

Exxon Valdez Oil Spill
Restoration Project Final Report

Construction of a Linkage Map for the Pink Salmon Genome

Restoration Project 96910
Final Report

Fred W. Allendorf
Kathy L. Knudsen
Kate R. Linder
Paul Spruell
Joe A. Tyburczy

Division of Biological Sciences
University of Montana
Missoula, MT 59812

for:

Alaska Department of Fish and Game
Habitat and Restoration Division
333 Raspberry Road
Anchorage, Alaska 99518

October 2006

The *Exxon Valdez* Oil Spill Trustee Council administers all programs and activities free from discrimination based on race, color, national origin, age, sex, religion, marital status, pregnancy, parenthood, or disability. The Council administers all programs and activities in compliance with Title VI of the Civil Rights Act of 1964, Section 504 of the Rehabilitation Act of 1973, Title II of the Americans with Disabilities Act of 1990, the Age Discrimination Act of 1975, and Title IX of the Education Amendments of 1972. If you believe you have been discriminated against in any program, activity, or facility, or if you desire further information, please write to: EVOS Trustee Council, 441 West 5th Avenue, Suite 500, Anchorage, Alaska 99501-2340; or O.E.O. U.S. Department of the Interior, Washington, D.C. 20240.

Exxon Valdez Oil Spill
Restoration Project Final Report

Construction of a Linkage Map for the Pink Salmon Genome

Restoration Project 96910
Final Report

Fred W. Allendorf
Kathy L. Knudsen
Kate R. Linder
Paul Spruell
Joe A. Tyburczy

Division of Biological Sciences
University of Montana
Missoula, MT 59812

for:

Alaska Department of Fish and Game
Habitat and Restoration Division
333 Raspberry Road
Anchorage, Alaska 99518

October 2006

Construction of a Linkage Map for the Pink Salmon Genome

Restoration Project 96910 Final Report

Study History: Project 96910 was initiated in March 1996 to construct a genetic linkage map for pink salmon (*Oncorhynchus gorbuscha*) and to use this map to evaluate effects of natural selection on the genome of this species. Such a map was proposed initially to provide the necessary platform to identify genetic damage in pink salmon inhabiting oiled streams following the March 1989 *Exxon Valdez* oil spill (EVOS). This research was designed to aid recovery efforts with pink salmon, including estimation of straying rates, description of stock structure, and testing if marine survival and other organismal measures of phenotypic variation have a genetic basis. This work was designed to support work with pink salmon under the project *Oil-Related Embryo Mortalities* (Restoration Study \191A). The objective of that project was to identify germline mutations in pink salmon exposed to oil. Genetic damage induced by oil may either be small changes in nucleotide sequence (microlesions) or large-scale changes in chromosome structure (macrolesions). Annual reports were submitted in 1997 through 2002. One thesis has been published: Pilgrim 1999). Four journal articles were published: Spruell et al. (1999), Lindner et al. (2000), Steinberg et al. (2002), and Funk et al. (2005).

Abstract: We have constructed a genetic linkage map for pink salmon (*Oncorhynchus gorbuscha*) and have experimentally investigated marine survival and fitness of pink salmon in Prince William Sound. We analyzed segregation of 596 DNA fragments in odd-year pink salmon. Of these markers, 553 were assigned to one of 44 linkage groups. We estimated gene-centromere distances for 312 loci using gynogenetic diploid progeny. In August 1998, we collected gametes and tissue from 150 pink salmon from Likes Creek and used single-pair mating to produce 75 families. In May 1999, approximately 48,000 individuals were marked and released into Resurrection Bay from the Alaska SeaLife Center. In August 1999, we collected 68 adult pink salmon from Likes Creek and produced 68 families. These families were raised at the Alaska SeaLife Center and approximately 24,000 fry were marked and released into Resurrection Bay in May 2000. Only 36 returning adults from the 1998 experimental cohort were collected in August 2000. In August 2001, 259 returning adults from the 1999 cohort were collected. We assigned the adult returns to family using genotype data from 10 loci. We found nearly random family survival and high heritability of body length. Morphological traits were all moderately to highly heritable, but egg number and egg weight were not heritable, suggesting that past selection has eliminated additive genetic variation in egg number and egg weight or that there is high environmental variance in these traits. Genetic correlations were similar for nonadjacent morphological traits and adjacent traits. Genetic correlations predicted phenotypic correlations fairly accurately, but some pairs of traits with low genetic correlations had high phenotypic correlations, and vice versa, emphasizing the need to use caution when using phenotypic correlations as indices of genetic correlations.

Key Words: adaptation, fitness, gene-centromere mapping, genetics, linkage map, marine survival, mutation, *Oncorhynchus gorbuscha*, pink salmon.

