1b	Ogo1c	Ogo2	Ogo3	Ogo4	Ogo5	Ogo6	Ogo8
Oncorhynchus									
gorbuscha	Y	Y	Y	Y	Y	Y	Y	Y	Y
nerka	-	Y	Y55	Y	Y*	Y	Y	Y	Y
keta	-	-	-	Y	Y	Y	Y	Y	Y
kisutch	Y*	Y	Y*	Y	Y	Y	Y	Y	Y
tshawytscha	Y*	-	-	Y	Y	Y	Y	Y	Y
mykiss	Y	-	Y	Y	Y*	Y	Y	Y	Y55
clarki	-	Y55	Y	Y	Y*	Y	Y	Y	Y
Salmo									
salar	Y	-	Y	Y	Y	Y	Y	-	Y
trutta	Y	-	Y	Y	Y*	Y	Y	-	-
Salvelinus									
malma	Y	Y	Y	Y	-	Y	Y	-	-
alpinus	Y	Y	Y	Y	-	Y	Y	-	-
fontinalis	Y	Y	-	Y	-	Y	Y	-	-
namaycush	Y	Y	Y55	Y	-	Y	Y	-	-
Thymallus									
arcticus	-	-	-	Y	-	Y55	Y	-	-

(Y) = amplified at designed annealing temperature, (-) = did not amplify, (55) = amplified at 55°C, (*) = amplified product at least 100bp smaller than clone

CHAPTER 3: GENETIC INTERPRETATION OF BROAD-SCALE MICROSATELLITE POLYMORPHISM IN ODD-YEAR PINK SALMON

Introduction

The pink salmon *Oncorhynchus gorbuscha* is the most abundant species of Pacific salmon and is of significant economic, ecological, and cultural importance to coastal communities of the Pacific rim (Heard 1991). Spawning populations are present in freshwater drainages on both east and west shores of the Pacific Ocean north of about 40°N (Heard 1991). Like most *Oncorhynchus*, pink salmon are anadromous and semelparous. They are philopatric and often exhibit temporal separation of spawning aggregations within drainages (Heard 1991; Bue et al. 1996).

Pink salmon in their native range exhibit a rigid two-year life cycle that has resulted in two reproductively isolated odd- and even-year lineages (Davidson 1934, Aspinwall 1974). Genetic studies using allozyme loci show divergence of the two lineages is the most significant genetic subdivision in pink salmon (Beacham et al. 1985; Beacham et al. 1988, Zhivotovsky et al. 1994). Spawning aggregates within each lineage also reveal significant population structure (Gharrett et al. 1988; Shaklee et al. 1991; Varnavskaya and Beacham 1992; Shaklee and Varnavskaya 1994; Hard et al 1996; but see Omelchenko 1994). Recent studies of mitochondrial DNA variation in Russian (Brykov et al. 1996) and Alaskan (Seeb et al. 1996) pink salmon also show evidence of genetic structure with the greatest variation occurring between odd- and even-year populations.

Estimates of genetic distance among large coastal aggregates of odd-year pink salmon (e.g. Russia, Alaska, British Columbia and Washington) generated from allozyme frequency data appear to support a model of genetic isolation by distance (Varnavskaya and Beacham 1992; Shaklee and Varnavskaya 1994). Phenetic analyses suggest North American odd-year populations form two distinct groups; populations from Alaska and Northern British Columbia and populations from southern British Columbia and

Washington State (Varnavskaya and Beacham 1992; Shaklee and Varnavskaya 1994). However, the statistical validity of these putative regional groupings and genetic isolation by distance have not been tested. In addition, no studies have examined the genetic structure of odd-year populations from a broad geographic range (e.g. North America) using a different class of genetic marker such as microsatellites. Microsatellite data would strengthen support for inferred genetic relationships based solely allozymes if the two marker classes provide concordant results (Avise 1994).

Microsatellites have potential for a variety of genetic studies of salmonids including population genetics and kinship analysis (Bentzen et al. 1994; O'Reilly and Wright 1995). However, it is not clear how to make best use of microsatellite data given the apparent high rate and stepwise mode of mutation (see Jarne and Lagoda 1996). Recent theoretical efforts have focused on developing new statistics to estimate genetic distance and genetic population structure from microsatellite data based on a stepwise mutation model (Slatkin 1995; Goldstein et al. 1995). Slatkin (1995) introduced the parameter R_{ST} as an index of subpopulation structure analogous to θ defined by Weir and Cockerham (1984). In contrast to θ, R_{ST} accounts for differences in allele size under a stepwise mutation model rather than simple identity or non-identity of allelic states under an infinite allele model (Slatkin 1995). Simulation studies reveal that R_{ST} is a more accurate estimate of genetic structure when migration is small, the difference in average coalescence times within and between populations is large, and the mutation rate is high (e.g. 10^{-3}) (Slatkin 1995). However, R_{ST} and θ approach equality as migration increases, the average coalescence times converge, and the mutation rate decreases (Slatkin 1995). Unfortunately, these simulations do not capture the complexity of interactions between continuous population parameters such as migration rate and average coalescence time, complicating the choice of which statistic to use. Further, these parameters are rarely known with any certainty. Only empirical analysis, comparing both statistics for estimating genetic population structure, will clarify which makes best use of the microsatellite data for a particular taxa (Forbes et al. 1995).

The purpose of this study was to use microsatellite loci to extend our knowledge of genetic variation in North American odd-year pink salmon. I sampled five loci in 13 populations (12 odd- and 1 even-year) from six geographic regions. My objectives were to test for significant inter-population allelic variation; compare θ and R_{ST} to determine which is the most appropriate measure of population subdivision; test for genetic isolation by distance among odd-year populations; test for significant regional and population level genetic structure. My results provide one of the first broad-scale estimates of intraspecific microsatellite polymorphism in Pacific salmon.

Methods

Sample collection and preparation

I sampled 12 populations comprising one pair from each of six geographic regions representing the North American range of odd-year pink salmon (Figure 3.1). In addition, I included one even-year population from Southcentral Alaska. Sample sizes ranged from 20 to 52 (Figure 3.1). Tissue samples (heart and liver) from adult pink salmon were stored at -80°C prior to DNA extraction. Total genomic DNA was extracted from approximately 50-100 mg of frozen heart or liver tissue using a Gentra Systems™ (Minneapolis MN) Puregene DNA isolation kit. Precipitated DNA was hydrated in TE buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0) and heated at 55°C for approximately 12 h. The DNA concentration was measured by spectrophotometry and diluted to 100 ng/µL for use in the polymerase chain reaction (PCR).

Microsatellite selection and multilocus genotyping

Primer pairs for 21 microsatellite loci were chosen among 51 previously screened in pink salmon (see Olsen et al. 1996; Scribner et al. 1996) and tested in at least four individuals to re-evaluate amplification potential. All PCRs were performed with a Perkin Elmer 9600 thermocycler in a 10 µL volume (10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM each dNTP, 0.5 units *Taq* polymerase, 0.1-0.5 µM each primer, and 100 ng DNA template). Primers were purchased from Operon Inc.

(Alameda, CA). One primer of each pair contained one of three fluorescent labels for fragment detection by the Perkin Elmer-Applied Biosystems Inc. (ABI) 377 Prism™ semi-automated fluorescent detection system.

I was able to amplify 11 loci, five of which were selected for use in this study based on quality of the PCR product and the presence of at least two alleles in at least four fish. Four loci (*One*μ3, Scribner et al. 1996; *Ots*1, D. Hedgecock, U.C.D., personal communication; μ*Sat*60, Estoup et al. 1993; *Ssa*85, O'Reilly et al. 1996) consisted of arrays comprised of dinucleotide repeats, and one (*Ssa*197, O'Reilly et al. 1996) of primarily tetra-nucleotide repeats. I tested PCR co-amplification (multiplexing) using various primer pair combinations following the methods described in Olsen et al. (1996). The multiplex system consisted of three PCRs (*One*μ3 and *Ots*1, anneal at 52°C; μ*Sat*60 and *Ssa*197, anneal at 58°C; *Ssa*85, anneal at 58°C). The following PCR profile was used: 1 cycle of (94°C (2 min)) + 7 cycles of (94°C (1 min) + X°C (30 sec) + 72°C (15 sec)) + Y cycles of (94°C (30 sec) + X°C (30 sec) + 72°C (15 sec)), where X (annealing temperature) and Y (cycles) varied among microsatellites.

Microsatellites were size fractionated using the ABI 377 Prism™ in GeneScan™ mode (ABI 1996b). For each sample approximately 0.5 μL from each of three PCRs was combined with 2.5 μL formamide, 0.50 μL 25 mM EDTA and 0.5 μL (1.0 fmol) Perkin-Elmer GS350 internal size standard in a 0.5 mL microcentifuge tube. The samples were denatured at 95°C for approximately 3 min, chilled on ice, and loaded on a 4.5% denaturing polyacrylamide gel. Approximately 2.5 μL from each sample was loaded per well. Each gel was run for approximately 2 h at 3000 V. Data were analyzed using the internal lane sizing standard and local Southern sizing algorithm in the GeneScan analysis software, ver. 2.1 (ABI 1996b). Scoring of alleles for each locus and tabulation of data for importing into statistical software was performed with Genotyper software, ver. 1.1 (ABI 1994).

Statistical analysis

The number of alleles and allelic range were computed for each locus and population. Tests for conformity to Hardy-Weinberg equilibrium (HWE), genotypic linkage disequilibrium, and independence between populations and allele frequency were performed for all populations using the probability test in GENEPOP ver. 3.1 (Raymond and Rousset 1995). A Markov chain method was used to provide an unbiased estimate of the "exact" *P*-value (Guo and Thompson 1992) except for tests of HWE at loci with fewer than 5 alleles. Statistical significance levels (α) for the probability tests were determined using sequential Bonferroni adjustments for simultaneous tests (Rice 1989).

I compared the statistics θ and ρ_{ST} as measures of subpopulation structure in pink salmon. Rousset (1996) showed that ρ_{ST} is analogous to R_{ST} but uses the same sample weighting scheme as θ. Each statistic was computed for all odd-year population pairs (66 pair) and all odd- x even-year population pairs (12 pair) using GENEPOP. A Wilcoxon signed ranks test (Zar 1984, pg. 153-154) was used to determine if ρ_{ST} was significantly greater than θ for the within lineage (odd- x odd-year) and between lineage (even- x odd-year) population pairs. I computed the coefficient of variation of θ and ρ_{ST} for all between region population pairs of odd-year pink salmon to estimate the precision of each statistic.

I tested for isolation by distance between odd-year populations from each regional sample location using θ and ρ_{ST} as measures of genetic distance. Geographic distance was measured as the shortest straight line coastal route between regions on a map of coastal North America (U.S. Geological Survey 1976). Statistical significance was tested using a Spearman rank correlation coefficient implemented in the "Mantel" test module in GENEPOP, and was based on 15,000 permutations of the data. The probability value of the observed data given the null hypothesis (no correlation between genetic distance and geographic distance) was calculated from the distribution of test statistics following the permutations.

I tested statistical significance of allele frequency clines at two loci (*One*μ3, μ*Sat*60). Both loci exhibited alleles at high frequencies that appeared to be geographically correlated. Such a pattern could result when migration follows a linear stepping stone model and mutation rate is much smaller than migration rate (Hartl and Clark 1989). I plotted frequency of the common allele versus geographic distance from Northwest Alaska and tested the significance of the slope of a best fit linear model using the statistical software STATISTICA (StatSoft Inc., Tulsa, OK.).

I conducted a hierarchical gene diversity analysis using an analysis of molecular variation (AMOVA) for diploid data (see Michalakis and Excoffier 1996) as implemented in ARLEQUIN ver. 1.0 (Schneider et al. 1996). A distance matrix of the number of different alleles for each pair of haplotypes was used to compute a global estimate of percent genetic variation within and between populations and θ for all twelve odd-year samples. I then tested for regional genetic structure at two levels by grouping odd-year populations and partitioning variation into within region (θ_{WR}) and between region (θ_{BR}) components. Two aggregation strategies were used for grouping populations based on geographic location. First, I grouped pairs of populations from the six sample locations to form six regional groups. Second, I grouped eight populations from the northern sample locations (Northwest Alaska, Southcentral Alaska, Southeast Alaska, North British Columbia) and four populations from the southern sample locations (South British Columbia, Northwest Washington) to form two regional groups. The second strategy emulated the genetic structure hypothesized in earlier allozyme studies (Varnavskaya and Beacham 1992; Shaklee and Varnavskaya 1994). Significance testing of the different gene diversity components followed the scheme in Excoffier et al. (1992): first, individuals are permuted among populations to obtain a null distribution of θ . Second, the regions are assumed real and individuals are permuted within each region without regard to population to obtain a null distribution of θ_{WR}. Third, the populations are assumed real and whole populations are permuted across regions to obtain a null

distribution of θ_{BR}. Fifteen thousand permutations were run to estimate the probability of having a higher value of θ, θ_{WR} and θ_{BR} than those observed by chance alone.

Results

Microsatellite genotyping

Microsatellite loci *Ots*1, μ*Sat*60 and *Ssa*197 were scored automatically using Genotyper. *One*μ3 amplification products exhibited single base "stutter" bands typical of variable adenylation (Magnuson et al. 1996); alleles at this locus were scored manually by selecting the longest fragment (size in nucleotide bases) as the "true" allele. *Ssa*85 amplification products exhibited extensive "stutter" so I developed criteria based on allelic patterns of parents and offspring from three full-sib families to manually score each individual. *Ssa*197 exhibited several unique characteristics so its results are described separately below.

*Genetic variation at microsatellite Ssa*197

Two observations distinguished *Ssa*197 from the other loci. First, I identified one tetra- and 16 tri-allelic individuals in eight populations. Such an outcome is possible under certain scenarios given the tetraploid ancestry of Pacific salmon. However, other hypothesis may also explain these results including a tandem duplication event and non-specific primer annealing. Second, I found an apparent null allele as indicated by statistically significant heterozygote deficiency in eight populations (Table 3.1) and incomplete transmission of parental alleles in two full-sib families (J. Seeb, unpublished data). Microsatellite loci exhibiting null alleles may be useful genetic markers in some contexts (see for example Paetkau and Strobeck 1995; Brookfield 1996); however, sequencing is necessary to fully verify their presence. Alternatively, primers can be relocated away from the sequence mutation. Both options were beyond the scope of this study. For this reason, and because the explanation for apparent tri- and tetra-allelism in *Ssa*197 was not resolved, I did not include it in further statistical analysis of microsatellite variation.

*Genetic variation at microsatellites One*μ3, *μSat*60, *Ots*1 *and Ssa*85

The total number of alleles per locus for all odd-year samples ranged from four (*One*μ3) to 53 (*Ssa*85) (Table 3.1). Two microsatellites, μ*Sat*60 and *One*μ3, exhibited reciprocal latitudinal trends in frequency for alleles 109 and 162 respectively (Appendix B). The 109 allele at μ*Sat*60 was most frequent in the Snake River population in Northwest Alaska (0.87), least frequent in the Gray Wolf River population in Northwest Washington (0.44) and averaged 0.66 among all odd-year samples. Conversely, the 162 allele at *One*μ3 was least frequent in the Nome River population in Northwest Alaska (0.53), most frequent in the Gray Wolf River population in Northwest Washington (0.85) and averaged 0.66 among all odd-year populations. The frequency of both alleles in the single even-year sample from Koppen Creek in Southcentral Alaska (μ*Sat*60 = 0.91 and *One*μ3 = 0.49) was most similar to the two Northwest Alaska populations.

The mean expected heterozygosity (H_E) at each locus for all odd-year pink salmon ranged from 0.44 (*One*μ3) to 0.96 (*Ssa*85) and averaged 0.68. *One*μ3 and μ*Sat*60 exhibited moderate polymorphism (H_E = 0.44 and 0.46) while *Ots*1 and *Ssa*85 were highly polymorphic (H_E = 0.86 and 0.96). The mean H_E among populations was less variable, ranging from 0.64 (Snake River) to 0.71 (Khyex River) (Table 3.1). H_E for the even-year population ranged from 0.16 (μ*Sat*60) to 0.97 (*Ssa*85) and averaged 0.63. Tests for HWE revealed two populations (Babine River and Gray Wolf River) with significant heterozygote deficiencies for *Ssa*85 (Table 3.1).

Probability tests of non-random associations between genotypes for all pairs of loci for each population resulted in only two probability values less than 0.05. Neither value was significant when the critical value (0.05) was adjusted for 78 simultaneous tests.

I found significant heterogeneity in allele frequencies among the 13 pink salmon samples ($P < 0.001$), among all odd-year samples ($P < 0.001$), and among pooled odd-year samples from each geographic region ($P < 0.001$). Allele frequencies in the pooled odd-year sample varied significantly from the single even-year sample ($P < 0.001$). I

found significant allele frequency heterogeneity among the two northwest Washington samples ($P < 0.001$) but not the other intra-regional pairs of odd-year samples.

Comparison of θ *and* ρ_{ST}

Values of ρ_{ST} and θ were generally quite low (Table 3.2). I found no significant difference between ρ_{ST} and θ (Wilcoxon signed ranks test, $P > 0.50$) for the 66 odd-year population pairs (open circles, Figure 3.2). However, ρ_{ST} was larger than θ (Wilcoxon signed ranks test, $P < 0.001$) for the cross lineage population pairs (solid squares, Figure 3.2). Global values for ρ_{ST} and θ over all loci were 0.026 and 0.022 for all odd-year samples and 0.169 and 0.032 for the pooled odd- x even-year sample. Values of ρ_{ST} and θ for each locus for all odd-year samples were 0.020 and 0.022 (*One*μ3), 0.007 and 0.019 (*Ots*1), 0.071 and 0.058 (μ*Sat*60), 0.028 and 0.007 (*Ssa*85). Estimates of the coefficient of variation (CV) for ρ_{ST} were more than twice the estimate for θ for all but the Southcentral Alaska x North British Columbia odd-year population pairs (Table 3.3).

Geographic patterns: Isolation by distance and allele frequency clines

Three Mantel tests were performed using θ and ρ_{ST} values from all odd-year population pairs. Geographic distances used for all inter-regional pairs of populations are shown in Table 3.3. The first test indicated a highly significant correlation between θ and geographic distance ($P < 0.0005$) for population pairs from all sample regions (Figure 3.3). For the second test I removed the Northwest Alaska populations and tested for correlation among the five southern regions. The probability value increased but was still significant ($P < 0.003$). The third test indicated a marginally significant correlation between ρ_{ST} and geographic distance ($P = 0.051$) for all population pairs from all regions.

A best fit linear model was used to explain the relationship between allele frequency and geographic location for microsatellites *One*μ3 (allele 162, $R^2 = 0.63$, $P < 0.002$) and μ*Sat*60 (allele 109, $R^2 = 0.74$, $P < 0.001$) (Figure 3.4). I removed the Northwest Alaska populations to test the geographic basis of allele frequency clines among the five southern regions. The probability values increased but were still significant (allele 162, $P < 0.004$; allele 109, $P = 0.034$).

Genetic population structure

The hierarchical gene diversity analysis indicated significant genetic population structure (Table 3.4). The estimate of percent genetic variation that occurred among all odd-year populations was 2.25% ($p < 0.001$), with the remainder (97.75%) occurring within populations. Both regional pooling schemes revealed significant between region (θ_{BR}) variation (two regions, $P < 0.008$; six regions, $P < 0.001$). The within region component (θ_{WR}) was significant for the two region scheme ($P < 0.001$) but not for the six region scheme ($P > 0.24$).

Discussion

Comparison of broad-scale polymorphism among loci

The wide range in polymorphism among the four dinucleotide loci in odd-year pink salmon is consistent with data from other salmonids (O'Reilly and Wright 1995) and is suggestive of widely varying mutational properties among loci. Allele length variants differ by multiples of two bp which is consistent with length change via replication slippage (of one, two or more repeats) and could favor a single or two phase mutation model (Schlotterer and Tautz 1992; DiRienzo et al. 1994). Mutation is generally considered a diversifying force enhancing variation among populations. However, if microsatellite array length is limited, a high mutation rate may act as a homogenizing factor, counteracting the diversifying effects of genetic drift (Garza et al. 1995; Nauta and Weissing 1996). Nauta and Weissing (1996) showed the rate at which allele frequency distributions converge is a function of the mutation rate, population size, time since divergence and maximum number of alleles. Using simulations, Nauta and Weissing (1996) demonstrated that measures of population divergence may be underestimated when the number of allelic states is constrained and mutation rate is high (say 10^{-3} to 10^{-4}). This relationship is likely oversimplified because the mutational properties of microsatellites are poorly understood. However, some empirical evidence may support their theoretical conclusions. Bowcock et al. (1994) showed that microsatellite loci with the highest heterozygosity (presumably due to higher mutation rates) had significantly

lower F_{ST} values in a broad-scale study of human populations. The present study, while not containing enough loci for statistical analysis, showed a similar trend. The two most polymorphic loci (*Ots*1 and *Ssa*85) exhibited the lowest values of θ. These results must be interpreted with caution though because they are based on a single statistic applied to loci with potentially different mutational properties. Loci with higher mutation rates (more polymorphic) may tend to have lower values of θ, which measures variance in allele frequency, irrespective of constraints on allele size. The fact that I saw no trend between locus heterozygosity and ρ_{ST}, that measures variance in allele size, suggests this is the case in my study. Analysis of additional loci and larger sample sizes are needed to adequately test the theoretical relationship between mutation rate and estimates of population divergence. This is especially true for teleost fishes in which microsatellite loci vary widely in heterozygosity (Brooker et al. 1994; O'Reilly and Wright, 1995) and the number of alleles often exceeds 20, well beyond the limitations imposed by Nauta and Weissing (1996).

Genetic population structure and patterns of gene flow

My comparison of ρ_{ST} and θ suggests the latter is the better index of population structure in odd-year pink salmon, regardless of geographic distance between populations. The values of the two measures were not significantly different among regions, but θ was more precise (lower coefficient of variation) presumably because relatively few new mutations have accrued within populations Slatkin (1995). Slatkin (1995) showed θ approaches ρ_{ST} as migration increases, average coalescence times within and between populations converge and mutation rate decreases. I could not estimate all three parameters from the present data. However, it was instructive to consider results of the odd-year pairs in light of results from the odd-year/even-year pairs in which no migrants are exchanged and the time since divergence may be as great as 1 million years (Brykov et al. 1996, but see also Zhivotovsky et al. 1994). The fact that ρ_{ST} was significantly greater than θ for the twelve odd-year/even-year pairs suggests the overall mutation rate is high and therefore a high migration rate and/or small difference in

average coalescence times are influencing genetic population structure in odd-year pink salmon. Further analysis should incorporate additional even-year populations and more loci to validate the relationship between ρ_{ST} and θ within and between lineages.