Project Data: *Description of data* -- We have two primary sets of data: one for the linkage map and one for the marine survival and fitness experiment. Data for the linkage map are the inheritance of DNA fragments in the haploid and gynogenetic diploid progeny of two pink salmon females (A95-103 and V96-13). Sixteen additional diploid families were tested for nonrandom segregation between all pair-wise combinations of 14 allozyme and three microsatellite loci. The haploid data set consists of 596 polymorphic DNA fragments loci in female A95-103 and 94 of her haploid progeny, and 123 polymorphic DNA fragments in female V96-13 and 90 of her haploid progeny. The diploid data set consists of genotypes of 70 gynogenetic diploid progeny from female A95-103 at 319 loci and of genotypes of 54 gynogenetic diploid progeny from female V96-13 at 40 loci. Data for the marine survival and fitness experiment are genotypes at ten PCR-based loci for 50 families (50 parent pairs with 10 embryos each) from the 1998 experimental release. An additional 36-40 embryos from seven of the 1998 families were analyzed at nine microsatellite loci to investigate mutation rates and patterns. The parents for the experimental cohort produced in 1998 were genotyped at 12 additional PCR-based loci and 34 allozyme loci. The 1999 parents were genotyped at 10 PCR-based loci and 30 allozyme loci. Their progeny, the 259 returning adult fish recovered in August 2001, were genotyped at the same 10 PCR-based loci and this information has been used to assign them to parental family. Four meristic characters, as well as body length, egg mass, and egg number were recorded for both sets of parents as well as the 36 marked adults collected in August 2000. The 259 marked adults collected in August 2001 were measured for body length, egg mass, and egg number. *Format* -- All data were entered as Excel spreadsheets. *Custodian* -- Contact Fred W. Allendorf, Division of Biological Sciences, University of Montana, Missoula, MT 59812.

Citation:

Allendorf, F.W., K. L. Knudsen, K. R. Linder, P. Spruell, and J.A. Tyburczy. 2004. Construction of a linkage map for the pink salmon genome. *Exxon Valdez Oil Spill Restoration Project Final Report* (Restoration Project 96910), Alaska Department of Fish and Game, Habitat and Restoration Division, Anchorage, Alaska.

TABLE OF CONTENTS

LIST OF TABLES	7
LIST OF FIGURES	8
LIST OF APPENDICES	9
EXECUTIVE SUMMARY.....	10
INTRODUCTION.....	14
OBJECTIVES.....	15
METHODS.....	15
RESULTS.....	17
DISCUSSION.....	21
CONCLUSIONS.....	24
ACKNOWLEDGMENTS.....	25
LITERATURE CITED.....	25
OTHER REFERENCES	32

LIST OF TABLES

Table		Page
1	A breakdown of linkage groups based on number of markers in each group and the average size in cM.	33
2	Summary of polymorphic loci detected by four different techniques.	34
3	Summary of linkages in normal diploid families between allozymes and microsatellites.	35
4	Comparison of interference estimates based on half-tetrad and haploid data from nine linkage groups. The number in parenthesis corresponds to the relative amount of interference for each method. The half-tetrad interference estimates are grouped into no(0), low (1-3), moderate (4-6), high (7-9) or in some cases complete interference. The haploid data range is grouped into no (0), low (0.01-0.33), moderate (0.34-0.66), high (0.67-0.99), and complete (1.0) interference.	36
5	Summary of adult progeny recovered from each family and parent, the number of alevins from each parent pooled prior to freshwater rearing, and the percentage of alevins from each parent recovered as adults.	37

LIST OF FIGURES

Figure		Page
1	Genetic linkage of pink salmon based on the inheritance of 602 loci. Numbers to the left indicate recombination rates (cM). Locus names to the right. Centromeres are indicated by black rectangles.	39
2	Figure 3. Comparison of gene-centromere distances (y) for odd and even year pink salmon. The line indicates the expectation if y values are equivalent in both years.	42
3	Diagram of our half-sib family experimental design. Numbers across the top represent females, numbers down the side represent males. The squares contain the family designation and the numbers of individuals used to make each family. The number of the dam is the family number in each half-sibling cross. The letter A or B for each family designates the sire. The parents of family 5A are dam 5 and sire 105, 5B are dam 5 and sire 106, 6A dam 6 and sire 105, and 6B are dam 6 and sire 106, and so forth.	43
4	Map of Resurrection Bay. Numbers indicate the location where experimental fish were collected in 2001 as designated below. LC designates Likes Creek, the location where the parents were collected.	44
5	Scatterplot of the number of progeny returning for sires (a) and dams (b). For example, there were two sires that produced 10 returning progeny. The line is the expected Poisson distribution.	45

LIST OF APPENDICES

- A. Pilgrim, K. 1999. Identification of the sex-determining locus in pink salmon. M.A., Organismal Biology and Ecology, University of Montana.
- B. Spruell, P., B.A. Greene, C. Habicht, K.L. Knudsen, K.R. Lindner, J.B. Olsen, K.L. Pilgrim, G.K. Sage, J.E. Seeb, and F.W. Allendorf. 1999. Inheritance of nuclear DNA markers in gynogenetic haploid pink salmon (*Oncorhynchus gorbuscha*). *Journal of Heredity* 90:289-296.
- C. Lindner, K.R., J. E. Seeb, C. Habicht, E. Kretschmer, D. J. Reedy, P. Spruell, and F. W. Allendorf. 2000. Gene-centromere mapping of 312 marker loci in pink salmon by half-tetrad analysis. *Genome* 43:538-549.
- D. Steinberg, E. K., K. R. Lindner, J. Gallea, A. Maxwell, J. Meng, and F. W. Allendorf. 2002. Rates and patterns of microsatellite mutations in pink salmon. *Molecular Biology and Evolution* 19:1198-1202.
- E. Funk, W.C., Tyburczy, J.A., Knudsen, K.L., Lindner, K.R., and F. W. Allendorf. 2005. Genetic basis of variation in morphological and life history traits of pink salmon (*Oncorhynchus gorbuscha*). *J. Heredity* 96:24-31.

EXECUTIVE SUMMARY

The project to construct a genetic linkage map of the pink salmon (*Oncorhynchus gorbuscha*) genome, and to use this map to study the marine survival and fitness in this species, is complete. This work resulted in four peer-reviewed publications in the journals (Appendices B through E) and a published thesis (Appendix A). These journal articles represent the core of this final report.

The linkage map will allow the evaluation of genetic impacts of the March 1989 *Exxon Valdez* oil spill on pink salmon populations and will help to document the recovery of affected populations in Prince William Sound. A linkage map will be essential for detecting and understanding causes of reduced egg and embryo survival in oiled areas. In addition, the markers that are mapped and characterized in detail will aid other recovery efforts with pink salmon, including estimation of straying rates, description of stock structure, and testing for a genetic basis of marine survival.

Elevated embryo mortality was detected in populations of pink salmon inhabiting oiled streams following the spill. These increased rates of mortality persisted through the 1993 field season, three generations after the spill. This suggests that genetic damage may have occurred as a result of exposure to oil during early developmental life-stages.

The genetic linkage map will provide the platform to address the genetic impact of the oil spill. The initial framework of the map used haploid progeny to avoid the difficulties associated with dominant markers that obscure recessive alternatives in diploids. Gynogenetic diploids from the same family were also examined to locate centromeres of chromosomes and facilitate the consolidation of the map.

Gametes and tissues of pink salmon were collected from the Armin F. Koernig hatchery in August of 1995 as well as the Solomon Gulch hatchery in August 1996. Families of gynogenetic haploid and diploid embryos were produced in cooperation with the Alaska Department of Fish and Game by mixing irradiated sperm with eggs from individual females. One family (A95-103) was chosen to be the primary reference family upon which initial mapping efforts were focused.

Linkage analysis of 596 DNA markers segregating in the gynogenetic haploids produced a genetic map comprising 44 linkage groups covering a distance of 4550 centiMorgans. Assuming a minimum distance of 28 centiMorgans for linkage detection and accounting for all the gaps and unlinked markers the minimum distance of the pink salmon genome is 6472 centiMorgans. The haploid pink salmon genome consists of approximately 2.72 million kilobase pairs, thus we estimate approximately 420 kilobase pairs per centiMorgan.

Thirteen allozyme loci have been added to the map using gynogenetic diploid and normal diploid data. Five allozyme loci are polymorphic in female A95-103 and thus could be tested for nonrandom segregation using the gynogenetic diploid data. The other eight

loci were placed on the map through classic linkage analysis of diploid pink salmon families. With the addition of these markers the linkage map consists of a total of 609 markers.

The microsatellites and genes of known function added to the linkage map serve as landmarks, or “anchor loci” and will facilitate comparisons between maps. These loci allow comparison of genetic linkage of odd- and even-year pink salmon, estimation of recombination rates of males and females, and incorporation of data from other salmonid linkage maps. The known genes will be of particular interest during the second phase of this project in which we examine selective effects of the marine environment on the pink salmon genome.

A complementary even-year map has been constructed. Development of this map, based on the segregation of loci in family V96-13, followed the same design as the odd-year map. This map enabled us to compare odd- and even-year pink salmon as well as add seven new markers to the odd-year map. This even-year map consists of 123 loci, 103 of which have been assigned to one of 33 linkage groups. One locus included on this map is a gene of known function (*MHCba2*).

We have generated a large number of markers distributed throughout the genome using haploid embryos and multilocus techniques. Due to their polyploid ancestry, salmonid genomes are large, therefore many markers will be required to span the entire genome of pink salmon. We have successfully created a genome map with 44 linkage groups. However despite the number of the markers examined, we were unable to consolidate the map enough to reduce the number of linkage groups to 26, the number of chromosome pairs in pink salmon ($2N = 52$). Additional markers must be mapped in order to consolidate the map. We have collected gene-centromere distances of 319 loci using gynogenetic diploids. Comparison of the genome maps of odd- and even-year fish has revealed no significant differences between them. We will submit a publication on the results of our mapping efforts.

We are now focusing on the marine survival and fitness portion of the study. Two experimental cohorts have been produced and their returning adult progeny collected. Both cohorts were hatched and released from the Alaska SeaLife Center in Seward. In August 1998, gametes and tissue from 150 pink salmon from Likes Creek were collected and 75 single-cross families were produced. Ten embryos from each family were analyzed to evaluate inheritance of genetic markers. A total of 48,329 individuals from 49 of these families were marked and released into Resurrection Bay in May, 1999. At the time of release, 1000 fry from the experimental families were randomly sampled for genetic analysis. In August 1999, gametes and tissue from 68 pink salmon from Likes Creek were collected and 68 half-sibling families were produced. A total of 24,216 fry from all 68 families were marked and released into Resurrection Bay in April 2000. A sample of 500 fry was collected at the time of release.