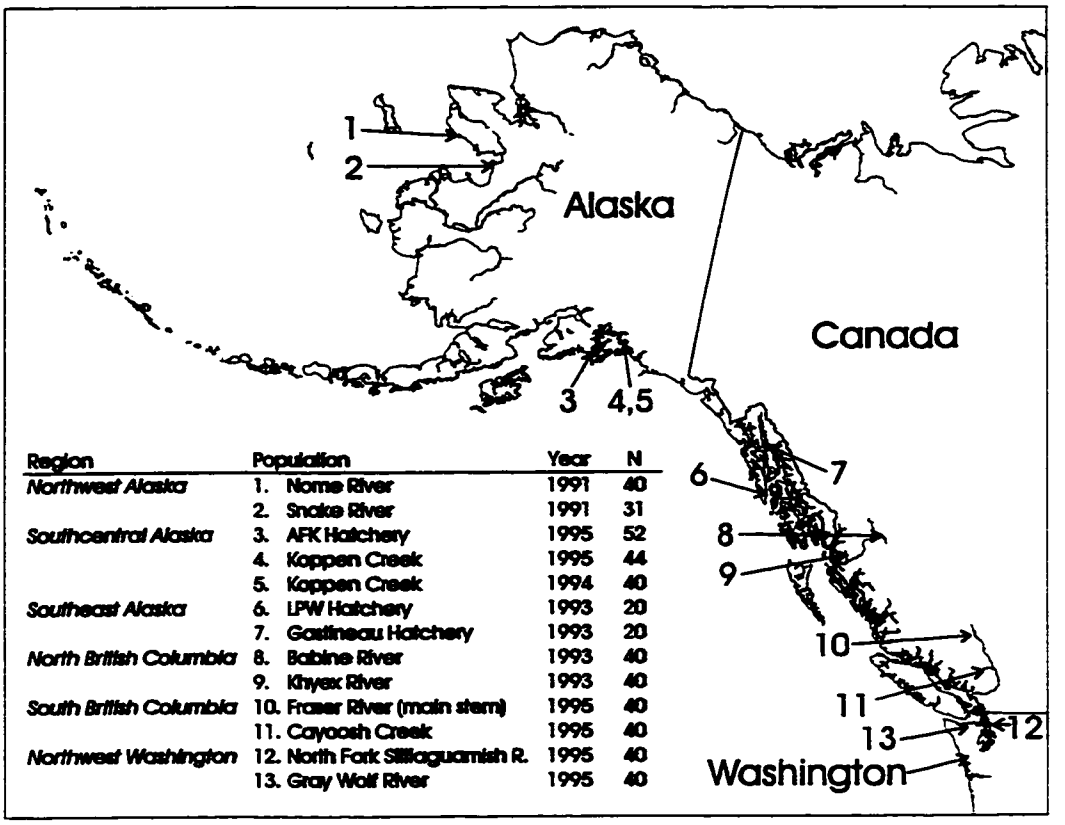
I estimated over 97.7% of the total genetic variation was common to all populations sampled. The estimate of subpopulation structure ($\theta = 0.023$, $P < 0.0001$) was low given the broad geographic range of my study. However, this values falls within the range reported for odd-year populations in earlier allozyme studies which is among the lowest for Pacific salmon (Beacham et al. 1988; Varnavskaya and Beacham 1992; Hard et al. 1996).

The geographic range to which pink salmon exhibit philopatry is a matter of debate (Varnavskaya and Beacham 1992; Omelchenko 1994; Zhivotovsky et al. 1994;). Some suggest that pink salmon are composed of unstable populations that fluctuate in time by citing the lack of regional heterogeneity in allele frequency at allozyme loci as well as tagging evidence of inter-population migration (see Omelchenko 1994 and references therein). Others support the concept of temporally stable pink salmon populations and show evidence of significant regional heterogeneity in allozyme allele frequency using a different suite of loci (Varnavskaya and Beacham 1992). The latter view is supported by recent studies of mtDNA variation in pink salmon showing little variation among rivers within regions but significant genetic heterogeneity between regions (Brykov et al. 1996; Seeb et al. 1996). I showed statistical support for genetically distinct regional spawning aggregates under two population aggregation schemes for North American odd-year pink salmon. The fact that θ_{BR} was significant for both regional aggregation schemes suggest hierarchical structure may exist beyond the two levels tested here. However, further refinement of the ARLEQUIN program and additional sampling is needed to test this. The fact that θ_{BR} was significant for the two region aggregation scheme lends some statistical support to the two putative North American population clusters in the dendrogram by Shaklee and Varnavskaya (1994). My microsatellite data suggests North American populations are temporally stable, at

least on a broad scale (~ 500-5000 km) but that regional differentiation is weak presumably due to migration. A better understanding of the dynamics of pink salmon populations will require further genetic analysis as well as additional migration studies. I did not detect significant intra-regional genetic variation for the six region pooling scheme using AMOVA. The probability test of allele frequency independence among population pairs within regions revealed only one instance of significant heterogeneity (Stillaguamish River and Gray Wolf River in Northwest Washington). Allozyme data show the Stillaguamish River and Gray Wolf River populations are part of two distinct lineages, Puget Sound and Hood Canal/Strait of Juan de Fuca, respectively (Busack and Shaklee 1995). In fact, the Puget Sound populations appear to be more closely related to Fraser River populations (South British Columbia populations in this study) than to Hood Canal/Strait of Juan de Fuca (Shaklee et al. 1991). The lack of differences in allele frequencies in any of the other intra-regional comparisons differs from other allozyme data (Beacham et al 1985; Beacham et al. 1988; Varnavskaya and Beacham 1992). This difference is best explained by the fact that my study included only two populations from each region and employed only four microsatellite loci. A more extensive analysis of microsatellite variation in a narrower geographic range using more populations and more loci may reveal significant intra-regional genetic heterogeneity.

My microsatellite data revealed two significant broad-scale trends that suggest the extent and direction of migration among North American odd-year pink salmon populations is related to distance and location. Significant allele frequency clines at two loci (Oneμ3, μSat60) suggest migration may operate under a linear stepping stone model where the allele frequency differences at the northern and southern ends of the geographic range are due to historical accident. It is possible that populations from Northwest Alaska and Northwest Washington / South British Columbia are descendent from two distinct lineages that persisted through glaciation in different refugia (Aspinwall 1974). Genetic evidence supporting northern and southern refugial populations have been reported for sockeye (Varnavskaya et al. 1994) and chum salmon (Seeb and Crane in press). Equal support for this hypothesis in odd-year pink salmon

will require additional sampling, including populations from Asia, and more loci. The significant relationship between θ and geographic distance for all odd-year population pairs provided, to my knowledge, the first statistical support for genetic isolation by distance (IBD) in pink salmon. Examples of significant correlation between genetic and geographic distance have been shown for chum (Kijima and Fujio 1982) and chinook salmon (Utter et al. 1993). In fact, Kijima and Fujio (1982) used their data to test hypothetical migration routes. My data suggests no more than the pattern of broad-scale gene flow is structured such that migration among regions increases as geographic distance decreases. In total my examination of microsatellite polymorphism provides additional support to a model of temporally stable population aggregations in odd-year pink salmon.



Region	Population	Year	N
Northwest Alaska	1. Nome River	1991	40
	2. Snake River	1991	31
Southcentral Alaska	3. AFK Hatchery	1995	52
	4. Koppen Creek	1995	44
	5. Koppen Creek	1994	40
Southeast Alaska	6. LPW Hatchery	1993	20
	7. Gastineau Hatchery	1993	20
North British Columbia	8. Babine River	1993	40
	9. Khyex River	1993	40
South British Columbia	10. Fraser River (main stem)	1995	40
	11. Cayoosh Creek	1995	40
Northwest Washington	12. North Fork Stillaguamish R.	1995	40
	13. Gray Wolf River	1995	40

Figure 3.1. Sample location, year and sample size (N) for pink salmon populations used in this study.

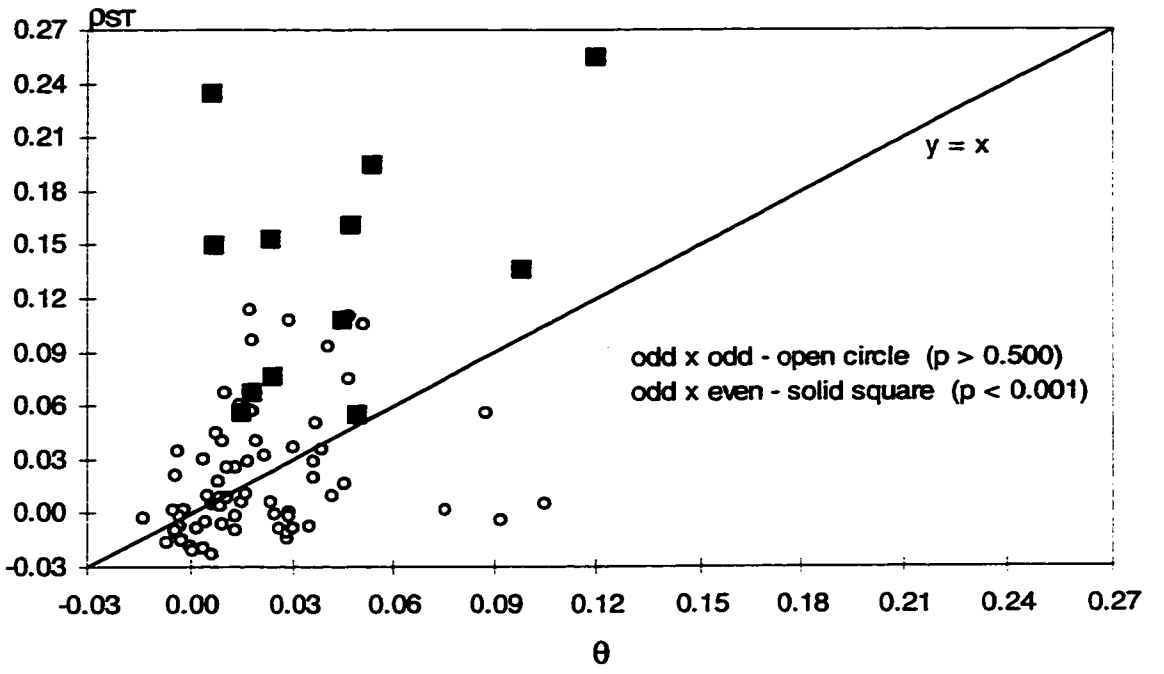


Figure 3.2. Scatter plot of ρ_{ST} versus θ for all pink salmon population pairs. Statistical significance was tested using a Wilcoxon signed ranks test.

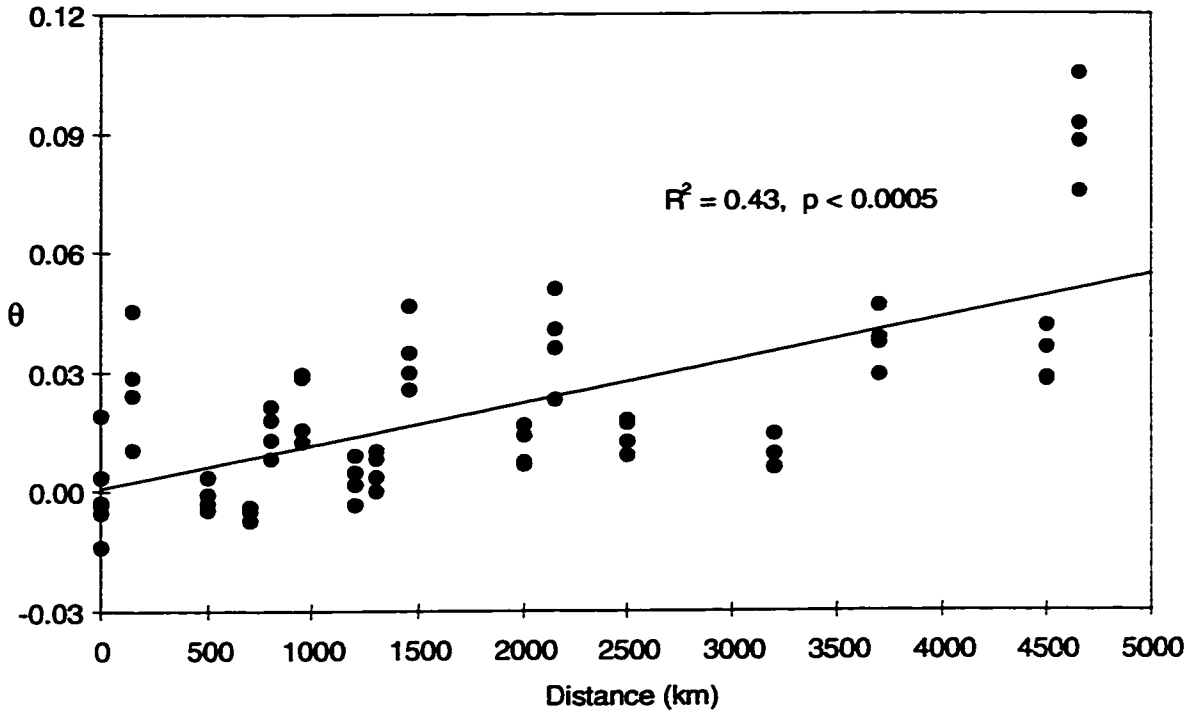


Figure 3.3. Scatter plot of θ versus geographic distance for all odd-year population pairs. Statistical significance was tested using a Spearman rank correlation coefficient implemented in the "Mantel" test module in GENEPOP, and was based on 15,000 permutations of the data.

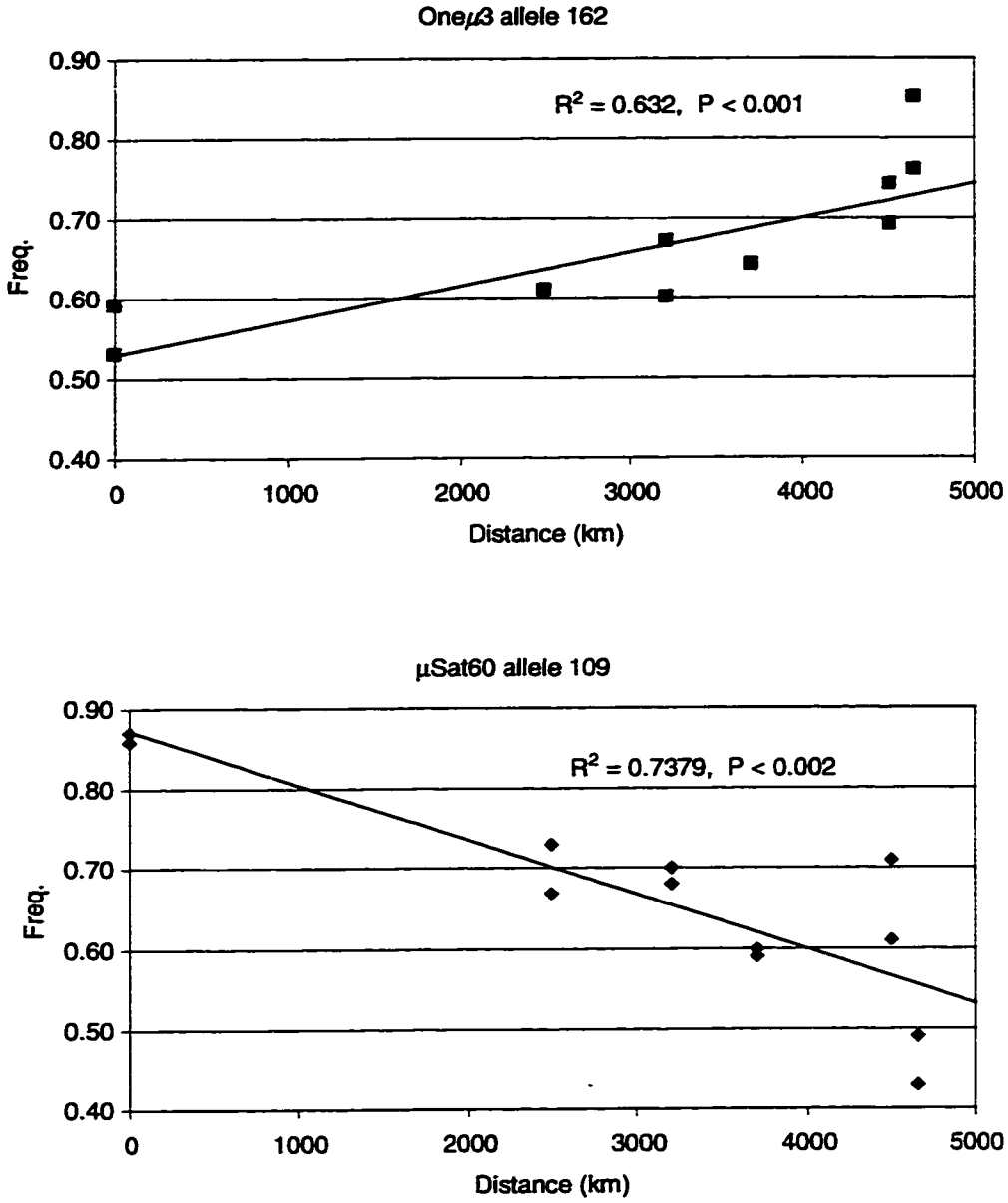


Figure 3.4. Allele frequency of *One*μ3 allele 162 and μ*Sat*60 allele 109 plotted against geographic distance from Northwest Alaska. The *P*-values were derived by testing the significance of slope of each linear regression.

Table 3.1. Expected heterozygosity (H_E) and number of alleles per locus (A) for each population-locus pair.

Region and population (sample year)[a,b]	Variable	Locus					
		*One*μ3	*Ots*1	μ*Sat*60	*Ssa*85	Avg.	*Ssa*197
Odd years							
Northwest AK							
Nome River (1991)	H_E	0.51	0.86	0.25	0.97	0.65	0.89*
	A	2	14	3	34	13	15
Snake River (1991)	H_E	0.50	0.84	0.24	0.97	0.64	0.92
	A	3	16	4	34	14	13
South-central AK							
AFK Hatchery (1995)	H_E	0.49	0.85	0.41	0.96	0.68	0.93*
	A	3	17	4	30	14	17
Koppen Creek (1995)	H_E	0.48	0.87	0.48	0.97	0.70	0.91*
	A	2	17	5	29	13	15
Southeast AK							
LPW Hatchery (1993)	H_E	0.49	0.84	0.49	0.96	0.70	0.91*
	A	2	10	4	20	9	12
Gastineau Hatchery (1993)	H_E	0.45	0.89	0.49	0.96	0.70	0.91*
	A	2	13	5	19	10	13
North BC							
Babine River (1993)	H_E	0.47	0.81	0.51	0.95*	0.69	0.90*
	A	2	8	4	21	9	14
Khyex River (1993)	H_E	0.47	0.85	0.54	0.96	0.71	0.91*
	A	2	13	4	30	12	15
South BC							
Fraser River (1995)	H_E	0.39	0.89	0.55	0.96	0.70	0.94
	A	2	13	4	28	12	17
Cayoosh Creek (1995)	H_E	0.44	0.87	0.43	0.94	0.67	0.93
	A	2	15	4	23	11	17
Northwest WA							
Stillaguamish River (1995)	H_E	0.37	0.87	0.56	0.94	0.69	0.93
	A	2	17	5	24	12	17
Gray Wolf River (1995)	H_E	0.26	0.87	0.62	0.94*	0.67	0.93*
	A	2	10	3	18	8	15
Odd-year average[c]	H_E	0.44	0.86	0.46	0.96	0.68	0.92
	A	2 (4)	14 (25)	4 (9)	26 (53)	11	15 (21)
Even year							
Southcentral AK	H_E	0.51	0.86	0.16	0.97	0.63	0.94
Koppen Creek (1994)	A	2	13	2	35	13	17

[a] Population-locus pairs with significant heterozygote deficiency based on Hardy-Weinberg equilibrium following sequential Bonferroni adjustment (initial $\alpha = 0.004$) are marked with an asterisk.

[b] AK = Alaska, BC = British Columbia, WA = Washington, AFK = Armin F. Koernig, and LPW = Little Port Walter.

[c] Total number of alleles for all odd-year samples are shown in parentheses.

Table 3.2. Pairwise ρ_{ST} (upper diagonal) and θ (lower diagonal) at four microsatellite loci in twelve odd- and one even-year pink salmon populations.

Population (sample year)	1	2	3	4	5	6	7
1. Nome River (1991)	----	0.001	0.113	0.097	0.067	0.006	0.109
2. Snake River (1991)	-0.002	----	0.040	0.026	0.005	-0.023	0.035
3. AFK Hatchery (1995)	0.017	0.009	----	-0.008	-0.010	0.034	-0.006
4. Koppen Creek (1995)	0.018	0.013	-0.004	----	-0.017	0.021	-0.009
5. LPW Hatchery (1993)	0.010	0.006	-0.005	-0.007	----	-0.002	-0.019
6. Gastineau Hatchery (1993)	0.015	0.006	-0.004	-0.005	-0.014	----	0.030
7. Babine River (1993)	0.047	0.038	0.009	0.002	-0.001	0.004	----
8. Khyex River (1993)	0.037	0.029	0.005	-0.004	-0.003	-0.005	-0.005
9. Fraser River (1995)	0.042	0.028	0.014	0.007	0.010	0.000	0.018
10. Cayoosh Creek (1995)	0.036	0.028	0.017	0.008	0.008	0.004	0.022
11. Stillaguamish River (1995)	0.088	0.075	0.036	0.023	0.030	0.026	0.016
12. Gray Wolf River (1995)	0.105	0.092	0.051	0.041	0.046	0.035	0.029
13. Koppen Creek (1994)	0.005	0.006	0.014	0.023	0.017	0.023	0.049

Table 3.2. continued

Population (sample year)	8	9	10	11	12	13
1. Nome River (1991)	0.050	0.009	0.028	0.056	0.005	0.236
2. Snake River (1991)	0.000	-0.014	-0.010	0.001	-0.004	0.151
3. AFK Hatchery (1995)	0.010	0.060	0.029	0.020	0.106	0.057
4. Koppen Creek (1995)	-0.002	0.044	0.017	0.006	0.093	0.077
5. LPW Hatchery (1993)	-0.016	0.025	0.004	-0.009	0.075	0.068
6. Gastineau Hatchery (1993)	-0.009	-0.021	-0.020	-0.009	-0.007	0.154
7. Babine River (1993)	0.002	0.057	0.032	0.011	0.107	0.056
8. Khyex River (1993)	----	0.008	-0.002	-0.010	0.036	0.109
9. Fraser River (1995)	0.008	----	-0.005	0.009	-0.002	0.196
10. Cayoosh Creek (1995)	0.013	0.004	----	-0.001	0.016	0.162
11. Stillaguamish River (1995)	0.012	0.010	0.024	----	0.040	0.137
12. Gray Wolf River (1995)	0.030	0.029	0.045	0.019	----	0.255
13. Koppen Creek (1994)	0.044	0.053	0.046	0.097	0.119	----

Table 3.3. Geographic distance and coefficient of variation (CV) of θ and ρ_{ST} for all between-region pairs of odd-year pink salmon from Alaska (AK), British Columbia (BC), and Washington (WA).

Regional pair	No. pop. pairs	Dist. (km)	CV θ	CV ρ_{ST}	CV ρ_{ST}/ CVθ
South BC x northwest WA	4	150	0.530	1.495	2.8
Southeast AK x north BC	4	500	3.349	6.715	2.0
South-central AK x southeast AK	4	700	0.271	3.501	12.9
North BC x south BC	4	800	0.386	1.098	2.8
North BC x northwest WA	4	950	0.412	1.415	3.4
South-central AK x north BC	4	1200	1.773	4.385	2.5
Southeast AK x south BC	4	1300	0.806	7.949	9.9
Southeast AK x northwest WA	4	1450	0.262	3.294	12.6
South-central AK x north BC	4	2000	0.428	0.492	1.1
South-central AK x northwest WA	4	2150	0.308	0.899	2.9
Northwest AK x south-central AK	4	2500	0.297	0.617	2.1
Northwest AK x southeast AK	4	3200	0.447	2.795	6.3
Northwest AK x north BC	4	3700	0.190	0.934	4.9
Northwest AK x south BC	4	4500	0.197	5.996	30.5
Northwest AK x northwest WA	4	4650	0.137	1.919	14.0

Table 3.4. Hierarchical gene diversity analysis of odd-year pink salmon over all loci for three population grouping strategies

Grouping strategy[a]	Source of variation[b]	σ^2	% of total	θ	θ_{BR}	θ_{WR}
None	Total	1.37299	100.00			
	Within populations	1.34207	97.75			
	Between populations	0.03092	2.25	0.0225*		
Six regions	Total	1.37609	100.00			
	Within populations	1.34207	97.53			
	Between populations	0.03402	2.47	0.0247*		
	Between regions	0.03239	2.35		0.0235*	
	Between pop. within regions	0.00163	0.12			0.0012
Two regions	Total	1.38512	100.00			
	Within populations	1.34207	96.89			
	Between populations	0.04305	3.11	0.0311*		
	Between regions	0.02447	1.77		0.0177*	
	Between pop. within regions	0.01858	1.34			0.0137*

[a] None = one group consisting of all populations; six regions = six groups consisting of population pairs from each region; two regions = two groups consisting of the eight northern populations and four southern populations.