In August 2000 we planned to collect experimental adult pink salmon as they returned to the fish pass at the Alaska SeaLife Center. Failure of the fish pass to attract any fish

forced us to modify these plans. A total of 36 marked pink salmon were collected by seining freshwater streams in upper Resurrection Bay and from recreational fishermen responding to an incentive. Based on the collection effort and the number of fish collected it is our belief that a significant number of marked pink salmon returned to Resurrection Bay, but due to our limited resources we were unable to collect more returning adults.

Though the fish pass at the Alaska SeaLife Center did attract fish 2001, probably due to increased water flow, it still failed to capture them. We collected fish through alternate means including seining nearby rivers, hook-and-line, and an incentive to recreational fishermen. We recovered 259 adult fish from the 1999 experimental cohort.

We genotyped the 36 returning marked pink salmon recovered in August 2000 at nine microsatellite loci and two genes of known function (*MHC α 1* and *GH2*) and successfully assigned them back to their family of origin.

The 259 fish from the 1999 cohort were genotyped at nine microsatellite loci and one gene of known function (*GH2*) which have been used to place them into families. Analysis based on this information has revealed nearly random family returns, and strong heritability of length in both males and females (0.38 and 0.41 respectively).

Several mutations were detected at two of the nine microsatellite loci (*SSA408* and *OGO1c*) in the embryos of the 1998 cohort during the analysis of the transmission of genetic markers used to evaluate parentage. Further analysis revealed a great deal of heterogeneity in microsatellite mutation rates and patterns. Mutation rate estimates ranged from 8.5×10^{-3} and 3.9×10^{-3} mutations per transmission in *SSA408* and *OGO1c* respectively, to 0.0 at the other seven loci.

The detection of large clusters of identical mutations at one locus, *SSA408*, indicated that the majority of mutant alleles identified reflected mutational events that occurred early in the differentiation of the germline. Evidence for a hypermutable allele at *SSA408* was also detected. Genetic analysis of the adult progeny collected from the 1999 cohort has revealed similar patterns of mutation. Mutations were detected at both *SSA408* and *OGO1c* at frequencies of 1.9×10^{-3} and 1.2×10^{-2} mutations per transmission. These rates are comparable to those found in the 1998 cohort, and as in that cohort, no mutations were detected at the other seven microsatellite loci. These findings indicate that microsatellite mutation dynamics are complex and likely vary substantially among loci.

The mapping portion of this project is complete. We released and collected both cohorts of experimental progeny. We have determined the genotypes of the parents used to generate both cohorts, the genotypes of returning adult progeny from both cohorts, as well as genotypes of embryos sampled from the 1998 cohort. We also collected morphological data on the parents and the returning adults the 1998 cohort.

Understanding the genetic basis of phenotypic variation is essential for predicting the direction and rate of phenotypic evolution. We estimated heritabilities and genetic correlations of morphological (fork length, pectoral and pelvic fin ray counts, and gill arch raker counts) and life history (egg number and individual egg weight) traits of pink salmon from Likes Creek, Alaska, in order to characterize the genetic basis of phenotypic variation in this species. Families were created from wild-caught adults, raised to the fry stage in the lab, released into the wild, and caught as returning adults and assigned to family using microsatellite loci and a growth hormone locus.

Morphological traits were all moderately to highly heritable, but egg number and egg weight were not heritable, suggesting that past selection has eliminated additive genetic variation in egg number and egg weight or that there is high environmental variance in these traits. Genetic correlations were similar for non-adjacent morphological traits and adjacent traits. Genetic correlations predicted phenotypic correlations fairly accurately, but some pairs of traits with low genetic correlations had high phenotypic correlations, and vice versa, emphasizing the need to use caution when using phenotypic correlations as indices of genetic correlations.

INTRODUCTION.

We proposed to construct a genetic linkage map for pink salmon (*Oncorhynchus gorbuscha*) and to use this map to evaluate effects of natural selection on the genome of this species. Such a map was proposed initially to provide the necessary platform to identify genetic damage in pink salmon inhabiting oiled streams following the March 1989 *Exxon Valdez* oil spill (EVOS). We also conducted a series of experiments based at the Alaska SeaLife Center (ASLC) to identify regions of the genome that affect various organismal traits and to test for the effects of natural selection on regions of the genome that include markers used to describe genetic population structure. This research was designed to aid recovery efforts with pink salmon, including estimation of straying rates, description of stock structure, and testing if marine survival and other organismal measures of phenotypic variation have a genetic basis.

Elevated embryo mortalities were detected in populations of pink salmon inhabiting oiled streams following the spill. These increased rates of mortality persisted through the 1993 field season, three generations after the oil spill, suggesting that genetic damage may have occurred as a result of exposure to oil during early developmental life-stages. The consequences of the putative genetic damage include impaired physiological function of individuals and reduced reproductive capacity of pink salmon populations (Bue et al. 1998).

This work was designed to support work with pink salmon under the project *Oil-Related Embryo Mortalities* (Restoration Study \191A). The objective of that project was to identify germline mutations in pink salmon exposed to oil. Genetic damage induced by oil may either be small changes in nucleotide sequence (microlesions) or large-scale changes in chromosome structure (macrolesions). A detailed genetic map for pink salmon would have been invaluable for interpreting the results of Restoration Study \191A in several ways. First, it would be possible by following the inheritance of any DNA lesions to determine if they are micro- or macro-lesions. Second, these lesions could be mapped to determine if they are randomly spread throughout the genome or if they occur at mutational "hot spots" that are susceptible to oil induced damage. However, Restoration Study \191A was not completed.

Molecular genetic techniques have been used extensively to describe population structure of Pacific salmon (Utter et al. 1993; Gharrett and Smoker 1994; Seeb et al. 1998). Genetic divergence among populations has been interpreted as largely reflecting the patterns of exchange of individuals among populations (gene flow) and random changes in frequency of selectively neutral alleles within populations (genetic drift) (Allendorf and Phelps 1981; Waples 1995). This is a useful approach that allows description of the pattern and amount of gene flow among populations.

This approach to describe population structure is based upon the assumption that the pattern and amount of divergence observed is not affected by natural selection or mutation. However, even weak natural selection may have a substantial effect on the pattern of genetic divergence among populations (Allendorf 1983). In addition, different mutation rates at marker loci may also effect the amount of genetic differentiation between populations, in particular if mutation rates at some loci are high (e.g., Jin and Chakraborty 1995). Thus, the high frequency of mutations that we have detected may also have a substantial effect on the amount and pattern of genetic divergence at some loci.

Our work resulted in four peer-reviewed publications in the journals (Appendices B through E) and a published thesis (Appendix A). These journal articles represent the core of this final report.

OBJECTIVES. This project originally had the following overall specific objectives:

1. Develop several hundred variable DNA markers in pink salmon and test them for Mendelian inheritance.
2. Construct a linkage map based upon joint segregation patterns of the DNA polymorphisms detected in previous objective.
3. Map putative lesions identified in Restoration Study \191A.
4. Test for Mendelian inheritance of markers throughout the genome in progeny of fish exposed to oil. Regions that show aberrant segregation ratios in progeny of fish exposed to oil and normal 1:1 ratios in fish not exposed to oil would be candidates for oil-induced lesions.
5. Test for regions of the genome that are associated with traits of adaptive significance (e.g., marine mortality or run-timing).
6. Test if protein markers (allozymes) are under natural selection such that they may not provide accurate information about the genetic structure and amount of gene flow among populations.

We completed Objectives 1 and 2 (Appendices A, B, C, and D). We did not pursue Objective 3 because Restoration Study /191A did not identify any putative lesions for mapping, and Objective 4 because this aspect of Restoration Study \191A was not completed. We completed Objectives 5 and 6 (Appendix E)

METHODS

Gametes for the inheritance studies and linkage map were collected from Prince William Sound in collaboration with the project Oil-Related Embryo

Mortalities (Restoration Study \191A). Embryo incubation took place at the Genetics Lab facilities of ADFG. The laboratory analyses were done at the University of Montana and the ADFG genetics lab in Anchorage.

We began in FY 1998 to use the ASLC Research Facilities at Seward for experiments designed to test for natural selection at loci throughout the genome of pink salmon. Sexually mature pink salmon used in the experimental matings in 1998 and 1999 were collected from Likes Creek in Resurrection Bay. The progeny were marked with an adipose fin clip and released into Resurrection Bay. Due to the failure of the fish pass at the ASLC to attract returning adults we adjusted our plans for recapturing returning adults from the 1999 cohort to include sampling in upper Resurrection Bay.

Our initial map was constructed using gynogenetic haploid and gynogenetic diploid progeny from an odd-year individual female (95-103). This is the same procedure that has been used to build the zebrafish linkage map (Postlethwait et al. 1994). Stanley (1983) reported that haploid embryos of Atlantic salmon will develop until just prior to the stage of hatching if development of the eggs is activated by sperm in which the DNA has been inactivated by UV-radiation. We have used this technique routinely with fishes of the genus *Oncorhynchus* (Forbes et al. 1994; Appendix B). This allows us to follow the segregation and linkage relationships in haploid progeny from females. The use of haploid progeny avoids possible difficulties of dominance with some types of DNA markers because recessive alleles are not obscured by their dominant alternatives in haploids (Lie et al. 1994). Our odd-year map is primarily based on 603 segregating markers in 94 haploid progeny from a single pink salmon female (A95-103) that returned to Armin F. Koernig hatchery in Prince William Sound in August 1995. We placed a number of so-called "anchor" loci on this map.

In addition we completed the construction of a linkage map based on the segregation pattern of 90 haploid individuals in an even-year female (V96-13). Odd- and even-year pink salmon are reproductively isolated due to the fixed two-year life cycle of this species (Aspinwall 1974). Beacham et al. (1988) report substantial allozyme and morphological evidence for differentiation of alternate brood years. In addition, Phillips and Kapuscinski (1988) and Phillips et al. (1999) detected chromosomal rearrangements between odd- and even-year populations that occur in the same geographical area. Furthermore, in a recent experimental study, Gharrett et al. (1999) demonstrated outbreeding depression in crosses between the two year classes. Together, these findings suggest that the alternate brood years are reproductively isolated and genetically distinct. Having linkage data from both odd- and even-year individuals will make it possible to map more markers and will allow us to determine whether linkage relationships are conserved between the reproductively isolated year classes.