[b] An asterisk denotes $P < 0.01$ of not greater than zero; BR = between regions; WR = within regions

CHAPTER 4: DETERMINING POPULATION IDENTITY OF INDIVIDUAL PINK SALMON IN A SUPPORTIVE BREEDING PROGRAM

Introduction

When natural populations are at risk of extinction or critical loss of genetic diversity, one management option is supportive breeding (Ryman et al. 1994). To improve survival, wild parents representing a fraction of the natural population are bred in captivity and their progeny are released, after some captive rearing, into natural habitat. Supportive breeding is often considered for restoration of Pacific salmon populations (*Oncorhynchus* spp.) because culture techniques and systems are well developed. Examples include endangered Snake River sockeye (Cummings et al. 1997) and Sacramento River winter run chinook (Hedrick et al. 1994).

If a population targeted for supportive breeding mingles with one or more other populations when founders are selected, identification error becomes a risk. Misidentification of individuals to population may result in progeny of mixed heritage, eroding genetic population structure and the existing genetic diversity and local adaptations of the target population (Waples 1994). Without a way to determine population affiliation, resource managers should consider alternatives to supportive breeding to maintain genetic diversity.

In the absence of reliable phenotypic identifiers, molecular genetic markers offer the only approach to identifying individuals to population. The most useful markers for estimating population identity are those with large allele-frequency differences among populations. Shriver et al. (1997) called such alleles "population specific alleles" (PSAs) and defined a large frequency differential as >50%. In the absence of selection the allele frequency differential varies randomly among loci and is a function of genetic drift. Thus high heterozygosity and large numbers of alleles are not necessarily indicative of usefulness for population-affiliation estimation. Marker types that tend to be less

polymorphic (e.g., allozymes) may be as useful as more variable markers (e.g. microsatellites).

Pink salmon in their native range are anadromous, semelparous and exhibit a rigid two-year life cycle that has resulted in two reproductively isolated lineages called odd- and even-year (Davidson 1934, Heard 1991). Puget Sound and the Strait of Juan de Fuca in Washington State are the southern limit of persistent spawning populations of pink salmon (Hard et al. 1996). Of the 14 recognized stocks (one even- and 13 odd-year) in this region, two are of special interest – the odd-year summer- and fall-runs of the Dungeness River in the Strait of Juan de Fuca. They represent the southwestern boundary of the species range and are widely separated from other pink salmon populations in Puget Sound (Shaklee et al. 1995). Their geographic location, unique life history, and genetic variation contribute significantly to the ecological and genetic diversity of pink salmon in this region (Hard et al. 1996). Since 1979 the fall-run has experienced a rapid decline in abundance. The number of potential breeders between 1987 and 1995 has averaged fewer than 400 individuals and twice dropped below 200 (Washington Department of Fish and Wildlife, unpublished data). In contrast the summer-run population has averaged over 6,000 potential breeders per cycle for the same period. Recent reviews have described the status of the fall-run as "critical" (Washington Department of Fisheries et al. 1993), at "moderate risk of extinction" (Nehlsen et al. 1991) and "at risk" (Hard et al. 1996). The drastic decline of Dungeness River fall pink salmon prompted resource managers to propose a supportive breeding program in 1997.

Stream survey data show the summer and fall populations are spatially isolated during spawning (Figure 1). However run timing data did not rule out the possibility of some temporal overlap during river entry, particularly in mid-August (Washington Department of Fish and Wildlife, unpublished data). Because the fall-run population spawns near the river mouth some interception of up-river migrating late summer-run fish could occur when selecting fall-run broodstock in August. A method was needed to distinguish fall from summer-run pink salmon.

This study was begun to evaluate molecular methods for selection of pink salmon broodstock. Two objectives were defined. The first was to evaluate the utility of two marker types (allozymes and microsatellites) for determining population structure and population identity of Dungeness River summer and fall pink salmon. These nuclear DNA markers were chosen because lab protocols are well described and they represent a range of polymorphism, from relatively low (allozymes) to high (microsatellites). The second objective was to use molecular markers for selecting fall-run pink salmon for broodstock from the 1997 return.

Methods

Sample collection and preparation

In 1995, tissue samples were collected from 100 summer- and 62 fall-run Dungeness River pink salmon after spawning. Selection of individuals for each population was based on location of capture (Figure 1). Samples from heart, liver, muscle, and eye were collected and stored at -80 °C for allozyme analysis by Washington Department of Fish and Wildlife personnel. Samples of fin were collected for DNA analysis and preserved in 100% ethanol.

In 1997 fin tissue for DNA analysis was collected from 171 pre-spawning adults (offspring of adults from the 1995 return) at a weir 0.6 km from the mouth (Figure 1). Sampling began August 13 and ended September 17. All fish were given a single external tag and transferred to Hurd Creek Hatchery where they were held until genotyping was complete and a determination of population identity was made. Each tag contained an alphanumeric code corresponding to the sample number assigned to each fin tissue. All fish were genotyped using microsatellites. Fish identified as fall pink salmon were retained for breeding and all others were released upstream of the weir.

Methods for preparation of tissue extracts for starch gel electrophoresis of allozyme loci were similar to those described by Aebersold et al. (1987). Total genomic DNA was isolated from 20-30 mg of fin tissue using a Gentra Systems™ (Minneapolis MN) Puregene DNA isolation kit. Precipitated DNA was hydrated in TE buffer (10 mM

Tris, 0.1 mM EDTA, pH 8.0) and heated at 55°C for approximately 12 h. The DNA concentration was measured by spectrophotometry, and then the extracts were diluted to 100 ng/μL for use in the polymerase chain reaction (PCR).

Genotyping

Twenty-five allozyme loci previously described as polymorphic in Dungeness River pink salmon were scored using procedures described by Shaklee et al. 1991. The loci were *mAAT-1**, *sAAT-3**, *sAAT-4**, *ADA-2**, *mAH-4**, *sAH**, *ALAT**, *CK-A1**, *CK-C1**, *CK-C2**, *FDHG**, *GDA**, *GPI-B2**, *G3PDH-1**, *G3PDH-2**, *sIDHP-2C**, *LDH-A1**, *LDH-B1**, *sMDH-B1,2**, *MPI**, *PEPD-2**, *PEP-LT**, *PGDH**, and *PGM-2**. Alleles were resolved using horizontal starch-gel electrophoresis and enzyme-specific histochemical staining procedures described by Shaklee et al. (1991) and Aebersold et al. (1987).

Seven microsatellite loci were scored using multiplexing procedures described by Olsen et al. 1996. The loci were *μSat*60-1, *μSat*60-2 (Estoup et al 1993): *Ots*1 (Banks et al. 1996); *Ogo*1a, *Ogo*3, *Ogo*4 (Olsen et al. 1998); *Ots*101 (Small et al. 1998). Primers for the locus *μSat*60 amplified two loci in pink salmon that differed in size by approximately 100 bp. The amplification of duplicate microsatellite loci has been reported in brown trout (Estoup et al. 1998a) and is not unexpected for salmonids given their tetraploid ancestry (Allendorf and Thorgaard 1984). The allelic range of the *μSat*60 duplicate loci did not overlap and the loci exhibited disomic inheritance, therefore each was treated as an independent segregating locus. They were distinguished based on size and designated using the numbering protocol of Shaklee et al. 1990 (e.g. *μSat*60-1, *μSat*60-2).

Microsatellites were size fractionated using an Applied Biosystems Inc. (ABI) 373A automated DNA sequencer operated in GeneScan™ mode (ABI 1993). Data were analyzed using the internal lane sizing standard and local Southern sizing algorithm in the GeneScan 672 software ver. 1.1 (ABI 1993). Scoring of alleles for each locus and

tabulation of data for importing into statistical software was performed with Genotyper software, ver. 2.0 (ABI 1996a).

Statistical analysis

Estimates of expected heterozygosity (H_E) were calculated for allozyme and microsatellite loci using equation 8.4 of Nei (1987). The average H_E was calculated as the sum of H_E across loci divided by the number of loci. Tests for conformity to Hardy-Weinberg expectation (HWE) and genotypic linkage disequilibrium analyses were performed using a probability test in GENEPOP ver. 3.1b (Raymond and Rousset 1995). Statistical significance levels (α) for the probability tests were determined using sequential Bonferroni adjustments for simultaneous tests (Rice 1989).

Unbiased estimates of Wright's F-statistics, F_{IS} F_{ST} F_{IT}, were computed using the statistics f, θ, and F (Weir and Cockerham 1984). Each statistic was computed for each locus and for all loci using FSTAT ver. 1.2 (Goudet 1995). The probability each statistic was greater than zero was tested in FSTAT by permuting the data set 10,000 times. Alleles were permuted within samples (individuals) to test f and among samples (individuals) to test θ and F.

Estimates of the allele frequency differential (δ) among fall and summer-run pink salmon were calculated for each locus as the summation of all allele frequency differences of like sign (Shriver et al. 1997).

An estimate of population identity (fall- or summer-run) was made for each individual using the computer program WHICHRUN (M. Banks, pers. com.). Likelihood estimates for run origin were evaluated by calculating expected multilocus frequencies for test sample genotypes using observed allele distributions from either fall- or summer-run samples taken in 1995. This assumes random mating, linkage equilibrium within populations and that both fall and summer are equally probable source populations. Single locus genotype frequencies were computed as p^2_i (homozygotes) and $2p_iq_i$ (heterozygotes) where p and q are allele frequencies at locus i. If an allele observed in test samples was missing from the baseline database of one or both populations a

hypothesized frequency of 1/(2n+1) was applied which assumed that the allele would be the next allele sampled. Log-likelihood ratios of multilocus genotype frequencies were computed for each individual using the formulae

$$LLR_{AFS} = \log_{10}\left(\prod_{i=1}^{i=n} Fi / \prod_{i=1}^{i=n} Si\right) \quad (1)$$

and

$$LLR_{ASF} = \log_{10}\left(\prod_{i=1}^{i=n} Si / \prod_{i=1}^{i=n} Fi\right) \quad (2)$$

(Shriver et al 1997) where n is the number of loci and Fi and Si are the genotype frequencies for individual A at locus i in the fall and summer populations respectively.

The usefulness of each marker type for population assignment was estimated by calculating the number of correct assignments at four critical values: LLR>0, LLR>1.3, LLR>2, LLR>3. Two approaches were used to establish a LLR selection criterion for fall-run broodstock in 1997. First, only individuals from the 1995 baseline file were used in the test sample. These fish were called "baseline". Second, 35 summer-run individuals were removed from the microsatellite baseline to create a true unknown test sample with genotypes not present in the baseline. These fish were called "unknown" and provided a more conservative estimate of the likelihood of miss classifying summer pink salmon as fall. The critical value was determined as the LLR above which no individuals in either test sample (baseline or unknown) were incorrectly assigned to summer- or fall-run.

Results

Population genetics

Microsatellites were much more polymorphic than allozymes in the two populations. (Table 1). The number of alleles per locus was 4-34 (mean = 23.1) for microsatellites and 2-3 (mean = 2.3) for allozymes. The mean H_E for summer and fall samples was 0.87 and 0.84 for microsatellites and 0.19 and 0.19 for allozymes. Only three allozyme loci had H_E greater than 0.5 in one or both populations.

Probability tests showed no significant departures from HWE in each population. When the populations were pooled a significant *P*-value was observed for the microsatellites ($P = 0.026$) but not for the allozymes ($P = 0.914$). Probability tests for genotypic linkage disequilibrium resulted in one significant *P*-value ($P < 0.0001$, *Ots*1 x *Ogo*4, fall-run) among 42 pairwise tests of seven microsatellite loci. No significant *P*-values were observed among 552 pairwise tests of 25 allozyme loci.

Values of θ were similar for both marker types, ranging from 0.005 to 0.076 (microsatellites) and -0.007 to 0.098 (allozymes) (Table 1). Results of the permutation tests indicated θ was significantly greater than zero at five microsatellite loci and one allozyme locus. The value of θ over all loci was 0.020 and significantly greater than zero ($P < 0.0001$) for both microsatellites and allozymes. Values of f and F over all loci were similar for both marker types and not significantly greater than zero.

Population assignment

Values of δ varied within marker types and were higher for microsatellites (0.248 - 0.522) than for allozymes (0.004-0.228) (Table 1). Probability plots of multilocus genotype frequencies showed greater delineation among populations for microsatellites than for allozymes (Figure 2). For microsatellites, the percentage of correct assignments for baseline and unknown test samples was 94.7% and 91.4% at LLR>0, 74.8% and 62.9% at LLR>1.3, 61.5% and 37.1% at LLR>2, and 31.8% and 11.4% at LLR>3. For allozymes, the percentage of correct assignments for baseline and unknown test samples was 75.9% and 62.9% at LLR>0, and 3.1% and 0.0% at LLR>1.3. No correctly assigned samples exceeded a LLR of 2 for the allozyme loci.

The effect on population assignment accuracy produced by varying the number of loci was tested at four LLR selection criteria (>0, >1.3, >2, >3) (Figure 3). Microsatellite loci were added incrementally in order of descending δ value and all allozyme loci were added at once. For LLR>0, the assignment accuracy using the most informative microsatellite (*Ogo*1a) was 0.768 and was greater than the assignment accuracy using all allozyme loci. In fact the assignment accuracy showed little increase (<3.9%) after

addition of the third microsatellite locus (Figure 3). For LLR>1.3 and >2, the increase in assignment accuracy was greatest (24.5% and 22.5%) after addition of the second microsatellite and increased minimally (0.6% and 0.5%) after addition of the sixth microsatellite. For LLR>3, the assignment accuracy increased steadily as loci were added. Addition of the allozyme loci provided noticeable increases in individual LLRs >2 and >3.

Incorrectly assigned fish were those with a LLR>0 for the wrong population. Incorrect assignments occurred in 5.3% of baseline and 8.6% of unknown test samples using microsatellites and 24.1% of baseline and 37.1% of unknown test samples using allozymes. In no case did the LLR exceed 1.3 for the wrong population. Therefore I used LLR>1.3 as my selection criterion for broodstock in 1997.

1997 Broodstock selection

One hundred seventy one pink salmon were genotyped using seven microsatellite loci and assigned to one of three categories: fall-run (76), indeterminate (73), and summer-run (22) (Figure 4). The distribution of LLR's for individuals assigned to the fall population was 23 (1.3<LLR<2), 37 (2<LLR<3), and 16 (LLR>3). Fifty-three (69.7%) pink salmon assigned to the fall population had LLR>2.

The summer-run pinks were most prevalent in the first week of sampling and declined in number over time (Figure 4). The percentages of individuals assigned respectively to summer and fall populations by week were 23.3% and 20.9% (8/13-8/19), 20.5% and 44.5% (8/20-8/26), 5.0% and 50.5% (8/27-9/02), and 6.9% and 65.5% (9/03-9/17). No individuals were assigned to the summer population after September 5.

Discussion

Marker comparison

My comparison of marker types identified two important practical considerations regarding their use for population assignment in salmon. First, estimators of mean allele frequency variation between populations such as θ are not reliable predictors of our ability to determine population origin of individuals. In the present study both marker

types provided similarly low estimates of population structure ($\theta = 0.02$). In contrast, the statistic δ, which estimates cumulative allele frequency differences and not mean variance in frequency, was larger for all microsatellite loci and a better predictor of the assignment accuracy differences between marker types. This suggests that loci with low mean variance in allele frequency can have large allele frequency differentials (δ) and be useful population affiliation markers. For populations exhibiting little genetic differentiation (such as pink salmon) loci exhibiting many alleles are preferred because δ is more likely to be large as a result of genetic drift at each allele (Shriver et al. 1997).

A second consideration is the number of loci to screen. Using my three most informative microsatellites ($\delta = 0.522, 0.422, 0.384$), I correctly assigned over 92% of baseline samples to population at LLR>0 but found little improvement with more loci. Similarly, Estoup et al. (1998a) correctly assigned over 80% (LLR>0) of individuals to population (brown trout) using the two most polymorphic loci from a seven locus panel. I found that additional loci did not increase the proportion of individuals correctly assigned to population but did increase individual LLR's, providing higher confidence in population assignments of some individuals. Shriver et al. (1997) selected five loci with the highest δ (>0.65) from a panel of 1,000 loci and, using simulation, correctly assigned 96% of individuals to ethnic group (African American versus European American) at LLR>3. My locus survey for Dungeness River pink salmon was not as extensive and only one locus (microsatellite *Ogo*1a) exhibited a δ greater than 0.50. Had I established a LLR>3 as my selection criteria then only 16 individuals would have been selected for broodstock in 1997. Clearly additional loci with high (>0.50) δ would be needed to achieve the level of confidence in population assignment used in human forensics. My initial results suggest this maybe possible but will require surveying more loci.

Although microsatellites provided greater assignment accuracy, the addition of allozymes resulted in a noticeable increase in individual LLRs >2 and >3 (Figure 3). Of the 25 allozyme loci used, 13 are detectable in skeletal muscle samples (including *PGDH* – the most informative) and therefore could presumably be reliably scored in non-lethal

biopsy samples. If surveying additional microsatellite loci is not possible then adding allozymes could be considered for future broodstock screening if there is a need to increase the selection criterion.

Population genetics and population assignment

Pink salmon exhibit low levels of intra-lineage population differentiation compared to other Pacific salmon (Allendorf and Utter 1979). Most studies indicate 3% or less of the total genetic variation at neutral markers is due to population differences in this species (e.g. Hard et al. 1996). I found significant but low (2%) inter-population genetic variation in Dungeness River pink salmon using two marker types. This level of genetic differentiation is well below that described in human races (Shriver et al. 1997 and references therein) and is likely due to recent post-glacial divergence and gene flow. Regardless, I have shown that reasonable estimates of population origin of individuals are possible using highly variable markers and that these markers can be useful tools in achieving conservation objectives. For example, I established that summer-run pink salmon are present in the lower Dungeness River after August 15. Had the selection strategy assumed all are fall-run, the breeding program would have resulted in a number of crosses between populations. Assuming the fall- to summer-run ratio (approximately 3:1) applies to the indeterminate samples, then a random single pair mating scheme would have resulted in approximately 56 % fall x fall, 38% fall x summer and 6% summer x summer families.

While my results are encouraging there are some important caveats. First, the usefulness of this tool for allocating individuals to populations will likely decrease as the number of populations increase, especially if the distance between populations decreases (Paetkau et al. 1995; Estoup et al. 1998a). Ours is a straightforward case in that, aside from low genetic differentiation, there are only two populations and they are geographically well separated from other conspecific populations. Banks et al. (1996) are developing microsatellite markers to assign individuals to four populations of Chinook salmon in the Sacramento River of which one is listed as endangered under the Endangered Species Act. In this case it is likely that δ will vary, due to random drift, at

each marker depending upon the population pair examined. In other words certain loci will distinguish individuals from some populations but not all. In such situations a more extensive screening of loci will probably be needed. Fortunately, microsatellites in cold water teleosts appear to be more polymorphic than in mammals (Brooker et al. 1994) suggesting suitable loci can be found for individual/population identification without the level of screening required for humans.

Second, other methods for establishing population identity criteria should be considered. I used a LLR threshold (LLR>1.3) determined empirically from the baseline data. This approach accounts for classification error due to genotypes of similar frequency in both populations that are present in the baseline data. However, it does not account for as yet unseen genotypes and for the sampling variance of all possible genotypes. An alternative to computing likelihood ratios is to bootstrap sample each database and compute a 95% confidence interval for the multilocus genotype frequency in each population (Weir 1995 pg. 218). This would allow for population assignments based on an estimate of genotypic sampling variance.

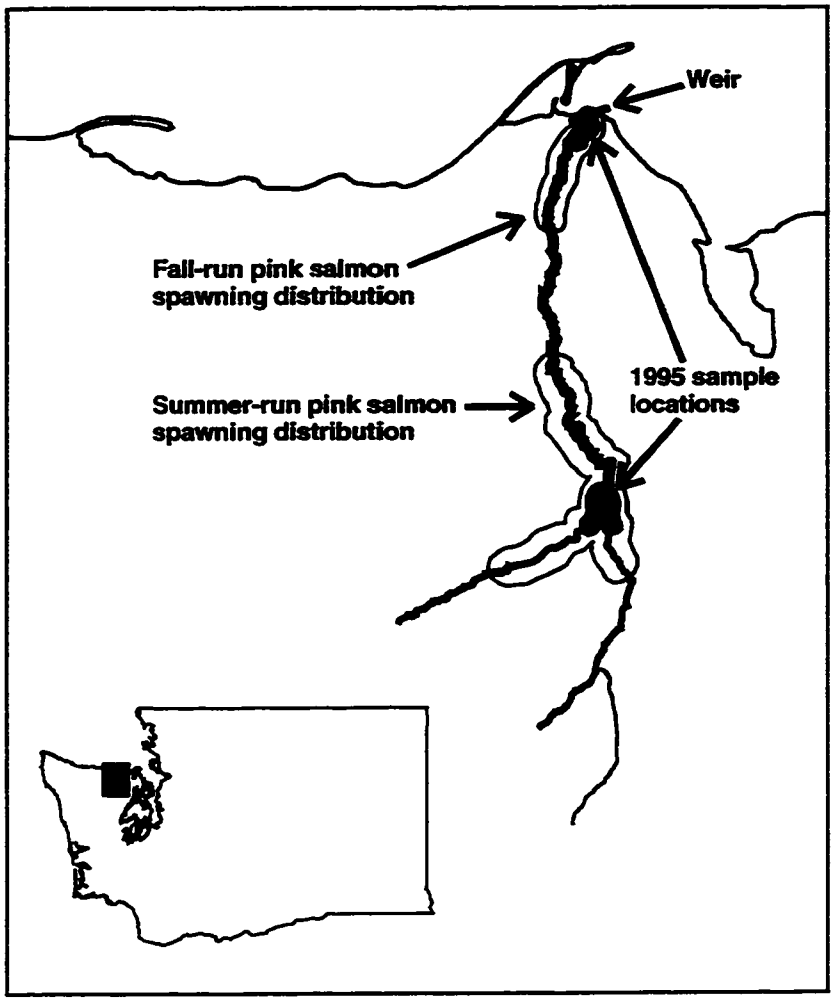


Figure 4.1. Map of the Dungeness River and Washington State (inset). Approximate spawning distribution and sample locations for summer- and fall-run pink salmon are shown.

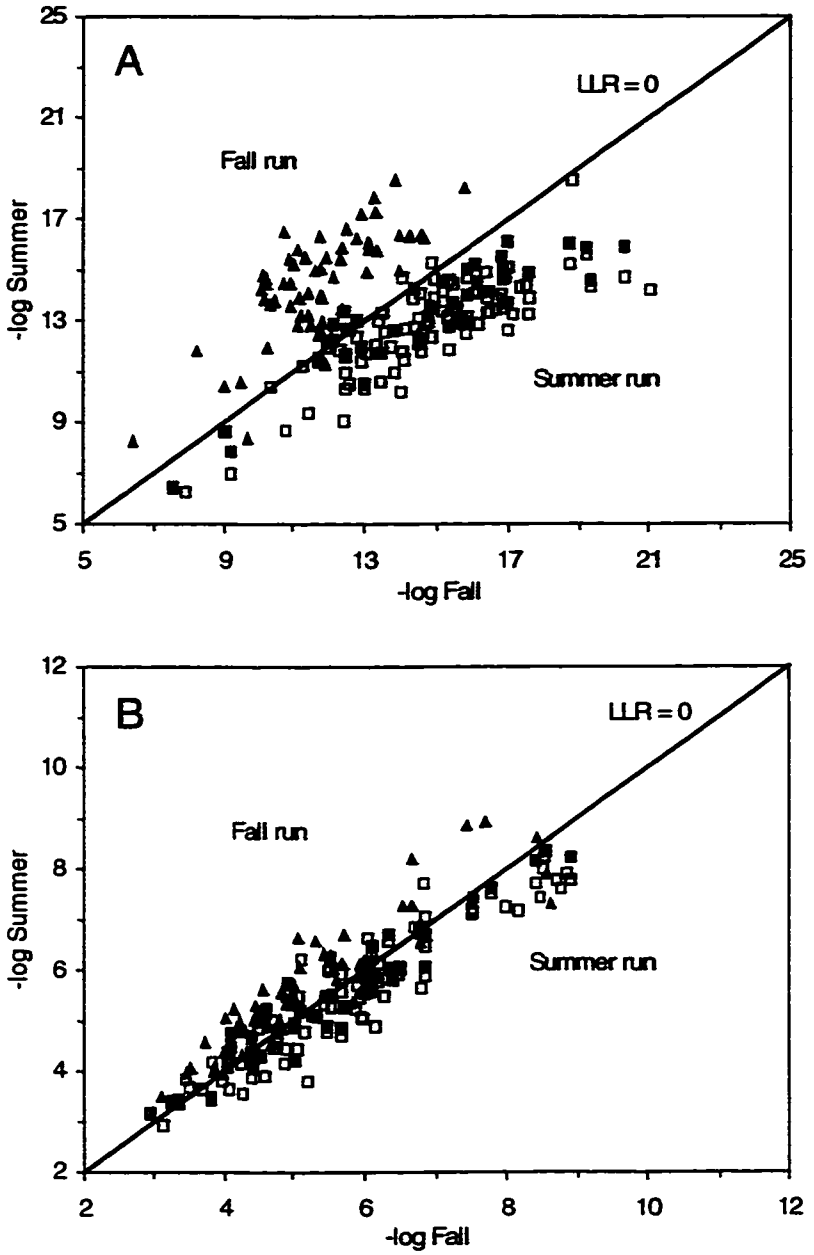


Figure 4.2. Probability plots of multilocus genotype frequencies ($-\log_{10}$) for baseline fall-run (solid triangle), baseline summer-run (open square) and unknown "summer-run" (gray square) using microsatellites (A) and allozymes (B).

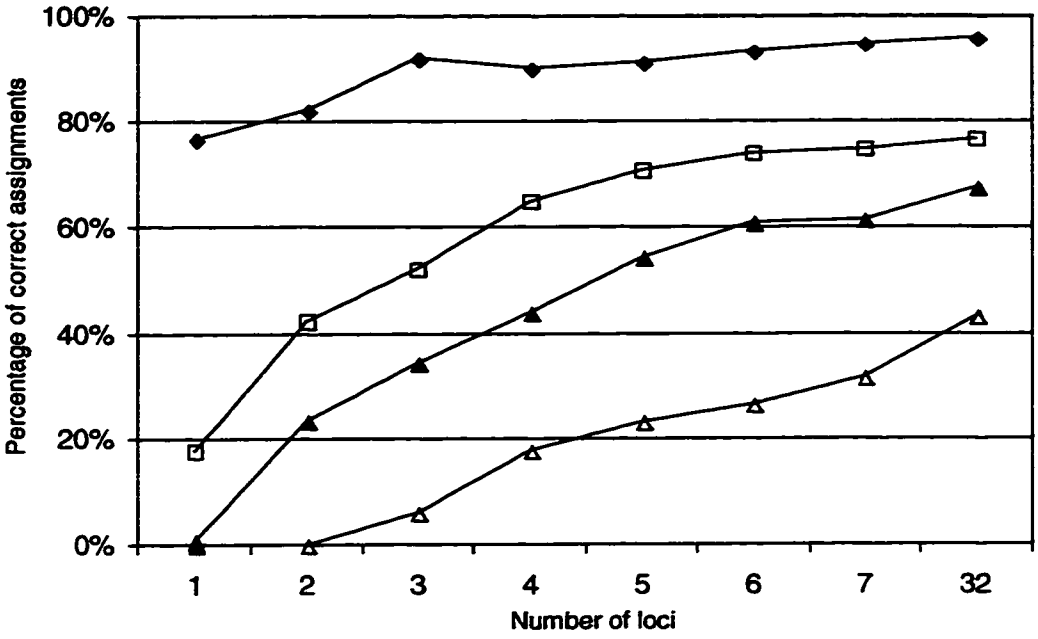


Figure 4.3. Population assignment accuracy versus number of loci for LLR>0 (solid diamond), LLR>1.3 (open square), LLR>2 (solid triangle), LLR>3 (open triangle). Microsatellite loci were added incrementally in order of descending δ and all allozyme loci were added at once.

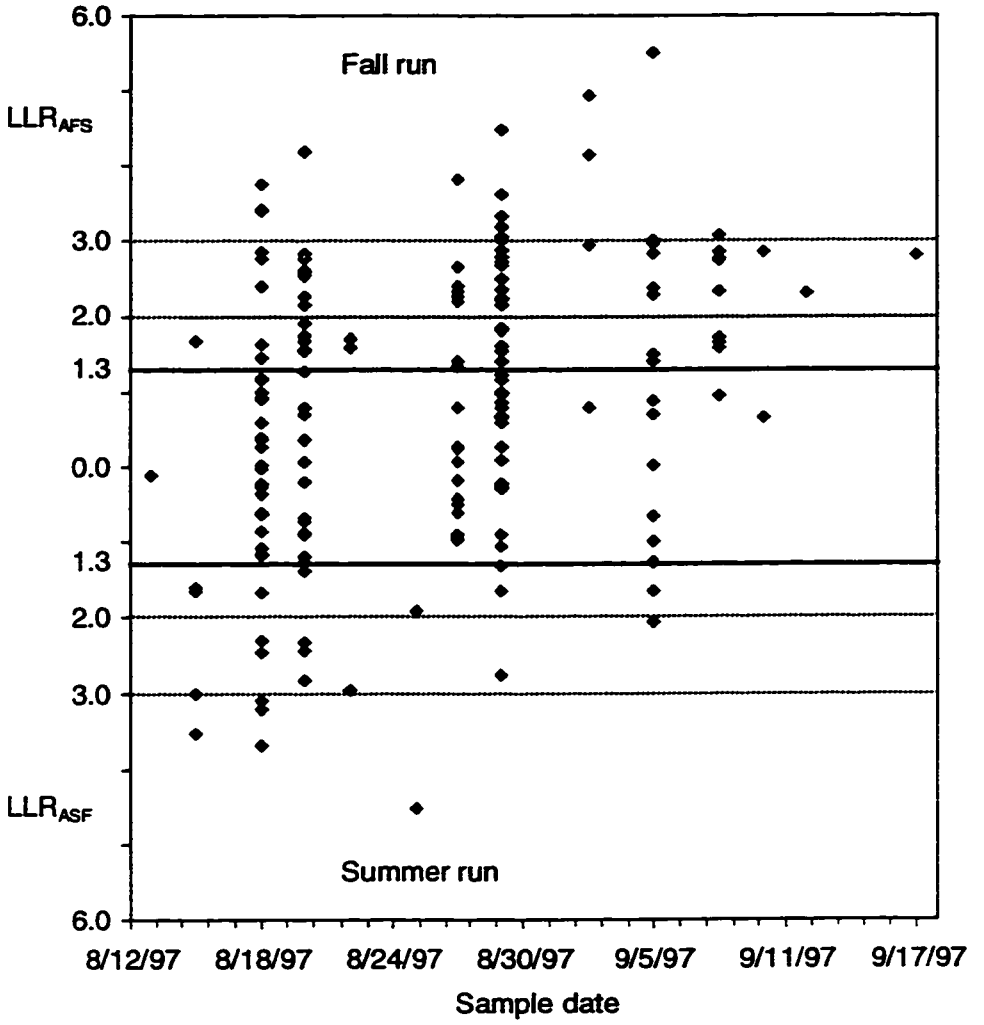


Figure 4.4. Distribution of LLR estimates for Dungeness River pink salmon in 1997. Individuals were classified as fall-run ($LLR_{AFS} > 1.3$), indeterminate ($LLR_{AFS} < 1.3$, $LLR_{ASF} < 1.3$) and summer-run ($LLR_{ASF} > 1.3$).

Table 4.1. Sample size, number of alleles per locus (A), expected heterozygosity (H_E),θ, and δ for microsatellites and allozymes. Values of θ significantly greater than zero are marked with an asterisk. *P*-values were determined using a Bonferroni adjustment for multiple tests.

	Sample size			H_E			
Locus	fall	sum	A	fall	sum	θ	δ
			Microsatellites				
*Ots*1	58	91	14	0.851	0.877	*0.011	0.248
*Ots*101	57	89	32	0.915	0.876	0.007	0.375
µ*Sat*60-1	58	93	4	0.495	0.603	*0.076	0.247
µ*Sat*60-2	57	93	23	0.834	0.912	*0.027	0.422
*Ogo*1a	57	91	34	0.932	0.961	*0.020	0.522
*Ogo*3	49	78	28	0.928	0.934	*0.013	0.384
*Ogo*4	57	91	27	0.929	0.920	0.005	0.293
avg			23	0.841	0.869	*0.020	0.356
			Allozymes				
mAAT-1	62	100	3	0.000	0.040	0.006	0.020
sAAT-3	61	100	2	0.406	0.486	0.031	0.131
sAAT-4	62	100	2	0.496	0.410	0.043	0.150
ADA-2	62	100	2	0.062	0.020	0.007	0.022
mAH-4	62	100	2	0.163	0.095	0.006	0.039
sAH	62	100	2	0.047	0.039	-0.006	0.004
ALAT	62	100	2	0.163	0.156	-0.006	0.004
CK-A1	62	100	2	0.047	0.030	-0.004	0.009
CK-C1	56	99	2	0.000	0.030	0.005	0.015
CK-C2	54	99	2	0.000	0.030	0.005	0.015
FDHG	62	100	2	0.163	0.205	-0.002	0.026
GDA	61	100	3	0.629	0.636	-0.002	0.061
GPI-B2	62	100	2	0.077	0.058	-0.005	0.010
G3PDH-1	62	100	2	0.163	0.156	-0.007	0.004
G3PDH-2	61	100	3	0.141	0.160	-0.006	0.011
sIDHP-2	62	100	3	0.529	0.483	0.004	0.086
LDH-A1	61	100	2	0.124	0.242	0.022	0.074
LDH-B1	62	100	2	0.000	0.039	0.010	0.020
sMDH-B1,2	62	100	3	0.214	0.175	-0.001	0.036
MPI	62	100	2	0.062	0.095	-0.004	0.018
PEPD-2	62	100	3	0.392	0.279	0.011	0.085
PEP-LT	62	100	2	0.000	0.039	0.010	0.020
PGDH	62	100	2	0.370	0.501	*0.098	0.228
PGM-2	62	100	2	0.238	0.131	0.020	0.067
avg			2.3	0.187	0.189	*0.020	0.048

CHAPTER 5: RECONSTRUCTING COMPLEX PEDIGREES IN CHINOOK SALMON USING MICROSATELLITES

Introduction

There is growing interest among biologists and ecologists in the use of genetic markers to reconstruct genealogical relationships in fish populations. Some applications include studying mating behavior and reproductive success (e.g. Morán et al. 1996; Foote et al. 1997; Mjølnerød et al. 1997, Thompson et al. 1998), estimating heritability of economically and ecologically important traits (e.g. Mousseau et al. 1998), comparing progeny performance and rearing strategies in hatchery populations (e.g. Ferguson and Danzmann 1998; Herbinger et al. 1999 in press), and evaluating sampling protocols for population genetic studies (e.g. Hansen et al. 1997). Highly polymorphic DNA markers, such as microsatellites, are frequently used because they have potential for high resolution of first order relationships (parent-offspring, and full sibs) (Wright and Bentzen 1994).

Accurate pedigree reconstruction using genetic markers is dependent on many factors including the number of loci genotyped, their heterozygosity, number of parents and parent pairs, and degree of relatedness of individuals in the parent population (Meagher and Thompson 1986; Chakraborty et al. 1988; Chakraborty and Jin 1993; Blouin et al. 1995). Few highly variable loci are required to resolve parent-offspring relationships if the number of parents and parent pairs is small. O'Reilly et al. (1998) used just four microsatellites (mean heterozygosity = 0.839) to assign unambiguous parentage to 791 of 792 (99.9%) offspring from 12 full sib Atlantic salmon families. The resolution obtainable with a set of genetic markers decreases, however, as pedigree complexity increases. Estoup et al. (1998b) conducted parentage analysis on two simulated mating schemes (10 x 11 and 1 x 20) with the same number of parents (21) and found the assignment success was lowest when the number of parent pairs was greatest (110). Conversely, for two mating schemes (1 x 20 and 5 x 4) with the same number of

parent pairs (20), the assignment success was lowest when the number of parents was greatest (21).

The two methods typically used to assign parentage with genetic data are exclusion and likelihood analysis. If only one candidate parent pair is genetically compatible with an offspring, then the assignment is unambiguous and based on exclusion alone. If multiple pairs are genetically compatible, then likelihood analysis is used to infer parentage (e.g. Meagher and Thompson 1987; Marshall et al. 1998; Prodöhl et al. 1998). The statistical basis for likelihood analysis is well described (e.g. Meagher and Thompson 1986; Meagher and Thompson 1987; Marshall et al. 1998), but has received little empirical support because *a priori* knowledge of true relationships is rarely available. Also, likelihood analysis may provide different results depending on whether single parent-offspring or parent pair-offspring likelihood ratios are used. Meagher and Thompson (1986) showed that the most likely single parents of each gender will often form the most likely parent pair, but empirical support is equivocal (e.g. Prodöhl et al. 1998). Finally, few studies have examined how full sibs in the candidate parent population affect parentage assignment success. In other words, how frequently are aunts and uncles incorrectly assigned as mother and father? This potential error is particularly relevant when estimating reproductive success of captive broodstock composed of groups of closely related individuals and in natural populations with large variance in family size.

I addressed these issues in the present study using both simulation and empirical data. My goals were the following: 1) to use computer simulation to determine how the number of loci, heterozygosity, and parent population size affect parentage assignment by exclusion and likelihood analysis in a random mating population; 2) to use actual microsatellite data to reconstruct a known chinook salmon pedigree and determine how close relatives in the parent population affect parentage assignment by exclusion and likelihood analysis; 3) to use actual microsatellite data from a known chinook salmon pedigree to compare single parent-offspring (SPO) and parent pair-offspring (PPO) likelihood based parentage assignments.

Methods

Sample population, tissue collection, and DNA preparation

Parentage analysis was performed on chinook salmon from a captive broodstock program that uses fish from Dungeness River in Washington State. Pre-emergent chinook salmon larvae were collected from 14 redds in 1993 to establish the broodstock (Smith and Wampler 1995). These fish were reared to maturity in freshwater hatchery tanks. The redd origin (i.e. assumed family identity) of each individual was retained using a combination of rearing isolation (juveniles) and physical tagging (adults). The first full-scale mating of these fish occurred in 1996 and comprised 102 adults (48 males and 54 females); all but one adult was obtained from the 14 redds sampled in 1993. The mating scheme consisted primarily of 3x3 factorial crosses that did not include individuals from the same redd (putative full sibs). A total of 134 crosses were made. Offspring from each cross were maintained in isolation through incubation and then tagged and placed in communal rearing tanks.

Tissue samples were collected from adults and their progeny for microsatellite genotyping. The samples from adults were taken from caudal fin tissue at the time of spawning. The samples from offspring consisted of whole fish taken from each family prior to placement in communal rearing tanks. One hundred offspring representing 18 families (3-12 progeny per family) were subsampled to evaluate parentage assignment success. The 18 families consisted of nine half sib pairs and represented the genetic contribution from all 14 redds (Table 5.1). All samples were preserved in 100% ethanol and stored in the laboratory at ambient temperature.

Total genomic DNA was isolated from 20-30 mg of tissue using procedures based on those for the Gentra Systems™ (Minneapolis MN) Puregene DNA isolation kit. Precipitated DNA was hydrated in TE buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0) and heated at 55°C for approximately 12 h. The DNA concentration was measured by spectrophotometry, and then the extracts were diluted to 100 ng/μL for use in the

polymerase chain reaction (PCR). DNA samples were prepared for all 102 adults and 100 offspring (Table 5.1).

Multilocus genotyping

Sixteen microsatellite loci were scored for each individual. PCR was carried out in 10 μL volumes (10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM each dNTP, 0.5 units *Taq* polymerase (Promega, Madison, WI), 0.05-0.35 μM each primer, and 100 ng DNA template) using a Perkin Elmer model 9600 thermo cycler. Microsatellites were amplified using the following series of thermal profiles in which X was the annealing temperature: 1) five cycles of 94°C (1 min) + X+n°C (30 s) + 72°C (15 s); 2) seven cycles of 94°C (1 min) + X°C (30 s) + 72°C (15 s); 3) 17 cycles of 94°C (30 s) + X°C (30 s) + 72°C (15 s); 4) 72°C (30 min). Profile 1 was a "touch down" PCR in which an initial annealing temperature of X+5°C decreased 1°C with each cycle. The purpose of the final 30 min incubation at 72°C was to promote complete adenylation of PCR products (Magnuson et al. 1996). Four multiplexed groups of loci were chosen using methods described by Olsen et al. (1996): two groups with an annealing temperature of 58°C, and two groups with an annealing temperature of 50°C (Table 5.2).

Microsatellites were size fractionated using an Applied Biosystems Inc. (ABI) 373A automated DNA sequencer operated in GeneScan™ mode (ABI 1993). Three lanes were required to genotype each individual because of the large allelic range of some loci. Data were analyzed using the internal lane sizing standard and local Southern sizing algorithm in the GeneScan 672 software ver. 1.1 (ABI 1993). Allele scoring and tabulation of data for importing into statistical software was performed with Genotyper software, ver. 2.0 (ABI 1996a).

Genotyping precision within and among gels was evaluated in two ways. First, one of four individuals from the adult sample was scored on every gel. If an allele was incorrectly scored at any locus for that individual, then the gel was rerun. Overall genotyping precision was measured for each locus and allele by calculating the standard deviation of fragment size estimates for each allele size category for each locus in Genotyper.

Statistical analysis

Estimates of population allele frequencies are required for parentage assignment using likelihood analysis and for unbiased estimates of relatedness using the statistic r (Queller and Goodnight 1989). These methods also assume population genotype frequencies conform to Hardy-Weinberg Expectation (HWE) for a random mating population and loci exhibit independent segregation (no linkage). In order to estimate allele frequencies and test these assumptions, a random sample of the population is needed. The adults in this study do not represent a random sample of the population because they were collected as juveniles from putative family groups (redds). In fact, collecting a population sample from chinook salmon in the Dungeness River would be problematic because live adults are rarely seen and difficult to capture if observed.

It was possible, however, to infer 12 multilocus genotypes for adults contributing to eight redds sampled in 1993. No more than 4 alleles per locus were observed in each of the eight redds – consistent with a single pair mating. An estimate of the genotype at each locus for each parent pair was made using 9-12 offspring genotypes from each redd, assuming Mendelian segregation of alleles. Thus, the probability of observing both alleles from a heterozygote parent was greater than 0.99 ($1-0.5^8$). Common genotypes were observed in four redd pairs, suggesting shared paternity or maternity. These shared genotypes were used to infer multilocus genotypes of the three parents assumed to be contributing to each redd pair. Genotype data from 27 outmigrants collected in 1993 were added to the 12 contrived genotypes to increase the population sample to 39. The outmigrating juveniles were presumed unrelated because they were sampled from various locations in the river.

The 39 genotypes were used to estimate population allele frequencies. Estimates of expected heterozygosity (H_E) were calculated for each microsatellite locus using the equation

$$H_E = 2n(1-\Sigma x_i^2)/(2n-1) \qquad (1)$$

where n is the number of individuals in the subpopulation and x_i is the frequency of the ith allele (Nei 1987, pg.178). The average H_E was calculated as the sum of H_E across loci divided by the number of loci. Tests for conformity to HWE and genotypic linkage disequilibrium analyses were performed using a probability test in GENEPOP ver. 3.1b (Raymond and Rousset 1995). Statistical significance levels (α) for the probability tests were determined using sequential Bonferroni adjustments for simultaneous tests (Rice 1989). Mendelian segregation was tested for all 16 loci in three families using a chi-square test. A minimum of 30 offspring was genotyped per family to assure that expected cell values were always greater than 5. The average exclusion probability (P_E) was estimated for each locus for a single unrelated parent-offspring pair using the computer program CERVUS (Marshall et al. 1998). The multilocus exclusion probability $P_E(C)$ was calculated using the equation

$$P_E(C) = 1 - \prod_{l=1}^{L} (1 - PE_l) \qquad (2)$$

where PE_l is the exclusion probability for the lth locus (Chakraborty et al. 1988). The relatedness statistic r, an index of kinship, was estimated for all pairs of adults and for groups of adults from the same redd using the computer program RELATEDNESS version 5.0.1 (Goodnight and Queller 1997).

Parentage assignment

The computer program PEDIGREE (C. Busack, unpublished manuscript) was used to determine how number of loci, heterozygosity, and parent population size affect parentage assignment in a simulated random mating population with no variance in reproductive success. A typical iteration began by creating a specified number of male and female k-locus genotypes (the parent population) from a random sample of a gamete pool generated by user-specified allele frequencies. Five hundred progeny genotypes were created by drawing a male and female parent at random and selecting one of two alleles at random from each locus from each parent. Each progeny was then assigned parentage from the parent population. If the true parents were the only non-excluded

pair, then the assignment was unambiguous. If multiple parent pairs were not excluded, then a parent pair-offspring (PPO) log-likelihood ratio (LOD) was computed for each non-excluded pair using the equation

$$\mathrm{LOD(QQ:UU)} = \sum_{l=1}^{L} \log_e[T_l(g_B|g_C,g_D)/P_l(g_B)] \qquad (3)$$

where QQ:UU is the probability the parent pair-offspring trio are related versus the probability they are not related, g_B is the offspring genotype, g_C and g_D are the parental genotypes, T_l is the Mendelian segregation probability for the lth locus, and P_l is the genotype probability for the lth locus (Meagher and Thompson 1986). The offspring was counted as correctly assigned if the true parent pair had the highest LOD score. One thousand iterations were completed per run. The mean and standard deviation of all unambiguous assignments (exclusion) and correct assignments (PPO likelihood analysis) were computed. The mean number of non-excluded parent pairs per individual was also computed. The assignment success for each run was defined as the mean percentage of offspring assigned to their true parent pair based on exclusion or PPO likelihood analysis. Runs were performed using three values for expected heterozygosity (0.667, 0.750. 0.875) and parent population size (50, 100, 200). Each run began with four loci, and others were added in increments of two until all assignments were unambiguous. All loci were assigned equifrequent alleles. A second set of simulations was performed using actual microsatellite frequency data from this study for 4, 6, 8, 10, 12, and 14 loci, and assuming a parent population size of 100.