Marine Survival and Fitness Experiment

This aspect of the research was performed at the ASLC research facilities. Approximately 50,000 and 24,000 marked fish were released in spring of 1999 and 2000 respectively. We collected 36 sexually mature adults in Resurrection Bay from the 1998 cohort and 259 sexually mature adults in Resurrection Bay from the 1999 cohort produced from wild pink salmon collected from Likes Creek. A sample of the fish was collected at release and will be analyzed so that their genetic characteristics prior to the marine phase of the life cycle can be compared to the returning adults. Mousseau et al. (1998) have used a similar approach to estimate heritabilities for weight, length, and age at sexual maturation in chinook salmon.

RESULTS

Gene-Centromere Map

We estimated recombination rates between 312 loci and their centromeres using half-tetrad analysis in a recently published manuscript (Appendix C). We produced the half-tetrads by initiating development with irradiated sperm and blocking the maternal second meiotic division (Appendix B). AFLPs were significantly more centromeric than loci identified by three other techniques (allozymes, microsatellites, and PINES). The near absence of AFLPs in distal regions could limit their utility in constructing linkage maps. A large proportion of loci had y values approaching 1.0, indicating near complete crossover interference on many chromosome arms. As predicted from models of chromosomal evolution in salmonids, all duplicated microsatellite loci that shared alleles (isoloci) had y values of nearly 1.0. This is consistent with previous data from allozyme loci.

Haploid Linkage Map

We assigned 546 of the 590 markers analyzed for segregation in family A95-103 to one of 44 linkage groups covering a distance of 4559 cM (Figure 1; Tables 1-3). Given the haploid number of 26 chromosomes for pink salmon, our mapping efforts produced 18 extra linkage groups. Taking into account the extra linkage groups and 43 unassigned markers as well as the distance to the telomeres we estimate the size of the pink salmon genome to be 6691 cM. The haploid pink salmon genome is approximately 2.72 billion base pairs or 2.72 million kilobase pairs (kpb; Johnson et al 1987b); thus, we estimate approximately 406 kbp/cM.

Each of the 35 microsatellite loci analyzed for linkage in female A95-103 was assigned to one of the 44 linkage groups. Segregation analysis of three loci *Fgt1-1,2*, *Ogo7-1,2*, and *OmyFGT276-1,2* identified these loci as being duplicated (Lindner et al. 2000). *OmyFGT276-1,2* was the only duplicated locus for which female A95-103 was polymorphic at both loci allowing both loci to be mapped. The two loci map to different linkage groups, LG18 and LG53. Additional loci such as, *μSat60-1,2*, *Ocl2-1,2*, *Fgt25*, and *Ssa20.19-1,2* where only one of the duplicated pair of loci segregates in female A95-103 were confirmed to be duplicated through segregation

analysis in different pink salmon families (unpublished data). Additional segregation analysis identified a PCR null allele at *Ssa197*, *One2*, and *OmyFGT276-2*. Half of the haploid progeny amplified the same allele present in female A95-103 and half did not amplify any product (Appendix B).

AFLPs and PINEs amplified the greatest number of reproducible polymorphic loci, 393 and 162 respectively (Table 2). A total of 519 of the AFLP and PINE polymorphisms are presence / absence differences. In addition, ten of the AFLP polymorphisms and eight PINE polymorphisms appeared to be caused by a length polymorphism within a fragment. For all of these polymorphisms, individuals have one of two different sized fragments produced by the same primer combination (Appendix B).

We analyzed 168 AFLP and 101 PINE loci in gynogenetic diploid progeny at which female A95-103 is heterozygous for the presence or absence of a product based on haploid progeny (Appendix B, Lindner et al. 2000). In addition, gene-centromere data was collected for five allozyme loci (*sAAT3*, *CKC2*, *ADA2*, *GDA1*, and *PEPD2*) and 34 microsatellite loci in female A95-103.

A total of 202 out of 304 loci met the gynogenetic diploid linkage analysis criteria discussed above. Only two of the five allozyme loci polymorphic in female A95-103 could be included with a $y \leq 0.79$ (*CKC2* $y=0.29$ and *GDA1* $y=0.35$). We detected 345 non-random associations between pairs of loci. Of these pairwise associations, 293 confirmed linkage previously detected with the haploid data. However, 52 non-random associations are between loci from two different linkage groups based on the haploid data. Ten pairs of linkage groups included two or more significant pairwise associations. Two separate linkage groups based on the haploid data were consolidated to one linkage group, LG40, based on the gynogenetic diploid analysis.