Actual genetic data from 14 microsatellite loci were used to reconstruct known parent-offspring relationships in chinook salmon and determine how the presence of siblings in the parent population affect parentage assignment. Null alleles were identified for two of the 16 loci (*Oki*3a, *Ots*102), so they were excluded from the analysis. Parentage analysis was conducted using 4, 6, 8, 10, 12, and 14 loci included in descending order of H_E. Each analysis began by computing single parent-offspring

(SPO) LOD scores for all non-excluded candidate male and female parents for each of the 100 offspring using the computer program CERVUS (Marshall et al. 1998). Offspring for which only one potential male and female parent were not excluded were assigned unambiguous parentage. Offspring with more than two non-excluded potential parents were assigned parentage using two approaches. First, the male and female with the highest LOD scores were identified as the most likely parents based on SPO likelihood analysis. Second, the male and female pair with the highest LOD score were identified as the most likely parents based on PPO likelihood analysis. The second approach required testing all possible parent pairs from the list of non-excluded single parents for each undefined offspring. The computer program PROBMAX (Danzmann 1998) was used to identify non-excluded candidate parent pairs. LOD scores were computed for each possible pair for each offspring using equation 3 above.

The methods above considered all possible crosses since knowledge of the breeding pairs was not considered: common in studies of natural populations. Nevertheless, the breeding pairs in this study were known: common for most captive broodstock programs. The computer program PROBMAX (Danzmann 1998) was used to assign parentage given the limited pool of known matings (134). The results of this approach were compared to the results above that considered 2,592 (54 x 48) possible breeding pairs.

Results

Genotyping precision

The mean standard deviation of fragment size estimates in all allele size categories for each microsatellite ranged from 0.08 bases (*One*μ10) to 0.43 bases (*Ots*100) and was 0.19 bases over all loci (Table 5.2). Fragment sizing precision was highest for dinucleotide loci, with the exception of *Ocl*1. The lower sizing precision of tetranucleotide alleles did not effect genotyping accuracy because most alleles differed by four bases, allowing for non-contiguous allele categories.

Segregation analysis

Multiple chi-square tests for Mendelian segregation resulted in two significant deviations at the $\alpha = 5\%$ level (famAB3/Ots4, $P = 0.01$; famAB3/Ots1, $P = 0.01$; Table 5.3). The tests were not significant when the α level was adjusted for 44 simultaneous tests (adjusted $\alpha = 0.001$).

A small number of samples possessed aberrant phenotypes. For example, 20 offspring from family W2 lacked an allele from one or both parents at microsatellite *Oki*3a. When both parents were assumed heterozygous with a single null allele (Callen et al. 1993), a Mendelian model of inheritance was not rejected ($P = 0.40$). Similar evidence of a null allele was also found for microsatellite *Ots*102 in one of 18 families used for parentage analysis (data not shown). Five offspring possessed three alleles at one or more loci, and when they exhibited a two allele phenotype the electropherogram peak heights in Genotyper differed by a factor of about two, suggesting a three-dose genotype. These offspring, from family AB3 (4) and AA1 (1), apparently received two maternal alleles, consistent with spontaneous triploidy (e.g. Thorgaard and Gall 1979; Miller et al. 1994). Therefore these offspring were not included in the allelic segregation test.

Expected heterozygosity, exclusion probabilities, and relatedness

The number of alleles per locus ranged from 5 (*Ots*1, *Ots*3, *Ots*4) to 19 (*Ots*100) (Table 5.2). Estimates of expected heterozygosity (H_E) ranged from 0.553 (*Ots*1) to 0.946 (*Ots*100) and averaged 0.783. Probability tests of Hardy-Weinberg expectation (HWE) at each locus showed two significant deviations at the $\alpha = 5\%$ level (*Ots*104, $P = 0.02$; *One*μ8, $P = 0.01$). The tests were not significant, however, when the α level was adjusted for fourteen simultaneous tests using the sequential Bonferroni procedure (adjusted $\alpha = 0.004$). A global test of HWE across loci using Fisher's method of multiple probability tests was not significant ($P = 0.09$). Probability tests for random association of alleles between all pairs of loci showed seven significant deviations (gametic

disequilibrium) when the critical level ($\alpha = 0.05$) was adjusted for 91 simultaneous tests using the sequential Bonferroni procedure (adjusted $\alpha = 0.0005$).

The average exclusion probabilities (P_E) for each locus for a single parent offspring pair are shown in Table 5.2. Loci ranked the same according to informative value whether by P_E or H_E, with the exception of *Ocl*1 and *Ots*104. The average exclusion probability for all loci, $P_E(C)$, exceeded 0.999.

Between 3 and 12 individuals from each redd were used as broodstock in 1996 (Table 5.4). Estimates of average relatedness among individuals from the same redd ranged from 0.199 (redd 10.4) to 0.654 (redd 15.7). The average relatedness within redd for all redds was 0.446. All estimates were significantly greater than zero (the expectation if unrelated) based on the 95% confidence interval generated from a bootstrap sample of all loci.

Parentage assignment – simulation

Mean H_E had a large impact on the number of loci required to identify parentage for 95% of all progeny (Table 5.5). For example, the minimum number of loci required for at least 95% assignment success in the largest population (200 parents) increased from 6 to 20 when mean H_E decreased from 0.875 to 0.667. Parent population size had much less affect on the number of loci required for high assignment success. The decrease in assignment success that resulted from increasing the parent population size two or four fold was usually remedied by adding just two loci (Table 5.5). One additional locus was probably sufficient in some cases, but this possibility was not tested.

In many simulation scenarios over 50% of all successful assignments were attributable to PPO likelihood analysis. This relationship was observed only once, however, when the assignment success exceeded 70% (Table 5.5). In fact, when the assignment success exceeded 80% the proportion due to PPO likelihood analysis fell to between 5% and 37%. The contribution of PPO likelihood analysis declined rapidly and was no greater than about 7% when assignment success exceeded 95%. This trend is clear in Figure 5.1 for a parent population of 100. The three pairs of curves show assignments based on exclusion and exclusion plus PPO likelihood analysis for mean H_E

= 0.875, 0.750, and 0.667. For each pair, the distance between the exclusion curve and exclusion plus PPO likelihood analysis curve is greatest below 80% assignment success and declines to near zero at 95% assignment success. Also note the slope of each pair of curves declines as locus heterozygosity declines, reflecting greater informational value in incremental increases of more variable loci.

Parentage assignment – chinook salmon captive broodstock

The parentage assignment success was generally lower in the chinook salmon population than in the simulated random mating population of equal size with the same allele frequencies (Figure 5.2). Offspring in the simulated population were assigned 97% and 99% unambiguous parentage (exclusion alone) using 6 and 8 loci respectively. Only 67% of offspring in the chinook salmon population were assigned unambiguous parentage using 6 loci. The percentage of chinook salmon offspring with unambiguous parentage increased as loci were added but did not reach 95% even using 14 loci. SPO likelihood analysis did not resolve parentage for chinook salmon offspring with multiple non-excluded parent pairs (Figure 5.2). In contrast, PPO likelihood analysis increased assignment success to 76% (6 loci), 93% (8 loci), 97% (10 loci), and 99% (14 loci).

The mean number of non-excluded parent pairs (MPP) was always greater in the chinook salmon population than in the simulated population (Figure 5.3). Estimates of pairwise relatedness (r) based on 14 loci, and the redd origin data, showed that most non-excluded false parents were full sibs of true parents (Table 5.6). The mean of relatedness estimates for all true parent/false parent pairs in each chinook salmon family exceeded 0.50, the expectation for full sibs, when eight or more loci were used for parentage analysis. The mean of relatedness estimates for all true parent/false parent pairs in the chinook salmon population gradually increased as loci were added for parentage analysis, suggesting the threshold r for false inclusion of full sibs increased with number of loci genotyped (Figure 5.3). With four loci there was a noticeable drop in mean relatedness (r = 0.35), an increase in the standard deviation of r, and a steep increase in MPP (Figure 5.3). There was also a large increase in the difference between the chinook salmon population MPP and the simulation population MPP. The increase was because, at four

loci, the non-excluded false parents included individuals related and unrelated (r = 0.00) to the true parents.

The parentage assignment success varied between families (Table 5.6). Family AA1 always had more possible parent pairs than other families, including those families with a similar number of sampled progeny (AB3, W2). The mean of relatedness estimates for true parent/false parent pairs in family AA1 were always greater than 0.50 and were generally higher than in other families. This relationship indicates the AA1 parent pair shared more low frequency alleles with their full sibs than did other parent pairs.

Finally, knowledge of the breeding pairs vastly improved assignment success in the chinook salmon population by reducing the number of possible parent pairs to 134. Assignment success for 100 progeny was 95% (4 loci), 97% (6 loci), 99% (8 loci), and 100% (10 or more loci). All assignments were unambiguous and PPO likelihood analysis did not resolve parentage in the few instances where multiple parent pairs were not excluded. Although nine pairs of half sib families were sampled, in no instance were half sibs incorrectly assigned the same parent pair.

Discussion

Comparison of factors influencing pedigree reconstruction

The simulation results underscore the fact that pedigree reconstruction is highly sensitive to mean H_E. A decrease in mean H_E of about 0.1 almost doubles the number of loci required to achieve a given level of assignment success. Also, the informative value of incremental increases in number of loci noticeably decreases as mean H_E decreases. These results corroborate Blouin et al. (1996) and are of practical importance to those conducting parentage analysis for two reasons. First, the results stress the importance of investing effort in marker development and selection. Even a modest increase in locus H_E can substantially reduce genotyping, especially in large-scale studies. Second, the results provide a good guide for assessing the number of loci needed for parentage analysis. Given the availability of polymorphic DNA markers like microsatellites, it

should be possible to attain a mean H_E within the range tested here (0.667 to 0.875) in most populations. Indeed, these values of H_E comprise a range in marker variability recently reported in actual parentage and kinship studies (e.g. O'Reilly et al. 1998; Estoup et al. 1998b; Marshall et al. 1998; Blouin et al. 1996).

In contrast, moderate differences in parent population size have much less effect on the assignment success, especially when mean H_E is high (0.750, 0.875). The simulation results show that when high assignment success is sought (e.g. 95%) there is little decrease in actual assignment success for a given set of gene markers if the parent population size doubles. This finding appears to contradict Estoup et al. (1998b) who describe a decrease in assignment success for a given set of gene markers when the parent population is doubled. The difference they describe, however, is most pronounced below 90% assignment success, as is the case in this study. The difference in frequency of correct assignments between a 5 x 4 mating scheme and a 10 x 11 mating scheme is very small near 95% assignment success (Estoup et al. 1998b). Indeed, four turbot microsatellites (mean $H_E = 0.89$) provide near 95% assignment success for both mating schemes. Similarly, I assigned unambiguous parentage to 95% of chinook salmon offspring using the four most variable microsatellites (mean $H_E = 0.89$) when the number of potential parent pairs was 134 (the actual broodstock mating scheme).

A basic assumption made when estimating the exclusion probability for a set of gene markers is that individuals in the parent population are unrelated, so that the probabilities of each individual are assumed to be independent and identically distributed (Chakraborty et al. 1988). In some parentage studies, however, this assumption is unrealistic. My results for chinook salmon provide one of the first empirical tests of the influence of parental full sibs on parentage assignment. One of the interesting general observations is the difference in unambiguous assignments between the simulated population and the chinook salmon population. For example, the simulation results suggest six loci with mean $H_E = 0.87$ are sufficient to achieve, on average, 97% unambiguous assignments when the parent population size is 100. In contrast, the actual proportion of chinook salmon offspring assigned parentage using six loci is much lower:

67% unambiguous assignments, and 76% assignments based on exclusion plus PPO likelihood analysis. The loss in resolution is due, in most cases, to non-excluded parent pairs composed of full sib relatives of the true parents (Figure 5.3, Table 5.6).

This dramatic difference in assignment success is counter to the simulation results of Marshall et al. (1998) who show related males in the parent population have a minor effect on paternity assignment. The difference in the two studies is likely due to the following reasons. First, Marshall et al. 1998 use likelihood analysis to infer paternity, so only full sib males of the true father are considered as alternative fathers. In the present study, PPO likelihood analysis is conducted so full sibs of both parents, and of both genders, are considered as alternative parents. Second, Marshall et al. (1998) consider relatedness estimates only as high as 0.50 for full sib males. Figure 5.3 suggests higher estimates of relatedness should be considered to account for full sib pairs that share more alleles than average. In the present study, average relatedness within all families used as broodstock is 0.446, but is as high as 0.654 for a single family and 0.809 for a single full sib pair. Third, Marshall et al. (1998) considered only five close relatives. This total may be to low for some studies; a larger number of relatives increases the probability of very similar full sib genotypes. In the present study, 43 of 102 parents are from four different redds (10-12 individuals per redd) (Table 5.4). Finally, I found that shared parentage (half sib offspring) does not affect assignment success which does agree with the Marshall et al. (1998).

Non-zero variance in reproductive success implies that candidate parental populations will be made up of families with varying numbers of siblings. This will tend to bias the overall assignment success in the population towards the assignment success of the families most represented in the parent population (Estoup et al. 1998b). The results presented here show that parentage assignment success for progeny whose parents are part of large families will often, but not invariably, be lower than for progeny derived from smaller families. For example, 43 of the chinook salmon parents are from four redds (10 – 12 individuals per redd), and seven of the 18 families sampled here have a parent from one of those redds (redd 15.9). Two of the three families (AA1, G2) with

more than one non-excluded parent pair at 14 loci have a parent from redd 15.9. Sometimes progeny whose parents are from relatively small families will have low assignment success if the mean relatedness estimate is higher than average. The parents of the third family (AB3) with more than one non-excluded parent pair at 14 loci are from redds 10.4 and 4.2b (6 and 7 individuals per redd). The false parent in this case is a male full sib of the true male parent, and the pair have a high relatedness value ($r = 0.66$). While these results represent only a single population sample, the data emphasize the loss of resolution that can occur due to variance in reproductive success and high parental relatedness compared to expectation for an unrelated parental population.

One factor influencing pedigree analysis not addressed in my study is genotyping error. Results obtained by O'Reilly et al. (1998) suggest a microsatellite genotyping error rate of about 2-3% per allele scored; however, the methods used to score genotypes in that study differed substantially from this one. While not eliminating the risk, I believe that the semi-automated system used here greatly reduces risk of some sources of genotyping error. For example the use of multiplex PCR reduces sample handling; scoring the same individual on each gel helps identify poor quality gels; the use of in-lane sizing standards improves allele sizing precision and allows for non-contiguous allele size categories that help reveal new alleles or miss-sized alleles.

Another source of error for microsatellites is mutation. In cases where parentage is truly unknown mutation may be a concern, but empirical evidence suggests the affect of mutation is minor (O'Reilly et al. 1998). No evidence of mutation was found in the inheritance screening conducted for this study.

Comparison of exclusion and likelihood analysis for pedigree reconstruction

Results of the simulation study reveal an interesting and important aspect of pedigree reconstruction within the range of H_E examined here. That is, the proportion of parentage assignments based on PPO likelihood analysis decreases rapidly above 80% assignment success. Ignoring for the moment the impact of full sibs in the parent population, this finding has a clear implication for study design. When high assignment success is sought (e.g. 95%), the need for population allele frequency data is reduced

because PPO likelihood analysis contributes little to parentage assignment. If allele frequency data are difficult to obtain, as in the present study, then assignments may be carried out based on exclusion alone with a minimal decrease in assignment success. For example, 10 loci with mean $H_E = 0.75$ (easily attainable for many species) provide 89% unambiguous assignments and 7% correct assignments based on PPO likelihood analysis when parent population size is 100 (Table 5.5).

When moderate assignment success is sought (e.g. 80%), the simulation results suggest PPO likelihood analysis contributes a larger proportion of all assignments and the need for good population allele frequency data increases. Also, my empirical data show PPO likelihood analysis contributes a larger fraction of all assignments when the parent population consists of many full sibs. At eight and 10 loci, PPO likelihood analysis substantially increased parentage assignment success for chinook salmon offspring even though an unconventional sample was used to obtain population allele frequency data (results of tests for gametic disequilibrium suggest the derived genotypes may not reflect a random mating population).

Results of the empirical comparison of likelihood methods (SPO and PPO) for inferring parentage of chinook salmon progeny with multiple compatible parent pairs provide new insight for pedigree reconstruction. Specifically, the results suggest SPO likelihood analysis does not increase assignment success over exclusion alone when the parent population consists of many full sibs. On the other hand, PPO likelihood analysis does increase assignment success. These empirical results differ somewhat from Meagher and Thompson (1986) who show, using simulation, that the most likely single parents of each gender will often form the most likely parent pair. Other empirical support is equivocal. Prodöhl et al. (1998) show that just 51% of candidate mothers and 30% of candidate fathers have both the highest SPO and PPO LOD scores in a population of armadillos. Prodöhl et al. (1998) could not assess success of the two approaches, however, because they had no *a priori* knowledge of true relationships. Further, they present no evidence of the degree of relatedness among individuals in the parent population. My results suggest the difference in assignment success between SPO and

PPO likelihood analysis can be due to closely related individuals in the parent population. Full sibs in the parent population affect parentage assignment by SPO likelihood analysis in two ways. First, same sex full sibs sometimes have higher SPO LOD scores than the true parent, but are excluded as part of a parent pair. Second, opposite sex full sibs sometimes have higher SPO LOD scores than the true opposite sex parent, but are excluded as part of a parent pair. In the chinook salmon population, opposite sex full sibs account for 47% of the difference in assignment success between SPO and PPO likelihood analysis at eight loci.

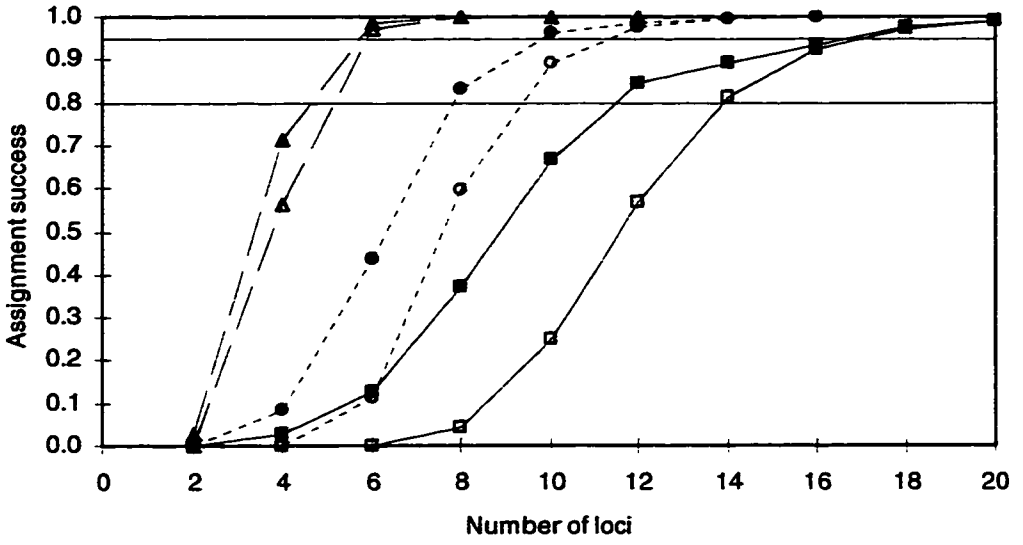


Figure 5.1. Simulation results showing relationship between parentage assignment success and number of loci genotyped. Data were generated by simulating a random mating population of 100 parents and 500 offspring using exclusion (△, □, ○) and exclusion + PPO likelihood analysis (▲, ■, ●), and assuming mean H_E of 0.667 (■, □), 0.750 (●, ○), 0.875 (▲, △). Horizontal lines indicate assignment success of 0.80 and 0.95.

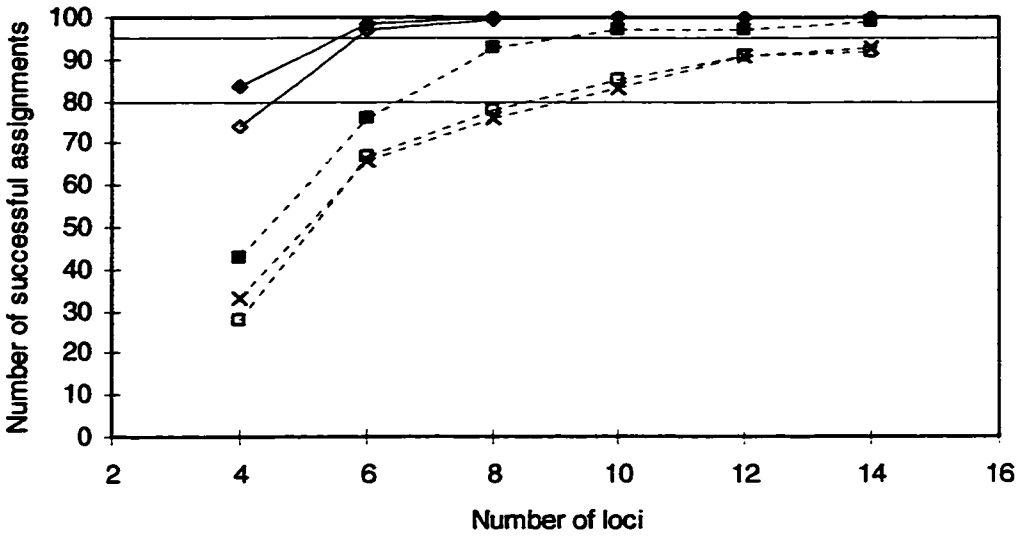


Figure 5.2. Relationship between parentage assignment success and number of loci for a simulated randomly mating population (◆,◇) and a chinook salmon captive broodstock (■, □, x). The simulation was performed using microsatellite allele frequency data from the source population used to create the captive broodstock. Mean H_E was 0.89 (4 loci), 0.87, (6 loci), 0.85 (8 loci), 0.83 (10 loci), 0.81 (12 loci), 0.78 (14 loci). Parentage analysis was conducted using exclusion (◇, □) and exclusion + PPO likelihood analysis (◆, ■) for both data sets, and SPO likelihood analysis (x) for the chinook salmon population only. Horizontal lines indicate assignment success of 0.80 and 0.95.

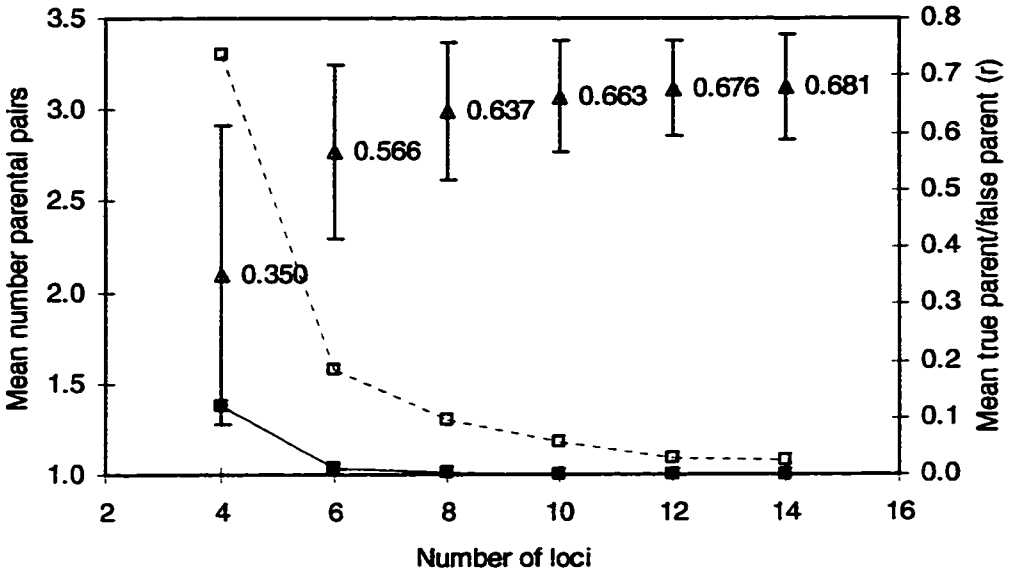


Figure 5.3. Relationship between mean number of non-excluded candidate parent pairs per offspring and number of loci typed for a simulated randomly mating population (■) and the chinook salmon population (□). The simulation was performed using microsatellite allele frequency data from the source chinook salmon population. Mean H_E was 0.89 (4 loci), 0.87, (6 loci), 0.85 (8 loci), 0.83 (10 loci), 0.81 (12 loci), 0.78 (14 loci). Also shown are mean relatedness (Δ) for all true parent/false parent pairs in the chinook salmon broodstock. Error bars denote standard deviation of the mean relatedness estimate.

Table 5.1. Eighteen chinook salmon families sampled for this study. The redd identity is provided for each parent which indicates distance (in miles) above river mouth. Redds less than 0.1 miles apart are labeled as a or b. The number of offspring typed for parentage analysis and segregation analysis are shown in columns PA and SA.

Family	female	redd	male	redd	PA	SA
AA1	F191	15.9	M183	15.7	11	32
AA2	F191	15.9	M192	10.4	5	
AB2	F194	4.2b	M192	10.4	5	
AB3	F194	4.2b	M198	17.6b	10	91
AD3	F215	6.2	M221	17.6b	4	
AE1	F216	17.4	M218	10.9	5	
AI1	F254	4.2b	M253	17.4	5	
AK1	F258	6.2	M253	17.4	4	
AM1	F292	4.2a	M281	15.9	5	
AM2	F292	4.2a	M299	9.0	5	
G1	F20	15.9	M23	10.9	4	
G2	F20	15.9	M24	9.4	3	
M1	F61	15.9	M61	4.2b	5	
M2	F61	15.9	M62	15.2	5	
T1	F98	10.9	M96	15.7	5	
T2	F98	10.9	M97	17.6a	5	
W1	F146	4.3	M152	15.7	4	
W2	F146	4.3	M159	9.4	10	33

Table 5.2. Fourteen microsatellite loci used for parentage analysis in chinook salmon. Abbreviations indicate tetranucleotide (T), dinucleotide (D), PCR annealing temperature ($T°_m$), PCR multiplex group (a, b), number of alleles (A), allele range in bases (R), mean standard deviation of fragment size estimates (in bases) in all allele categories (MSD), expected heterozygosity (H_E), and average exclusion probabilities (P_E) for a single unrelated parent-offspring pair.

Locus		Reference	$T°_m$	A	R	MSD	H_E	P_E
*Ots*100	T	Nelson et al. 1998	58a	19	214-402	0.43	0.946	0.768
*Ots*101	T	Small et al. 1998	50a	17	147-281	0.36	0.892	0.618
*Ots*2	D	Banks et al. 1999 in press	50a	10	69-105	0.12	0.870	0.565
*Ocl*1	D	Condrey and Bentzen 1998	58a	10	149-179	0.22	0.847	0.504
*Ots*104	T	Nelson et al. 1999 in press	50a	14	157-323	0.39	0.845	0.523
*One*μ8	D	Scribner et al. 1996	58a	11	157-191	0.15	0.828	0.488
*Ogo*4	D	Olsen et al. 1998	58b	8	136-184	0.13	0.796	0.411
*Omy*325	D	O'Connell (pers. comm.)	58a	8	85-145	0.14	0.767	0.380
*Ogo*2	D	Olsen et al. 1998	58b	7	210-262	0.14	0.765	0.360
*Ots*108	T	Nelson et al. 1999 in press	50b	13	100-298	0.16	0.735	0.356
*One*μ10	D	Scribner et al. 1996	50a	6	134-156	0.08	0.734	0.312
*Ots*3	D	Banks et al. 1999 in press	50a	5	85-105	0.15	0.728	0.299
*Ots*4	D	Banks et al. 1999 in press	58b	5	140-162	0.10	0.655	0.235
*Ots*1	D	Banks et al. 1999 in press	50b	5	178-196	0.15	0.553	0.152
*Ots*102[1]	T	Nelson et al. 1999 in press	50a	-	-	-	-	-
*Oki*3a[1]	T	A. Spidle, unpublished	58a	-	-	-	-	-
mean				9.86		0.19	0.783	
P_E(C)								0.999

[1] excluded from analysis because of null allele.

Table 5.3. Results of segregation analysis at 16 microsatellite loci in three chinook salmon families.

		Parental genotypes		Observed offspring genotypes										
Locus	Family	female	male	aa	ab	ac	ad	bb	bc	bd	cc	cd	n	*P*
*Oki*3a	AB3	152/223 (ad)	195/215 (bc)		23	23				28		19	93	0.63
	W2	151/null (ac)	223/null (bc)		11	6			9		5		31	0.40
	AA1	219/223 (cd)	175/195 (ab)			10	8		8	6			32	0.80
*Ots*102	AB3	223/328 (ab)	339/339 (cc)			39			54				93	0.12
	W2	201/247 (ab)	201/201 (aa)	19	14								33	0.38
	AA1	255/279 (ac)	255/263 (ab)	8	11	7			4				30	0.34
*Ots*100	AB3	270/314 (ab)	270/320 (ac)	26	31	17			18				92	0.12
	W2	226/310 (ac)	290/346 (bd)		12		10		3			6	31	0.10
	AA1	262/274 (ab)	274/274 (bb)		18			14					32	0.48
*Ots*101	AB3	183/191 (ab)	183/215 (ac)	14	30	24			24				92	0.13
	W2	259/259 (bb)	227/345 (ac)		19				14				33	0.38
	AA1	195/215 (ab)	215/215 (bb)		12			19					31	0.21
*Ots*2	AB3	69/87 (ac)	85/105 (bd)		21		25		18			27	91	0.54
	W2	87/93 (ab)	87/103 (ac)	6	7	11			9				33	0.62
	AA1	83/93 (bc)	69/83 (ab)		7	9		9	5				30	0.69
*Ocl*1	AB3	162/168 (ab)	168/174 (bc)		31	18		20	24				93	0.24
	W2	161/169 (ac)	161/163 (ab)	9	7	8			7				31	0.95
	AA1	159/163 (ac)	161/163 (bc)		5	12			8		7		32	0.35
*Ots*104	AB3	253/263 (bc)	201/201 (aa)		35	47							82	0.19
	W2	203/223 (bd)	195/219 (ac)		8		9		7			8	32	0.97
	AA1	210/221 (ab)	243/247 (cd)			8	4		11	8			31	0.36
*One*μ8	AB3	167/183 (ac)	177/177 (bb)		44				49				93	0.60
	W2	177/179 (cd)	161/173 (ab)			7	11		5	8			31	0.49
	AA1	177/177 (bb)	175/177 (ab)		19			13					32	0.29
*Ogo*4	W2	136/154 (ac)	142/162 (bd)		7		11		7			8	33	0.73
	AA1	166/170 (cd)	136/164 (ab)			9	7		5	9			30	0.69
*Omy*325	W2	91/91 (aa)	99/123 (bc)		16	15							31	0.86
*Ogo*2	AB3	224/226 (bc)	220/224 (ab)		24	26		23	18				91	0.68
	W2	224/226 (ab)	224/262 (ac)	7	7	12			8				34	0.57
	AA1	224/226 (ab)	224/224 (aa)	19	12								31	0.21
*Ots*108	AB3	108/120 (ab)	170/170 (cc)			44			46				90	0.83
	W2	158/174 (bc)	108/108 (aa)		15	17							32	0.72
	AA1	108/108 (aa)	174/178 (bc)		18	11							29	0.19
*One*μ10	AB3	138/144 (ab)	144/150 (bc)		27	26		14	26				93	0.18
	W2	138/140 (ab)	144/144 (cc)			16			18				34	0.73
	AA1	140/144 (ab)	144/150 (bc)		9	10		8	4				31	0.44
*Ots*3	AB3	87/91 (ab)	91/95 (bc)		23	30		24	14				91	0.12
	W2	91/95 (ac)	93/95 (bc)		9	7			6		11		33	0.62
	AA1	95/97 (bc)	93/93 (aa)		14	13							27	0.85
*Ots*4	AB3	148/148 (bb)	144/148 (ab)		58			33					91	0.01
	W2	144/148 (ab)	144/148 (ab)	9	13			11					33	0.42
*Ots*1	AB3	184/194 (ab)	194/194 (bb)		56			31					87	0.01
	W2	188/194 (bc)	184/184 (aa)		14	18							32	0.48
	AA1	184/184 (aa)	184/194 (ab)	12	19								31	0.21

Table 5.4. Estimates of average within redd relatedness (r) among (n) individuals sampled in 1993. The 95% confidence interval provided for each estimate was generated by bootstrap sampling of loci.

			95% C.I.	
Redd	n	r	upper	lower
17.6a	6	0.464	0.605	0.322
17.6b	6	0.499	0.658	0.339
17.4	7	0.563	0.703	0.423
15.9	11	0.449	0.520	0.377
15.7	5	0.654	0.813	0.495
15.2	7	0.339	0.530	0.148
10.9	10	0.363	0.552	0.174
10.4	3	0.199	0.393	0.004
9.4	8	0.378	0.467	0.290
9.0	5	0.456	0.643	0.269
6.2	10	0.350	0.438	0.262
4.3	4	0.311	0.455	0.167
4.2a	12	0.508	0.631	0.386
4.2b	7	0.541	0.701	0.381
ES3	1			

Table 5.5. Summary of parentage analysis for three simulated populations with (N) parents, (P) parent pairs, and 500 offspring at three values of mean locus heterozygosity (H_E), allele number (A), and allele frequency (p). Each row summarizes three simulation runs of 1000 iterations each. The mean number and standard deviation (sd) of all unambiguous assignments (E) and additional correct assignments using PPO likelihood analysis (L) are shown for each run.

	N=50, P=625					N=100, P=2,500					N=200, P=10,000				
Loci	E	sd	L	sd	Total (%)	E	sd	L	sd	Total (%)	E	sd	L	sd	Total (%)
						$H_E = 0.875$, A = 8, p = 0.125									
4	391	19.4	40	11.4	431(86)	281	17.1	76	10.4	357(71)	125	11.9	111	10.7	236(47)
6	493	4.2	3	2.4	496(99)	485	4.9	7	3.1	492(98)	468	6.5	15	4.1	482(96)
8	500	0.0	0	0.0	500(100)	499	1.4	0	0.0	499(100)	498	1.5	1	1.1	499(100)
10	500	0.0	0	0.0	500(100)	500	0.0	0	0.0	500(100)	500	0.0	0	0.0	500(100)
12	500	0.0	0	0.0	500(100)	500	0.0	0	0.0	500(100)	500	0.0	0	0.0	500(100)
						$H_E = 0.750$, A = 4, p = 0.250									
4	10	4.3	99	19.2	110(22)	0	0.7	41	8.3	41(8)	0	0.1	14	4.2	14(3)
6	185	26.2	152	15.5	338(68)	56	11.2	163	14.5	219(44)	7	2.9	104	12.1	111(22)
8	401	19.7	60	12.3	461(92)	298	21.1	119	13.8	416(83)	147	15.4	182	12.0	330(66)
10	476	6.5	15	5.1	491(98)	446	9.9	35	7.5	481(96)	384	13.3	76	10.1	459(92)
12	494	2.9	2	1.8	497(99)	488	4.0	5	2.4	493(99)	473	5.8	12	3.5	485(97)
14	499	1.4	0	0.0	499(100)	497	1.7	0	0.0	497(99)	494	2.5	1	0.8	495(99)
16	500	0.0	0	0.0	500(100)	499	1.2	0	0.0	499(100)	499	1.3	0	0.0	499(100)
						$H_E = 0.667$, A = 3, p = 0.333									
4	0	0.5	37	10.1	37(7)	0	0.0	13	4.5	13(3)	0	0.0	4	2.3	4(1)
6	14	6.4	129	18.2	142(28)	1	1.0	63	10.3	64(13)	0	0.1	25	5.6	25(5)
8	102	23.9	196	17.9	297(59)	21	7.1	164	16.3	185(37)	2	1.5	92	11.5	94(19)
10	263	32.3	149	21.5	411(82)	124	22.2	210	15.1	334(67)	30	7.5	192	13.7	222(44)
12	391	23.4	71	15.3	462(92)	285	25.0	138	16.2	422(84)	144	19.5	208	13.4	353(71)
14	458	11.9	18	5.8	476(95)	406	16.3	41	7.7	447(89)	303	19.6	84	9.7	387(77)
16	484	6.0	2	1.9	486(97)	463	9.1	5	2.6	468(94)	417	13.2	12	3.8	429(86)
18	494	3.1	0	0.0	494(99)	487	4.6	0	0.0	487(97)	469	6.9	1	0.9	470(94)
20	498	1.8	0	0.0	498(100)	495	2.6	0	0.0	495(99)	489	3.7	0	0.0	489(98)

Table 5.6. Number of offspring (n) and genetically possible parent pairs (P) in each chinook salmon family, mean of relatedness estimates (r) for all true parent/false parent pairs, number of offspring correctly assigned by exclusion (E), number of offspring correctly assigned by PPO likelihood analysis (L).

		14 loci				12 loci				10 loci				8 loci				6 loci				4 loci			
fam	n	P	r	E	L	P	r	E	L	P	r	E	L	P	r	E	L	P	r	E	L	P	r	E	L
AA1	11	4	0.69	6	5	4	0.69	6	4	7	0.67	2	8	6	0.67	2	8	12	0.60	2	3	21	0.56	0	3
AA2	5	1		5	0	1		5	0	3	0.64	3	2	3	0.64	3	2	3	0.64	3	1	11	0.21	1	0
AB2	5	1		5	0	1		5	0	1		5	0	1		5	0	1		5	0	13	0.24	0	0
AB3	10	2	0.66	8	2	2	0.66	8	2	2	0.66	8	2	3	0.56	5	5	5	0.59	5	0	13	0.30	1	1
AD3	4	1		4	0	1		4	0	1		4	0	1		4	0	1		4	0	6	0.48	1	0
AE1	5	1		5	0	1		5	0	1		5	0	1		5	0	2	0.35	3	0	13	0.17	0	0
AI1	5	1		5	0	1		5	0	1		5	0	1		5	0	2	0.29	4	1	3	0.47	0	5
AK1	4	1		4	0	1		4	0	1		4	0	1		4	0	3	0.56	3	1	6	0.47	3	0
AM1	5	1		5	0	1		5	0	1		5	0	1		5	0	1		5	0	8	0.53	3	0
AM2	5	1		5	0	1		5	0	1		5	0	4	0.63	4	0	4	0.63	4	0	6	0.57	2	1
G1	4	1		4	0	1		4	0	1		4	0	1		4	0	1		4	0	8	0.35	0	0
G2	3	2	0.67	2	0	2	0.67	2	0	2	0.67	2	0	4	0.51	1	0	4	0.51	1	0	8	0.49	0	0
M1	5	1		5	0	1		5	0	1		5	0	1		5	0	1		5	0	3	0.19	3	0
M2	5	1		5	0	1		5	0	1		5	0	1		5	0	1		5	0	3	0.39	4	0
T1	5	1		5	0	1		5	0	1		5	0	1		5	0	4	0.59	3	2	10	0.33	1	3
T2	5	1		5	0	1		5	0	1		5	0	2	0.66	4	0	3	0.62	3	0	7	0.51	2	0
W1	4	1		4	0	2	0.66	3	0	2	0.66	3	0	4	0.71	2	0	5	0.60	0	0	10	0.32	0	2
W2	10	1		10	0	1		10	0	1		10	0	1		10	0	3	0.44	8	1	4	0.29	7	0

BIBLIOGRAPHY

ABI (Applied Biosystems Inc.). 1993. GeneScan 672 Software Users Manual Revision A. Perkin-Elmer Corp., Foster City, Ca.

ABI (Applied Biosystems Inc.). 1994. Genotyper 1.1 DNA fragment analysis software. Users manual Revision A. Perkin-Elmer Corp., Foster City, Ca.

ABI (Applied Biosystems Inc.). 1996a. Genotyper 2.0 Users Manual. Perkin-Elmer Corp., Foster City, Ca.

ABI (Applied Biosystems Inc.). 1996b. ABI Prism GeneScan™ Analysis Software, Version 2.1 Users Manual. Perkin-Elmer Corp., Foster City Ca.

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. NMFS 61.

Allendorf, F. W., and F. M. Utter. 1979. Population genetics. Pages 407-454 *in* W. S. Hoar, and D. J. Randall, editors. Fish Physiology. Vol. VIII, Academic Press, N.Y.

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

Allendorf, F.W., Ryman, N., and Utter, F. 1987. Genetics and fishery management: past, present, and future. Pages 1-20 *in* N. Ryman, and F. Utter, editors. Population genetics and fishery management. Seattle, Washington: University of Washington Press.

Angers, B., L. Bernatchez, A. Angers, and L. Desgroseillers. 1995. Specific microsatellite loci for brook char reveal strong population subdivision on a microgeographic scale. Journal of Fish Biology 47 (Supplement A):177-185

Aspinwall, N. 1974. Genetic analysis of North American populations of the pink salmon, *Oncorhynchus gorbuscha*, possible evidence for the neutral mutation-random drift hypothesis. Evolution 28:295-305.

Avise, J.C. 1994. Molecular Markers, Natural History and Evolution. Chapman and Hall, Inc.

Banks, M. A., B. A. Baldwin, and D. Hedgecock. 1996. Research on chinook salmon (*Oncorhynchus tshawytscha*) stock structure using microsatellites. Bulletin of the National Research Institute of Aquaculture 2:5-9.

Banks, M.A., M.S. Blouin, B.B. Baldwin, V.K. Rashbrook, H.A. Fitzgerald, S.M. Blankenship, and D. Hedgecock. 1999. Isolation and inheritance of novel microsatellites in chinook salmon (*Oncorhynchus tshawytscha*). Journal of Heredity (in press).

Beacham, T.D., R.E. Withler, 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.

Beacham, T.D., R.W. 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.

Bentzen, P., Harris, A.S., and Wright, J.M. 1991. Cloning of hypervariable minisatellite and simple sequence microsatellite repeats for DNA fingerprinting of important aquacultural species of salmonids and tilapia. Pages 243-262 *in* T. Burke, G. Dolf, A.J. Jeffreys, and R. Wolff, editors. DNA fingerprinting: approaches and applications. Basel, Switzerland: Birkhauser Verlag.

Bentzen, P., D.B. Morris, and J.M. Wright. 1994. Development and use of variable number tandem repeat markers for population and aquacultural genetics of salmonids. Pages 85-90 *in*: L.K. Park, P. Moran, and R.S. Waples, editors. Application of DNA technology to the management of Pacific salmon; proceedings of the workshop. U.S. Dep. Commer. NOAA Tech. Memo. NMFS-NWFSC-17.

Blouin, M.S., M. Parson, V. Lacaille, and S. Lotz. 1996. Use of microsatellite loci to classify individuals by relatedness. Molecular Ecology 5:393-401.

Botstein, D., R.L. White, M. Skolnick, and R.W. Davis. 1980. Construction of a genetic linkage map using restriction fragment length polymorphisms. American Journal of Human Genetics 32:314-321.

Bowcock, A.M., A. Ruiz-Linares, J. Tomfohrde, E. Minch, J.R. Kidd, and L.L. Cavalli-Sforza. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368:455-457.

Brooker, A.L., D. Cook, P. Bentzen, J.M. Wright, and R.W. Doyle. 1994. Organization of microsatellites differs between mammals and cold-water teleost fishes. Canadian Journal of Fisheries and Aquatic Sciences 51:1959-1966.

Brookfield, J.F.Y. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology 5:453-455.

Brownstein, M.J., J.D. Carpten, and J.R. Smith. 1996. Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. BioTechniques 20:1004-1010

Brykov, V.L.A., N. Polykova, L.A. Skurikhina, and A.D. Kukhlevsky. 1996. Geographical and temporal mitochondrial DNA variability in populations of pink salmon. Journal of Fish Biology 48:899-909.

Bue, B.G., S. Sharr, S.D. Moffitt and A.K. Craig. 1996. Effects of the *Exxon Valdez* oil spill on pink salmon embryos and preemergent fry. Pages 619-627 *in*: S. Rice, R. Spies, D. Wolfe, and B. Wright, editors. Proceedings of the Exxon Valdez Oil Spill Symposium. American Fisheries Society, Symposium 18, Bethesda Maryland.

Busack, C. B., and J.B. Shaklee. 1995. Genetic diversity units and major ancestral lineages of salmonid fishes in Washington. State of Washington Department of Fish and Wildlife. Tech. Report RAD 95-02.

Callen, D.F., A.D. Thompson, Y. Shen, H.A. Phillips, R.I. Richards, J.C. Mulley, and G.R. Sutherland. 1993. Incidence and origin of "null" alleles in the (AC)n microsatellite markers. American Journal of Human Genetics 52:922-927.

Carvalho, G.R., and L. Hauser. 1994. Molecular genetics and the stock concept in fisheries. Reviews in Fish Biology and Fisheries 4(3):326-350.

Chakraborty, R., T.R. Meagher, and P. Smouse. 1988. Parentage analysis with genetic markers in natural populations. 1. The expected proportion of offspring with unambiguous paternity. Genetics 118:527-536.

Chakraborty, R., and L. Jin. 1993. Determination of relatedness between individuals using DNA fingerprinting. Human Biology 65(6):875-895.

Condrey, M.J., and P. Bentzen. 1998. Characterization of coastal cutthroat trout (*Oncorhynchus clarki clarki*) microsatellites and their conservation in other salmonids. Molecular Ecology 7(6):787-788.

Cummings, S.A., E. L. Brannon, K. J. Adams, and G. H. Thorgaard. 1997. Genetic analyses to establish captive breeding priorities for endangered Snake River sockeye salmon. Conservation Biology 11:662-669.

Danzmann, R.G. 1997. PROBMAX: A computer program for assigning unknown parentage in pedigree analysis from known genotypic pools of parents and progeny. Journal of Heredity 88(4):333.

Davidson, F.A. 1934. The homing instinct and age at maturity of pink salmon (*Oncorhynchus gorbuscha*). U.S. Bureau of Fisheries Bulletin 48:27-29.

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 of the USA 91:3166-3170.

Edwards, A., H.A. Hammond, L. Jin, C.T. Caskey, and R. Chakraborty. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12:241-253.

Ellegren, H. 1995. Mutation rates at porcine microsatellite loci. Mammalian Genome 6:376-377.

Estoup, A., P. Presa, F. Kreig, 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.

Estoup, A., C. Tailliez, J.M. Cornuet, and M. Solignac. 1995. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifer* and *Bombus terrestris* (Apidae). Molecular Biology and Evolution 12(6):1074-1084.

Estoup, A., F. Rousset, Y. Michalakis, J. M. Cornuet, M. Adriamanga, and R. Guyomard. 1998a. Comparative analysis of microsatellite and allozyme markers: a case study

investigating microgeographic differentiation in brown trout (*Salmo trutta*). Molecular Ecology 7:339-353.

Estoup, A., K. Gharbi, M. SanCristobal, C. Chevalet, P. Haffray, and R. Guyomard. 1998b. Parentage assignment using microsatellites in turbot (*Scophthalmus maximus*) and rainbow trout (*Oncorhynchus mykiss*) hatchery populations. Canadian Journal of Fisheries and Aquatic Sciences 55:715-725.

Excoffier, L., P.E. Smouse and J.M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Ferguson, M.M., and R.G. Danzmann. 1998. Role of genetic markers in fisheries and aquaculture: useful tools or stamp collecting? Canadian Journal of Fisheries and Aquatic Sciences 55:1553-1563.

FitzSimmons, N.N., C. Moritz, and S.S. Moore. 1995. Conservation and dynamics of microsatellite loci over 300 million years of marine turtle evolution. Molecular Biology and Evolution 12(3):432-440.