Linkage analysis conducted for allozyme and microsatellite loci in 16 normal diploid pink salmon families resulted in the addition of 11 allozyme loci to the haploid map. Five allozyme loci are linked to microsatellite loci already placed on the map using the haploid data (Table 3). One allozyme locus, *CKC2* is linked to two linkage groups. Based on the gene-centromere data from female A95-103 *CKC2* is linked to three loci in LG53, *OmyRGT43* ($\chi^2 = 22.03$), *ACG/CAA240* ($\chi^2 = 14.54$), and *ACC/CAA106* ($\chi^2 = 13.88$). Based on the normal diploid data from family A95-103 this locus is linked to $\mu Str60-1,2$ in LG 25 ($rf = 0.348$, $\chi^2 = 4.26$; Table 3).

Interference: The degree of interference was estimated in nine linkage groups based on both half-tetrad (gynogenetic diploid) data as well as haploid data (Table 4). The two methods used to estimate interference agree on the degree of interference in only one of the nine comparisons (LG8; Table 4). In four cases (linkage groups 27, 33, 34, 57) the estimates were very close, either high or complete interference. In three of the four cases (linkage groups 2, 25, and 40) when the two methods do not agree the half-tetrad analysis resulted in high or complete interference and the haploid data results moderate to no

interference. In one comparison (LG5) the half-tetrad data resulted in no interference and the haploid data resulted in moderate interference.

Interference analysis of half-tetrads across six loci in LG2 results in an estimate of high interference. If a subset of only two loci from this group is analyzed using the half tetrad method the interference estimate is no interference.

Comparison of Odd- and Even-Year Linkage Map

We described the segregation of 590 markers in haploid progeny from female A95-103; we also mapped 13 allozyme loci in the same female. We assigned 546 of the 590 DNA markers and all of the allozyme loci to one of 44 linkage groups covering a distance of 4559 cM. Given the haploid number of 26 chromosomes for pink salmon, our mapping efforts produced 18 extra linkage groups. Taking into account the extra linkage groups and 44 unassigned markers as well as the distance to the telomeres we estimate the size of the pink salmon genome to be 6691 cM. The haploid pink salmon genome is approximately 2.72 million kilobase pairs (kpb; Johnson et al 1987b); thus, we estimate approximately 406 kbp/cM. These results are consistent with our expectations when comparing to maps constructed in other fishes.

We completed construction of a linkage map for even-year pink salmon from Prince William Sound. We analyzed the segregation pattern of 85 loci in an even-year family (V96-13) and assigned 63 of 85 loci to one of 22 linkage groups. One gene of known function, *MHCB^a2*, is assigned to a linkage group that consists of one microsatellite and two PINE loci.

We estimated gene-centromere distances for 13 additional microsatellite loci in both odd and even year pink salmon (Figure 2). Most loci fall near our expectation for equal y 's in both years. One locus, *FGT34*, appears to be telomeric in the even year family and centromeric in the odd year family. *RGT1* displays the opposite pattern, appearing to be proximal in the even year family but distal in the odd year family.

Mutation Analysis

Our results have provided exciting and important information about mutation processes in microsatellites that are accepted for publication in the journal *Molecular Biology and Evolution* (Appendix D). Our experimental design depends upon being able to place returning adults into their correct family on the basis of their multiple-locus genotypes. We tested this by examining inheritance data at 11 loci (nine microsatellites and two genes of known function) for 10 progeny from each of the 50 families that were released in spring of 1999. In the process of analyzing the inheritance data, we detected several mutations at two of the microsatellite loci (*SSA408* and *OGO1c*), indicating that these loci have particularly high mutation rates. Furthermore, at *SSA408* the mutations detected were not distributed randomly among families. Rather, clusters of identical mutant alleles were found in certain families, suggesting they may have resulted from mutation events occurring very early in gametogenesis, prior to meiosis.

Our results have important significance for the use of microsatellite loci in management. Mutations are expected to have a substantial effect on the amount and pattern of genetic divergence among populations if the mutation rate approaches the rate of migration among populations (see discussion in Allendorf and Seeb 2000). Not surprisingly, the number of mutations detected was correlated with the number of alleles in the sample. We detected mutations at the two loci that have the greatest number of alleles in the parental population (*OGO1c* and *SSA408*). The mutation rate estimates at *OGO1c* and *SSA408* (3.7×10^{-3} and 5.4×10^{-3}) are at the high end of the range of 10^{-3} to 10^{-6} reported for other organisms (Dallas 1992, Weber and Wong 1993, Schug et al. 1997). The variability of these two loci makes them powerful tools for assigning parentage. We were able to unambiguously assign parentage to 35 of the 36 returning fish from the 1998 cohort based on these two loci alone. However, given the high probability of mutation at these loci, our results indicate that it is important to use a combination of low and high variability markers for parentage analysis. Our mutation analysis also suggests that *OGO1c* and *SSA408* are inappropriate as markers for analysis of stock structure in pink salmon.

Marine Survival and Fitness Experiment: 1998 cohort

In August 1998, 150 (75 male and 75 female) mature pink salmon were collected from Likes Creek, Resurrection Bay, and transported to the ASLC for controlled matings. We made 75 families of full-sibs by crossing one male and one female. One hundred progeny from each family were collected to test marker inheritance for parentage analysis. We then selected 50 of these families on the basis of egg number and survival during incubation for the release experiment. These families were pooled together into a single tank in March shortly after hatching. In May 1999, approximately 1,500 progeny from each of these 50 single-pair mating families were marked and released from the ASLC facility.