Foote, C.J., G.S. Brown, and C.S. Wood. 1997. Spawning success of males using alternative mating tactics in sockeye salmon (*Oncorhynchus nerka*). Canadian Journal of Fisheries and Aquatic Sciences 54:1785-1795

Forbes, S.H., K.L. Knudson, and F.W. Allendorf. 1994. Genetic variation in coho salmon detected by PCR amplification of growth hormone gene introns. Pages 29-34 *in*: L.K. Park, P. Moran, and R.S. Waples, editors. Application of DNA technology to the management of Pacific salmon; proceedings of the workshop, U.S. Dep. Commer. NOAA Tech. Memo. NMFS-NWFSC-17.

Forbes, S.H., J.T. Hogg, F.C. Buchanan, A.M. Crawford, and F.W. Allendorf. 1995. Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. Molecular Biology and Evolution 12(6):1106-1113.

Garza, J.C., M. Slatkin, and N.B. Freimer. 1995. Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. Molecular Biology and Evolution 12:594-603.

Gharrett, A.J., C. Smoot, and A.J. McGregor. 1988. Genetic relationship of even-year Northwestern Alaskan pink salmon. Transactions of the American Fisheries Society 117:536-545.

Goldstein, D.B., A. Ruiz-Linares, L.L. Cavalli-Sforza, M.W. Feldman. 1995. An evaluation of genetic distances for use with microsatellite loci. Genetics 139:463-471.

Goldstein, D.B., and D.D. Pollock. 1997. Launching microsatellites: A review of mutation processes and methods of phylogenetic inference. Journal of Heredity 88:335-342.

Goodnight K.F., and D.C. Queller. 1997. Relatedness version 5.0. Distributed by author. Keck Center for Computation Biology, Rice University.

Goudet, J. 1995. FSTAT version 1.2: a computer program to calculate F statistics. Journal of Heredity 86(6):485-486.

Guo, S.W., and E.A. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361-372.

Gyllensten, U., and A.C. Wilson. 1987. Mitochondrial DNA of salmonids: Inter- and intraspecific variability detected by restriction enzymes. Pages 277-299 *in* N. Ryman, and F. Utter, editors. Population genetics and fishery management. Seattle, Washington: University of Washington Press.

Hansen, M.M., E.E. Nielsen, and K.-L.D. Mensberg. 1997. The problem of sampling families rather than populations: relatedness among individuals in samples of juvenile brown trout *Salmo trutta* L. Molecular Ecology 6:469-474.

Hard, J.J, R.G. Kope, W.S. Grant, F.W. Waknitz, L.T. Parker, and R.S. Waples. 1996. Status review of pink salmon from Washington, Oregon and California. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-NWFSC-25, 131p.

Hartl, D.L. and A.G. Clark. 1989. Principles of population genetics, 2nd edition. Sinauer Associates, Inc., Sunderland, Mass.

Hauge, X.Y., and M. Litt. 1993. A study of the origin of "shadow bands" seen when typing dinucleotide repeat polymorphisms by the PCR. Human Molecular Genetics 2(4):411-415.

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. UBC Press.

Hedrick, P. W., D. Hedgecock, and S. Hamelberg. 1994. Effective population size in winter-run chinook salmon. Conservation Biology 9:615-624.

Herbinger, C.M., P.T. O'Reilly, R.W. Doyle, J.M. Wright, and F. O'Flynn. 1999. Early growth performance of Atlantic salmon full-sib families reared in single family tanks versus in mixed family tanks. Aquaculture (in press).

Hoelzel, A.R., and A. Green. 1992. Analysis of population-level variation by sequencing PCR-amplified DNA. Pages 159-186 *in* A.R. Hoelzel, editor. Molecular Genetic Analysis of Populations. A Practical Approach. New York: Oxford University Press, pp. 159-186.

Jarne, P. and J.L. Lagoda. 1996. Microsatellites, from molecules to populations and back. Trends in Ecology and Evolution 11(10):424-429.

Kijima, A. and Y. Fujio. 1982. Correlation between geographic distance and genetic distance in populations of chum salmon *Oncorhynchus keta*. Bulletin of the Japanese Society of Scientific Fisheries 48(12):1703-1709.

Levinson, G., and G.A. Gutman. 1987. Slipped-strand mispairing: A major mechanism for DNA sequence evolution. Molecular Biology and Evolution 4(3):203-221.

Louis, E.J., and E.R. Dempster. 1987. An exact test for Hardy-Weinberg and multiple alleles. Biometrics 43:805-811.

Magnuson, V.L. and eight co-authors. 1996. Substrate nucleotide-determined non-templated addition of adenine by *Taq* DNA polymerase: Implications for PCR-based genotyping and cloning. Biotechniques 21:700-709.

Marshall, T.C., J. Slate, L.E.B. Kruuk, and J.M. Pemberton. 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7:639-655.

McConnell, S.K., P. O'Reilly, L. Hamilton, J.M. Wright, and P. Bentzen. 1995a. Polymorphic microsatellite loci from Atlantic salmon (*Salmo salar*): genetic differentiation of north American and European populations. Canadian Journal of Fisheries and Aquatic Sciences 52:1863-1872.

McConnell, S.K., L. Hamilton, D. Morris, D. Cook, D. Paquet, P. Bentzen, and J.M. Wright. 1995b. Isolation of salmonid microsatellite loci and their application to the population genetics of Canadian east coast stocks Atlantic salmon. Aquaculture 137:19-30.

Meagher, T.R., and E. Thompson. 1986. The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. Theoretical Population Biology 29:87-106.

Meagher, T.R., and E. Thompson. 1987. Analysis of parentage for naturally established seedlings of *Chamaelirium luteum* (Liliaceae). Ecology 68(4):803-812.

Michalakis Y., and L. Excoffier. 1996. A generic estimation of population subdivision using distances between alleles with special reference for microsatellites. Genetics 142:1061-1064.

Miller, G.D., J.E. Seeb, B.G. Bue, and S. Sharr. 1994. Saltwater exposure at fertilization induces ploidy alterations, including mosaicism, in salmonids. Canadian Journal of Fisheries and Aquatic Sciences 51(suppl. 1):42-49

Mjølnerød, I.B., I.A. Fleming, U.H. Refseth, and K. Hindar. 1997. Mate and sperm competition during multiple-male spawnings of Atlantic salmon. Canadian Journal of Zoology 76:70-75.

Morán, P. A.M. Pendás, E. Beall, and E. García-Vázquez. 1996. Genetic assessment of the reproductive success of Atlantic salmon precocious parr by means of VNTR loci. Heredity 77:655-660.

Morris, D.B., K.R. Richard, and J.M. Wright. 1996. Microsatellites from rainbow trout (*Oncorhynchus mykiss*) and their use for genetic study of salmonids. Canadian Journal of Fisheries and Aquatic Sciences 53:120-126.

Mousseau, T.A., K. Ritland, and D.D. Heath. 1998. A novel method for estimating heritability using molecular markers. Heredity 80:218-224.

Nauta, M.J. and F.J. Weissing. 1996. Constraints on allele size at microsatellite loci: Implications for genetic differentiation. Genetics 143:1021-1032.

Nehlsen, W., J. E. Williams, and J. A. Lichatowich. 1991. Pacific salmon at the crossroads: stocks at risk from California, Oregon, Idaho, and Washington. Fisheries 16(2):4-20

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

Nelson, R.J., T. Beacham and M.P. Small. 1998. Microsatellite analysis of the population structure of a Vancouver Island sockeye salmon (*Oncorhynchus nerka*) stock complex using non-denaturing gel electrophoresis. Molecular Marine Biology and Biotechnology 7(4):312-319.

Nelson, R.J., and T.D. Beacham. 1999. Isolation and cross species amplification of microsatellites useful for study of Pacific salmon. Animal Genetics. (in press).

Nielsen, J.L., C.A. Gan, J.M. Wright, and K. Thomas. 1994. Biogeographic distributions of mitochondrial and nuclear markers for southern steelhead. Molecular Marine Biology and Biotechnology 3:281-293.

O'Reilly, P., and J.M. Wright. 1995. The evolving technology of DNA fingerprinting and its application to fisheries and aquaculture. Journal of Fish Biology 47(Supplement A):29-55.

O'Reilly, P.T., L.C. Hamilton, S.K. McConnell, and J.M. Wright. 1996. Rapid detection of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. Canadian Journal of Fisheries and Aquatic Sciences 53:2292-2298.

O'Reilly, P.T., C. Herbinger, and J.M. Wright. 1998. Analysis of parentage determination in Atlantic salmon (*Salmo salar*) using microsatellites. Animal Genetics 29:363-370.

Olsen, J. B., J. K. Wenburg, and P. Bentzen. 1996. Semiautomated multilocus genotyping of Pacific salmon (*Oncorhynchus* spp.) using microsatellites. Molecular Marine Biology and Biotechnology 5(4):259-272.

Olsen, J. B., P Bentzen, and J. E. Seeb. 1998. Characterization of seven microsatellite loci derived from pink salmon. Molecular Ecology 7(8):1087-1089.

Omelchenko, V.T. 1994. Population organization of pink salmon, *Oncorhynchus gorbuscha*: Analysis of concepts. Journal of Ichthyology 34(7):83-91.

Ostrander, E.A., G.F. Sprague Jr., and J. Rine. 1993. Identification and characterization of dinucleotide repeat (CA)n markers for genetic mapping in dog. Genomics 16:207-213.

Paetkau, D. and C. Strobeck. 1995a. The molecular basis and evolutionary history of a microsatellite null allele in bears. Molecular Ecology 4:519-520.

Paetkau, D., W. Calvert, I. Stirling, and C. Strobeck. 1995b. Microsatellite analysis of population structure in Canadian polar bears. Molecular Ecology 4:347-354.

Pepin, L., Y. Amigues, A. Lepingle, J. Berthier, A. Bensiad, and D. Vaiman. 1995. Sequence conservation of microsatellites between *Bos taurus* (cattle), and *Capra*

hircus (goat) and related species. Examples of use in parentage testing and phylogeny analysis. Heredity 74:53-61.

Petes, T.D., P.W. Greenwell, M. Dominska. 1997. Stabilization of microsatellite sequences by variant repeats in the yeast Saccharomyces cerevisiae. Genetics 146:491-498.

Prodöhl, P.A., W.J. Loughry, C.M. McDonough, W.S. Nelson, E.A. Thompson, and J.C. Avise. 1998. Genetic maternity and paternity in a local population of armadillos assessed by microsatellite DNA markers and field data. The American Naturalist. 151:7-19.

Queller, D.C., and K.F. Goodnight. 1989. Estimating relatedness using genetic markers. Evolution 43(2) 258-275.

Raymond, M. and F. Rousset. 1994. GENEPOP manual version 1.1. Laboratoire de Genetique et Environment URA CNRS 327 Institut des Sciences de l'Evolution CC 065, USTL Place E. Bataillon 34095 Montpellier cedex 05, France.

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(1):223-225.

Rousset, F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. Genetics 142:1357-1362.

Rozen S, and H. Skaletsky. 1996. Primer3 design program. Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html

Ryman, N., P. E. Jorde, and L. Laikre. 1994. Supportive breeding and variance effective population size. Conservation Biology 9:1619-1628

Sakamoto, T., N. Okamoto, Y. Nakamura, and T. Sato. 1994. Dinucleotide-repeat polymorphism in DNA of rainbow trout and its application in fisheries science. Journal of Fish Biology 44:1093-1096.

Schlotterer, C., and D. Tautz. 1992. Slippage synthesis of simple sequence DNA. Nucleic Acids Research 20:211-215.

Schneider, S., J.M. Kueffer, D. Roessli, and L. Excoffier. 1996. Arlequin: a software package for population genetics. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Schwengel, D.A., A.E. Jedlicka, E.J. Nanthakumar, J.L. Weber, and R.C. Levitt. 1994. Comparison of fluorescent-based semi-automated genotyping of multiple microsatellites with autoradiographic techniques. Genomics 22:46-54

Scribner, K.T., J.R. Gust, and R.L. Fields. 1996. Isolation and characterization of novel microsatellite loci: cross-species amplification and population genetic applications. Canadian Journal of Fisheries and Aquatic Sciences 53:685-693.

Seeb, J.E., C. Habicht, W.B. Templin, and L.W. Seeb. 1996. Genetics of populations of pink salmon inhabiting Prince William Sound. Annual Report, Restoration Project 94320D and 95320D, Alaska Department of Fish and Game, Anchorage, AK. 90pp.

Seeb, J.E., C. Habicht, J.B. Olsen, P. Bentzen, J.B. Shaklee, and L.W. Seeb. 1998. Allozyme, mtDNA, and microsatellite variants describe structure of populations of pink and sockeye salmon in Alaska. N. Pac. Anadr. Fish Comm. Bull. No. 1:300-318.

Seeb, L.W., and P. A. Crane. 1999. High genetic heterogeneity in chum salmon in Western Alaska, the contact zone between Asia and North America. Transactions of the American Fisheries Society (In press).

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.

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(Supplement 1):158-171.

Shaklee, J. B., J. Ames, and D. Hendrick. 1995. Genetic diversity units and major ancestral lineages for pink salmon and Washington. Chapter B in C. Busack and J. B. Shaklee, editors. Genetic diversity units and major ancestral lineages of salmonid fishes in Washington. State of Washington Department of Fish and Wildlife. Tech. Report RAD 95-02.

Shriver, M.D., M. W. Smith, L. Jin, A. Marcini, J. M. Akey, R. Deka, and R. E. Ferrel. 1997. Ethnic-affiliation estimation by use of population-specific DNA markers. American Journal of Human Genetics 60:957-964

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457-462.

Small, M.P., T.D. Beacham, R.E. Withler, and R. J. Nelson. 1998. Discriminating coho salmon (Oncorhynchus kisutch) populations within the Fraser River, British Columbia using microsatellite DNA markers. Molecular Ecology 7:141-155.

Smith, C.J., and P. Wampler. 1995. Dungeness river chinook salmon rebuilding project progress report 1992-1993. Northwest Fisheries Resource Bulletin. Project report series no. 3. WDFW 500 Capital Way N., Olympia, Wa. 98504-3150. 72p.

Tautz, D. 1989. Hypervariability of simple sequences as a general source of polymorphic DNA markers. Nucleic Acids Research 17:6463-6471.

Tessier, N., L. Bernatchez, P. Presa, and B. Angers. 1995. Gene diversity analysis of mitochondrial DNA, microsatellites and allozymes in landlocked Atlantic salmon. Journal of Fish Biology 47 (Supplement A):156-163

Thompson, C.E., W.R. Poole, M.A. Matthews, and A. Ferguson. 1998. Comparison, using DNA profiling, of secondary male contribution in the fertilisation of wild and

ranched Atlantic salmon (*Salmo salar*) ova. Canadian Journal of Fisheries and Aquatic Sciences 55:2011-2018.

Thorgaard, G.H., and G.A.E. Gall. 1979. Adult triploids in a rainbow trout family. Genetics 93:961-973.

U.S. Geological Survey. 1976. Subsurface temperature of North America [map]. 1:1,000,000. Washington D.C.: Dept. of Interior.

Urquhart, A., N.J. Oldroyd, C.P. Kimpton, and P. Gill. 1995. Highly discriminating heptaplex short tandem repeat PCR system for forensic identification. Biotechniques 18:116-121.

Utter, F., P. Aebersold, and G. Winans. 1987. Interpreting genetic variation detected by electrophoresis. Pages 21-46 *in* N. Ryman, and F. Utter, editors. Population genetics and fishery management. Seattle, Washington: University of Washington Press.

Utter, F.M., J.E. Seeb, and L.W. Seeb. 1993. Complementary uses of ecological and biochemical genetic data in identifying and conserving salmon populations. Fisheries Research 18:59-76.

Utter, F.M. 1994. Perspectives of molecular genetics and fisheries into the 21st century. Reviews in Fish Biology and Fisheries 4(3):374-378.

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.

Varnavskaya, N.V., C.C. Wood, and R.J. Everett. 1994. Genetic variation in sockeye salmon (*Oncorhynchus nerka*) populations of Asia and North America. Canadian Journal of Fisheries and Aquatic Sciences 51:132-146.

Waples, R. S. 1994. Genetic considerations in recovery efforts for Pacific salmon. Conservation Biology 8:884-886

Washington Department of Fisheries, Washington Department of Wildlife, and Western Washington Treaty Tribes. 1993. 1992 Washington State salmon and steelhead stock inventory (SASSI). Wash. Dep. Fish Wildl., Olympia, 212 p. + 5 regional vols. (Available from Washington Department of Fish and Wildlife, P.O. Box 43151, Olympia, WA 98501) 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. NMFS 61.

Weber, J.L., and C. Wong. 1993. Mutation of human short tandem repeats. Human Molecular Genetics 2(8):1123-1128.

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

Weir, B. S. 1995. Genetic data analysis II. Sinauer Associates, Inc., Sunderland, Mass.

Wenburg, J., J. Olsen, and P. Bentzen. 1996. Multiplexed systems of microsatellites for genetic analysis in coastal cutthroat trout (*Oncorhynchus clarki clarki*) and steelhead (*Oncorhynchus mykiss*). Molecular Marine Biology and Biotechnology 5(4):273-283.

Wright, J.M. 1992. DNA fingerprinting in fishes. Pages 57-91 *in* P.W. Hochachka, and T.P. Mommsen, editors. Biochemistry and Molecular Biology of Fishes, Vol. 2. Elsevier Science Publishers B.V.

Wright J.M, and P. Bentzen. 1994. Microsatellites: genetic markers for the future. Reviews in Fish Biology and Fisheries 4(3):384-388.

Zar, J.H. 1984. Biostatistical analysis, 2nd edition. Prentice Hall, Englewood Cliffs, New Jersey.

Zhivotovsky, L.A., A.J. Gharrett, A.J. McGregor, M.K. Glubokovsky, and M.W. Feldman. 1994. Gene differentiation in Pacific salmon (*Oncorhynchus sp.*): Facts and models with reference to pink salmon (*O. gorbuscha*). Canadian Journal of Fisheries and Aquatic Sciences 51(Supplement 1):223-232.

Ziegle, J.S., Y. Su, K.P. Corcoran, L. Nie, P.E. Mayrand, L.B. Hoff, L.J. McBride, M.N. Kronick, and S.R. Diehl. 1992. Application of automated DNA sizing technology for genotyping microsatellite loci. Genomics 14:1026-1031.

APPENDIX A: PANEL OF 35 SALMONID MICROSATELLITE PRIMER PAIRS.

Locus	Primer Sequence F>Forward, R>Reverse	Reference	Source Species
Fgt1	F>5'-AGA-TTT-ACC-CAG-CCA-GGT-AG	Sakamoto et al. 1994	*Oncorhynchus*
	R>5'-CAT-AGT-CTG-AAC-AGG-GAC-AG		*mykiss*
Omy77	F>5'-CGT-TCT-CTA-CTG-AGT-CAT	Morris et al. 1996	"
	R>5'-GGG-TCT-TTA-AGG-CTT-CAC-TGC-A		
Omy78	F>5'-ACT-CCA-GCA-CAC-CTG-TCT-CC	Michael O'Connell, Biology	"
	R>5'-TGT-CTC-AGT-GCT-CTT-TCC-C	Dept., Dalhousie University,	
Omy87	F>5'-TCC-TGG-TCT-GGT-GCA-GG	Halifax, N.S. Canada B3H-4JI	"
	R>5'-ATT-AAC-TCC-GTT-CCA-GCC-G	personal communication	
Omy207	F>5'-ACC-CTA-GTC-ATT-CAG-TCA-GG	"	"
	R>5'-GAT-CAC-TGT-GAT-AGA-CAT-CG		
Omy293	F>5'-CAC-AGA-GTG-CGA-TCG-TGG	"	"
	R>5'-GGT-ACT-AAT-GTT-AAG-CTC-GAG		
Omy325	F>5'-TGT-GAG-ACT-GTC-AGA-TTT-TGC	"	"
	R>5'-CGG-AGT-CCG-TAT-CCT-TCC-C		
PuPuPy	F>5' -ATG-CAG-CGG-ATG-TAG-GGG-GA	Morris et al. 1996	"
	R>5'-TTA-AGT-GAA-AAG-ACG-TAA-GTC		
Oneμ1	F>5'-GTC-TTA-CCA-AAT-GTC-TTC-CTC-CT	Scribner et al. 1996	*Oncorhynchus*
	R>5'-GCC-ATT-TAG-CAT-ACG-ATT-TTA-TC		*nerka*
Oneμ2	F>5'-GGT-GCC-AAG-GTT-CAG-TTT-ATG-TT	"	"
	R>5'-CAG-GAA-TTT-ACA-GGA-CCC-AGG-TT		
Oneμ8	F>5'-AAC-ATT-CTG-GGA-TGA-CAG-GGG-TA	"	"
	R>5'-CTG-TTC-TGC-TCC-AGT-GAA-GTG-GA		
Oneμ10	F>5'-ATG-GGG-AAC-AGA-AGA-GGA-AT	"	"
	R>5'-CTG-TAG-GTG-TGA-AAT-GTA-TTT-AAA		
Oneμ11	F>5'-GTT-TGG-ATG-ACT-CAG-ATG-GGA-CT	"	"
	R>5'-TCT-ATC-TTT-CCT-GTC-AAC-TTC-CA		
Oneμ14	F>5'-AGA-AAC-ATG-AGA-ACA-GTC-TAG-GT	"	"
	R>5'-CCT-TAT-GAG-TTT-GGT-CTC-CAT-GT		
Ots1	F>5'-GGA-AAG-AGC-AGA-TGT-TGT-T	D. Hedgecock, Bodega Bay	*Oncorhynchus*
	R>5'-TGA-AGC-AGC-AGA-TAA-AGC-A	Marine Laboratory, University	*tshawytscha*
Ots2	F>5'-ACA-CCT-CAC-ACT-TAG-A	of California, Davis, personal	"
	R>5'-AAT-ATC-CTT-CAC-ACT-G	communication	
Ots3	F>5'-CAC-ACT-CTT-TCA-GGA-G		"
	R>5'-AGA-ATC-ACA-ATG-GAA-G		
Ots4	F>5'-GAC-CCA-GAG-GAC-AGC-ACA-A	"	"
	R>5'-GGA-GGA-CAC-ATT-TCA-GCA-G		
Ots5	F>5'-ACA-GCA-GTC-TAC-ATT-GAC-C	"	"
	R>5'-TGT-TCA-TTA-AAA-CCA-AAA-A		
Ots6	F>5'-TCT-CTT-CCA-GCA-CCA-CAC-A	"	"
	R>5'-AGA-CAG-TTT-TTC-CAC-ATC-C		

Appendix A. continued

Locus	Primer Sequence P>Forward, R>Reverse	Reference	Source Species
Sfo8	P>5'-CAA-CGA-GCA-CAG-AAC-AGG	Angers et al. 1995	*Salvelinus*
	R>5'-CTT-CCC-CTG-GAG-AGG-AAA		*fontinalis*
Sfo12	P>5'-GGT-TTT-GAA-GAG-TGA-CAG	"	"
	R>5'-CCC-GTT-TCA-CAA-TCA-GAG		
Sfo18	P>5'-TGG-TGT-ATC-CTG-CTC-CTG	"	"
	R>5'-TGG-AAT-GTG-TGT-CTG-TTT-TCT		
Sfo23	P>5'-GTG-TTC-TTT-TCT-CAG-CCC	"	"
	R>5'-AAT-GAG-CGT-TAC-GAG-AGG		
Ssa4	P>5'-ATT-AGG-CAG-CAG-CAG-GCT-GC	McConnell et al. 1995a	*Salmo*
	R>5'-TGT-TCA-CTC-ACT-GAC-ACG-CG		*salar*
Ssa14	P>5'-CCT-TTT-GAC-AGA-TTT-AGG-ATT-TC	"	"
	R>5'-CAA-ACC-AAA-CAT-ACC-TAA-AGC-C		
Ssa85	P>5'-AGG-TGG-GTC-CTC-CAA-GCT-AC	O'Reilly et al. 1996	"
	R>5'-ACC-CGC-TCC-TCA-CTT-AAT-C		
Ssa171	P>5'-TTA-TTA-TCC-AAA-GGG-GTC-AAA-A	"	"
	R>5'-GAG-GTC-GCT-GGG-GTT-TAC-TAT		
Ssa197	P>5'-GGG-TTG-AGT-AGG-GAG-GCT-TG	"	"
	R>5'-TGG-CAG-GGA-TTT-GAC-ATA-AC		
Ssa202	P>5'-CTT-GGA-ATA-TCT-AGA-ATA-TGG-C	"	"
	R>5'-TTC-ATG-TGT-TAA-TGT-TGC-GTG		
Ssa289	P>5'-CTT-TAC-AAA-TAG-ACA-GAC-T	McConnell et al. 1995a	"
	R>5'-TCA-TAC-AGT-CAC-TAT-CAT-C		
Ssa293	P>5'-TGG-TTA-TTT-GTT-TCC-AGA-G	"	"
	R>5'-ATC-AGA-TAC-ACA-GAG-ACG-G		
µSat15	P>5'-TGC-AGG-CAG-ACG-GAT-CAG-GC	Estoup et al. 1993	*Salmo*
	R>5'-AAT-CCT-CTA-CGT-AAG-GGA-TTT-GC		*trutta*
µSat60	P>5'-CGG-TGT-GCT-TGT-CAG-GTT-TC	"	"
	R>5'-GTC-AAG-TCA-GCA-AGC-CTC-AC		
µSat73	P>5'-CCT-GGA-GAT-CCT-CCA-GCA-GGA	"	"
	R>5'-CTA-TTC-TGC-TTG-TAA-CTA-GAC-CTA		

APPENDIX B: ALLELE FREQUENCIES FOR ODD-YEAR PINK SALMON POPULATIONS

Locus ,allele (bp)[a,b]	Northwest Alaska		South-central Alaska.		Southeast Alaska		North British Columbia		South British Columbia		Northwest Washington		Even-year
	nom91	sna91	afk95	kop95	lpw93	gas93	bab93	khy93	fra95	cay95	sti95	gra95	kop94
*One*μ3													
(N)	(39)	(31)	(52)	(44)	(20)	(20)	(40)	(40)	(40)	(40)	(40)	(40)	(40)
154	0	0.016	0	0	0	0	0	0	0	0	0	0	0
156	0	0	0.010	0	0	0	0	0	0	0	0	0	0
162	0.526	0.597	0.605	0.614	0.600	0.675	0.638	0.638	0.738	0.688	0.762	0.850	0.488
168	0.474	0.387	0.385	0.386	0.400	0.325	0.362	0.362	0.262	0.312	0.238	0.150	0.512
*Ots*1													
(N)	(39)	(31)	(52)	(44)	(19)	(20)	(40)	(40)	(40)	(40)	(40)	(40)	(40)
214	0	0	0	0	0	0	0	0.013	0	0	0	0	0
216	0	0.016	0.010	0	0	0.025	0	0.013	0.025	0	0.013	0.075	0
218	0	0.032	0	0	0	0	0.038	0.037	0	0	0.025	0	0
220	0.090	0.016	0.038	0.045	0.026	0.050	0	0.025	0	0.013	0.013	0.013	0.050
222	0	0.016	0	0.011	0.026	0.025	0	0	0.150	0.087	0.188	0.024	0.025
224	0.205	0.194	0.144	0.193	0.263	0.225	0.225	0.225	0.125	0.224	0.150	0.212	0.124
226	0.064	0.016	0.010	0.011	0	0	0	0	0	0.012	0	0	0.050
228	0.244	0.275	0.183	0.171	0.185	0.125	0.113	0.087	0.150	0.062	0.037	0.088	0.237
230	0.076	0.032	0	0.023	0	0.025	0.024	0.038	0.100	0.037	0.024	0.050	0.012
232	0.154	0.226	0.279	0.239	0.238	0.200	0.287	0.263	0.200	0.162	0.250	0.112	0.225
234	0.037	0.016	0.076	0.046	0.079	0.100	0.087	0.050	0.025	0.038	0.050	0.175	0.050
236	0.026	0.016	0.114	0.091	0.105	0.100	0.213	0.138	0.025	0.025	0.112	0.188	0.138
238	0.013	0.016	0.019	0.057	0	0.025	0.013	0.025	0.100	0.213	0.037	0.063	0.025
240	0.026	0.032	0.029	0.023	0.026	0.050	0	0.062	0.063	0.050	0.037	0	0
242	0	0	0	0.011	0	0	0	0	0.013	0.013	0.012	0	0.038
244	0.013	0	0.010	0.023	0	0.025	0	0.024	0.012	0	0.013	0	0.013
246	0.026	0.065	0.029	0.023	0	0.025	0	0	0.012	0.013	0.013	0	0
248	0	0	0.010	0.011	0.026	0	0	0	0	0.038	0.013	0	0.013
250	0	0	0.010		0.026	0	0	0	0	0.013	0.013	0	0
252	0.013	0.016	0	0	0	0	0	0	0	0	0	0	0
254	0	0.016	0	0	0	0	0	0	0	0	0	0	0
256	0	0	0	0.011	0	0	0	0	0	0	0	0	0
260	0.013	0	0.010	0.011	0	0	0	0	0	0	0	0	0
264	0	0	0.010	0	0	0	0	0	0	0	0	0	0
268	0	0	0.019	0	0	0	0	0	0	0	0	0	0
μ*Sat*60													
(N)	(40)	(31)	(52)	(44)	(19)	(20)	(40)	(40)	(40)	(40)	(40)	(40)	(40)
103	0.063	0.016	0	0	0	0	0	0	0	0	0	0	0
105	0	0	0	0	0	0	0	0	0.088	0.012	0.012	0	0
107	0	0	0	0.011	0	0	0	0	0.024	0	0.012	0	0
109	0.863	0.871	0.731	0.670	0.684	0.700	0.600	0.588	0.613	0.713	0.488	0.438	0.913

Appendix B. continued

Locus ,allele (bp)[a,b]	Northwest Alaska		Southcentral Alaska.		Southeast Alaska		North British Columbia		South British Columbia		Northwest Washington		Even-year
	nom91	sna91	afk95	kop95	lpw93	gas93	bab93	khy93	fra95	cay95	sti95	gra95	kop94
111	0	0	0	0.023		0.025	0.012	0.025	0	0	0.025	0.137	0
113	0.074	0.097	0.240	0.284	0.211	0.175	0.375	0.337	0.275	0.263	0.463	0.425	0.087
115	0	0.016	0	0	0.053	0.025	0	0	0	0	0	0	0
117	0	0	0.019	0.012	0.052	0.075	0.013	0.050		0.012	0	0	0
121	0	0	0.010	0	0	0	0	0	0	0	0	0	0
*Ssa*85													
(N)	(38)	(31)	(52)	(44)	(18)	(19)	(40)	(40)	(40)	(40)	(40)	(40)	(40)
137	0	0	0	0	0	0	0	0	0	0	0	0	0.013
141	0	0	0	0	0	0	0	0	0	0	0	0	0.013
147	0	0	0	0	0	0	0	0	0	0.013	0	0	0.013
151	0	0	0	0	0	0	0	0	0	0	0	0	0.025
153	0	0.016	0	0.011	0	0	0	0	0	0	0	0	0.013
155	0.013	0	0	0	0	0	0	0	0	0	0	0	0
157	0	0	0.058	0	0	0	0.013	0.025	0	0	0.013	0	0.025
159	0	0	0	0	0	0	0.075	0.013	0.013	0	0	0	0
161	0	0.016	0	0	0	0	0.013	0.025	0	0	0	0	0.050
163	0	0.016	0.010	0	0	0	0	0	0	0	0	0	0.025
165	0	0	0.010	0	0.028	0	0.050	0	0	0	0	0	0
167	0	0	0.010	0	0.056	0.053	0	0.013	0	0	0	0	0.038
169	0	0	0	0	0	0	0	0	0	0	0	0	0.038
171	0	0	0	0.057	0.028	0	0	0	0	0	0	0	0
173	0.013	0.016	0	0.011	0	0	0	0.013	0	0	0	0	0.013
175	0.013	0.016	0	0	0	0	0	0	0	0	0.025	0	0
177	0.013	0.016	0.010	0	0	0	0	0	0	0	0.013	0	0
179	0	0	0.010	0.034	0.056	0	0	0	0	0	0	0	0.013
181	0.013	0.016	0.029	0	0	0.026	0	0.025	0	0	0	0	0.013
183	0.039	0	0	0	0	0	0	0	0	0.013	0	0	0.038
185	0.013	0.016	0.029	0.034	0	0	0	0.013	0.013	0	0	0	0
187	0.026	0.048	0.010	0.011	0	0	0	0	0.013	0	0.013	0	0.025
189	0.013	0.032	0.048	0.034	0.028	0.026	0.013	0.038	0.025	0	0.013	0.013	0.038
191	0.013	0.032	0.019	0.034	0	0	0.013	0.038	0.038	0.013	0.050	0	0.088
193	0.039	0.032	0.058	0.034	0.028	0	0	0	0.050	0.038	0	0.100	0.013
195	0	0.032	0.029	0.023	0	0	0	0	0.013	0.113	0	0.025	0.050
197	0.066	0	0.029	0.091	0.028	0.026	0.075	0.063	0.013	0.063	0.113	0.013	0.013
199	0.039	0.016	0.038	0.034	0.056	0.079	0.025	0.050	0.088	0.075	0.050	0	0.038
201	0.013	0	0.106	0.034	0.083	0.079	0.025	0	0.063	0.063	0.088	0.075	0.063
203	0.013	0.032	0.048	0.057	0.056	0.079	0.075	0.075	0.100	0.125	0.063	0.075	0
205	0.013	0.048	0.048	0.057	0.056	0.105	0.050	0.113	0.075	0.013	0.088	0.075	0
207	0.026	0.016	0.019	0.045	0.139	0.026	0.113	0.050	0.025	0.075	0.088	0.050	0.025
209	0.026	0.016	0.038	0.068	0.028	0.053	0.075	0.075	0.075	0.113	0.150	0.063	0

Appendix B. continued

Locus allele	Northwest Alaska		Southcentral Alaska.		Southeast Alaska		North British Columbia		South British Columbia		Northwest Washington		Even-year
(bp)[a,b]	nom91	sna91	afk95	kop95	lpw93	gas93	bab93	khy93	fra95	cay95	sti95	gra95	kop94
211	0	0.032	0.058	0.068	0	0.053	0.088	0.038	0.050	0.013	0.025	0.100	0.088
213	0.053	0.113	0	0.023	0.083	0.079	0.063	0.038	0.038	0.063	0.050	0.013	0.025
215	0.053	0.032	0.038	0.034	0.056	0.053	0.075	0.075	0.050	0.025	0.013	0.063	0
217	0.079	0.048	0.019	0.045	0.028	0.026	0.038	0.025	0.038	0	0.025	0.075	0.013
219	0.053	0.032	0.067	0.034	0.056	0.026	0.063	0.025	0.038	0.025	0.013	0.138	0.025
221	0.013	0.065	0.048	0.023	0	0	0	0.038	0.013	0.038	0.025	0.038	0.038
223	0.026	0.032	0.029	0.011	0	0.105	0.038	0.013	0.025	0	0.013	0.013	0.025
225	0.066	0.048	0.010	0	0	0	0.013	0	0.013	0	0	0.025	0.013
227	0	0.016	0.029	0.023	0	0	0	0.025	0	0	0	0	0.013
229	0.039	0.016	0	0.011	0.028	0.026	0	0.013	0.038	0.038	0.013	0	0.050
231	0.013	0.032	0	0	0	0.053	0.013	0.013	0	0	0	0	0.013
*Ssa*85													
(N)	(38)	(31)	(52)	(44)	(18)	(19)	(40)	(40)	(40)	(40)	(40)	(40)	(40)
233	0.013	0	0.038	0.034	0	0	0	0.013	0.013	0.025	0	0	0
235	0.053	0.016	0	0	0.056	0	0	0.013	0	0	0.013	0	0
237	0.039	0.016	0	0.011	0	0	0	0.025	0.038	0	0	0	0.013
239	0.026	0	0	0	0.028	0	0	0	0	0	0	0	0
241	0	0	0	0	0	0	0	0	0.025	0.013	0	0	0
243	0	0	0	0	0	0	0	0	0.013	0	0.025	0	0
245	0	0	0	0	0	0	0	0	0	0	0.025	0	0.013
247	0	0	0.010	0	0	0	0	0	0	0	0	0.050	0
249	0.013	0	0	0	0	0.026	0	0.013	0.013	0.013	0	0	0
251	0.026	0.032	0	0	0	0	0	0	0	0.025	0	0	0
253	0	0.016	0	0.011	0	0	0	0.013	0	0	0	0	0
255	0.026	0	0	0	0	0	0	0	0	0.013	0	0	0
263	0	0.016	0	0	0	0	0	0	0	0	0	0	0
*Ssa*197													
(N)	(39^{1})	(29^{1})	(49^{2})	(40^{0})	(19^{0})	(18^{1})	(37^{1})	(37^{0})	(40^{0})	(35^{5})	(37^{3})	(37^{3})	(40^{0})
124	0	0.017	0	0	0	0	0	0.027	0	0	0	0	0.013
128	0.026	0	0.041	0	0.026	0	0	0	0	0.014	0	0	0
132	0.064	0.069	0.061	0.050	0	0.028	0.014	0.054	0.063	0.071	0.054	0	0.063
136	0.013	0.052	0.020	0.038	0.053	0.028	0.095	0.041	0.025	0.114	0.027	0.054	0.038
140	0.167	0.103	0.082	0.113	0.079	0.028	0.041	0.041	0.063	0.029	0.014	0.095	0.088
144	0.192	0.155	0.092	0.175	0.079	0.194	0.203	0.149	0.113	0.014	0.122	0.041	0.100
148	0.192	0.121	0.102	0.125	0.132	0.083	0.041	0.041	0.075	0	0.027	0.054	0.100
152	0.077	0.121	0.092	0.088	0.105	0.083	0.041	0.068	0.063	0.029	0.095	0.041	0.075
156	0.038	0.121	0.071	0.075	0.211	0.111	0.149	0.068	0.075	0.129	0.122	0	0.100
160	0.051	0.052	0.112	0.100	0.132	0	0.068	0	0.025	0.071	0.068	0	0.088
164	0.051	0.069	0.092	0.100	0.026	0.028	0.014	0.176	0.038	0.057	0.122	0.081	0.088
168	0.026	0.017	0.051	0.025	0	0.028	0.122	0.054	0.050	0.143	0.081	0.081	0.050

Appendix B. continued

Locus allele (bp)[a,b]	Northwest Alaska		Southcentral Alaska.		Southeast Alaska		North British Columbia		South British Columbia		Northwest Washington		Even-year
	nom91	sna91	afk95	kop95	lpw93	gas93	bab93	khy93	fra95	cay95	sti95	gra95	kop94
172	0.013	0.052	0.061	0.013	0.105	0.167	0.095	0.135	0.075	0.057	0.041	0.108	0.075
176	0.038	0.052	0.031	0	0.026	0.083	0.068	0.014	0.063	0.043	0.041	0.108	0.038
180	0.026	0	0.051	0.013	0.026	0.056	0.041	0.068	0.100	0.014	0.054	0.108	0.013
184	0.026	0	0.010	0.038	0	0.083	0	0.041	0.088	0.086	0.095	0.014	0
188	0	0	0.010	0.025	0	0	0.014	0	0.038	0.043	0.014	0.041	0.025
192	0	0		0.025	0	0	0	0.027	0.038	0.071	0.014	0.122	0.038
196	0	0	0.020	0	0	0	0	0	0.013	0.014	0.014	0	0
200	0	0	0	0	0	0	0	0	0	0	0	0.041	0
204	0	0	0	0	0	0	0	0	0	0	0	0.014	0.013

[a] Population codes were derived using the name (first three letters) and sample date (last two numbers) from Figure 3.1.

[b] Number of individuals with more than two alleles at Ssa197 are denoted by superscript above the sample size.

JEFFREY BRIAN OLSEN

EDUCATION

Ph.D. Fisheries, 1999, University of Washington, Seattle.
Dissertation title: Genetic Interpretation of Microsatellite Polymorphism in Pacific Salmon: Case studies in population genetics and kinship analysis

B.Sc. Fisheries, 1981, University of Washington, Seattle

EMPLOYMENT HISTORY

1998-present

Title: Molecular Geneticist, Alaska Department of Fish and Game.
Duties: Plan and conduct genetic studies and monitoring efforts needed for management and conservation of Alaska's fish and game resources.

1994-1998

Title: Research assistant, University of Washington School of Fisheries.
Duties: Conduct research on conservation genetics of Pacific salmon using novel molecular markers

1995-1996

Title: Student intern, Alaska Department of Fish and Game, Genetics Division.
Duties: Conduct population genetic survey of pink salmon using microsatellites

1989-1994

Title: Operations Manager, Prince William Sound Aquaculture Corporation. Cordova, Alaska.
Duties: Coordinate operations of five salmon hatcheries in the Prince William Sound and Copper River region. Participate in short (1-5 years) and long term (5-20 year) planning of the PWS/CR enhancement program. Coordinate review and revision of fish culture goals and criteria for pink, chum, chinook, coho, and sockeye salmon. Assist with development and design of a 10 million smolt sockeye hatchery. Supervise five hatchery managers and an annual operating budget of 4.5 million dollars.

1987-1989

Title: Hatchery Manager, Esther Island Hatchery, Prince William Sound Aquaculture Corporation, Cordova Alaska
Duties: Responsible for short term planning and day to day administration of operations at a four species (pink, chum, coho, chinook) salmon hatchery.

1985-1987

Title: Assistant Hatchery Manager, Esther Island Hatchery, Prince William Sound Aquaculture Corporation, Cordova Alaska'

Duties: Responsible for day to day fish culture activities and direct supervision of the fish culture staff.

PUBLICATIONS AND PROFESSIONAL REPORTS

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 127535-550.

Olsen, J.B., P. Bentzen, and J.E. Seeb. 1998. Characterization of seven microsatellite loci derived from pink salmon. Molecular Ecology 7(8):1087-1089.

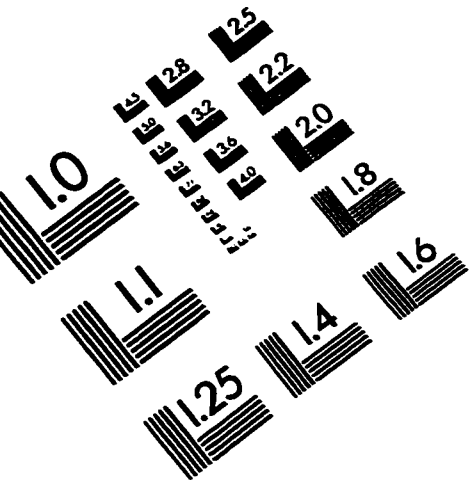
Seeb, J.E., C. Habicht, J.B. Olsen, P. Bentzen, J.B. Shaklee, and L.W. Seeb. 1998. Allozyme, mtDNA, and microsatellite variants describe structure of populations of pink and sockeye salmon in Alaska. N. Pac. Anadr. Fish Comm. Bull. No. 1.

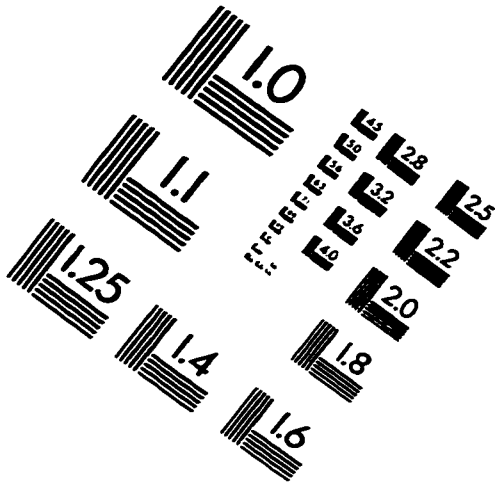
Olsen, J.B., J.K. Wenburg, and P. Bentzen. 1996. Semi-automated multilocus genotyping of Pacific salmon (*Oncorhynchus* spp.) using microsatellites. Molecular Marine Biology and Biotechnology 5(4):259-272.

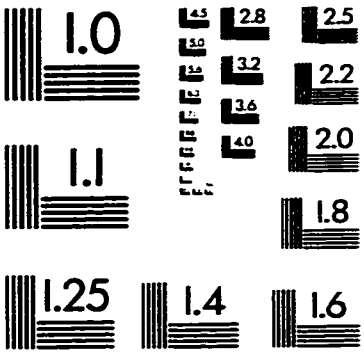
Wenburg, J.K., J.B. Olsen, and P. Bentzen. 1996. Multiplexed systems of microsatellites for genetic analysis in coastal cutthroat trout (*Oncorhynchus clarki clarki*) and steelhead (*Oncorhynchus mykiss*). Molecular Marine Biology and Biotechnology 5(4):273-283.

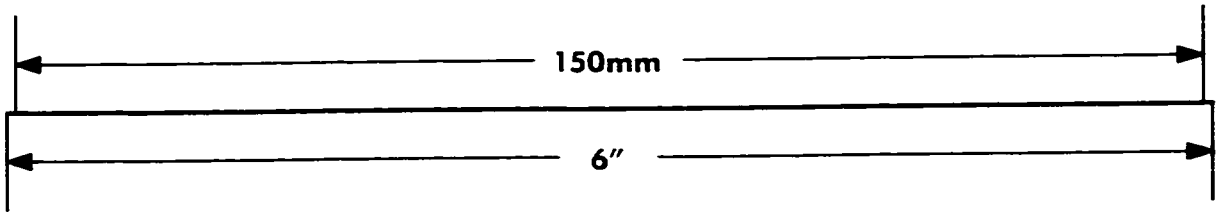
Olsen, J.B. 1994. Development of the pink salmon enhancement program in Prince William Sound, Alaska. Pages 57-68 *in*: W.W. Smoker, editor. Proceedings of the Sixteenth North East Pacific Pink and Chum Salmon Workshop, February 1993. Alaska Sea Grant.

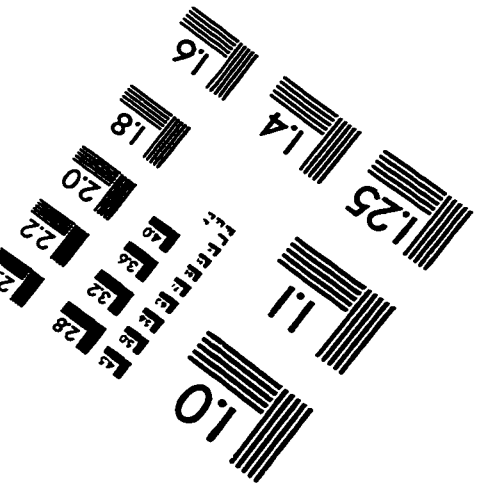
IMAGE EVALUATION
TEST TARGET (QA-3)

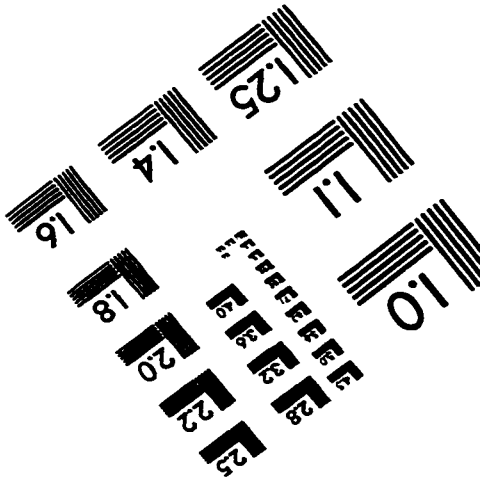












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989