Progeny from this experiment returned in August 2000. We had anticipated a return rate of 2%, for a total of 1,000-2,000 individuals expected to be recovered for genetic and morphological analyses (approximately 30 fish per family). However, no fish returned to the ASLC fish pass, and we captured a total of 36 fish throughout Resurrection Bay. These 36 fish were placed into 30 families on the basis of 10 microsatellite loci. This sample size was too small to answer the questions that we are addressing.

Marine Survival and Fitness Experiment: 1999 cohort

We repeated this experiment with odd-year pink salmon in August 1999. We collected 68 adults (34 females and 34 males) from Likes Creek, and released their marked progeny from the ASLC in May 2000. This cohort returned in the summer of 2001. We used a different experimental mating scheme with these fish to allow a more powerful genetic analysis of the progeny. Each male and each female was crossed with two individuals in a series of 2 x 2 diallel crosses (Figure 3).

We collected 259 sexually mature adults from this cohort that returned to Resurrection Bay in summer 2001 (Figure 4). We identified the parents of these fish based upon genotypes at eight microsatellite loci and a growth hormone locus. The distribution of returning progeny (i.e., reproductive success) of the 34 males and females is close to expected with random reproductive success (Table 5; Figure 5). The only apparent exception is that there are 2 males that each produced an exceptionally large number of returning progeny.

The results from quantitative genetic analysis of these progeny are described in Appendix E.

DISCUSSION

Evaluation of Even-year Families for Mapping

We placed a gene of known function, *MHC β 2*, on the even-year linkage map. This gene is currently linked to two PINE loci and one of a duplicated microsatellite locus (*STR60-2*). Unfortunately, *STR60-2* is not mapped on our more comprehensive odd-year map. Further work is necessary in order to place *MHC β 2* on the odd-year map.

There are two classes of *MHC* genes, class I and class II. Class I *MHC* is involved with the ability of the body to recognize altered "self" cells and *MHC* class II is involved in recognizing foreign invaders. Studies of the organization of *MHC* suggest that the class I and II regions are not linked in bony fishes (Sato et al. 2000). In addition, this gene is a candidate for analysis in the marine survival and fitness experiment.

Comparative mapping

Pink salmon are unique in that they exhibit a rigid two-year life cycle that has resulted in two reproductively isolated odd- and even-year lineages (Aspinwell 1974). Beacham et al. (1988) found substantial allozyme and morphological evidence for differentiation of alternate brood years. In addition, Phillips and Kapuscinski (1988) and Phillips et al. (1999) detected chromosomal rearrangements between odd- and even-year populations that occur in the same geographical area. Furthermore, in a recent experimental study, Gharrett et al. (1999) demonstrated outbreeding depression in crosses between the two year classes. Together, these findings suggest that the alternate brood years are reproductively isolated and genetically distinct.

Our genetic analysis of the odd and even-year stocks from Likes Creek did not detect any differences in recombination fraction at linked loci between year class. The comparison of both haploid linkage data and gene-centromere distances between odd and even-year classes support findings that gene order is highly conserved (Graf 1989). The similarity in gene order between these two year classes also supports the incorporation of

results from the even-year map onto the more comprehensive odd-year map. Finer resolution mapping with a greater number of loci is necessary to determine the existence and location of any differences between these year classes.

Of the 41 microsatellite loci on the pink salmon linkage map, 27 are included on the rainbow trout map (Sakamoto et al. 2000). Two of these loci included on our map are one of a duplicated pair in pink salmon, but are only known to have a single copy in the rainbow trout map. It is unknown which copies are included on our map. A comparison of the odd year pink salmon linkage map and the rainbow trout map (Sakamoto et al. 2000) is discussed in Appendix 4.

Differences in recombination rates

The analysis of recombination rates in pink salmon detected large differences between individuals. Sakamoto et al. (2000) suggest that this might be a result of ancestral tetrasomic inheritance and pseudolinkage. When homeologous chromosomes pair and exchange material, the resulting homologous chromosomes are less similar to each other than when homologous pairing occurs (Allendorf and Danzmann 1997). Presumably this makes it more difficult in subsequent generations for pairing and exchanges to occur resulting in a lower rate of recombination in those individuals produced from parents in which multivalent pairing occurred.

Previous studies in salmonids have detected differences in recombination rates between males and females (Wright et al. 1983, and Sakamoto et al. 2000). Due to large differences detected between individuals within each sex we compared the average recombination rate of females to that of males at each locus. Initial results agree with Sakamoto et al. (2000); females have a higher recombination rate at loci located close to the centromere ($y < 0.17$; Table 3). Due to our small data set we are unable to draw conclusions for loci that are farther from the centromere ($y > 0.71$).

In tetraploid species such as pink salmon, it has been suggested that the difference in recombination rate between sexes is due to constraints imposed on crossing-over during multivalent pairing (Sakamoto et al. 2000). Multivalent pairing has only been reported in males and generally occurs in the telomeric region (Wright et al. 1983; Allendorf and Danzmann 1997). It has been suggested that multivalent pairing in males explains the tendency for males to have a higher rate of recombination than females in telomeric regions. Recombination in the telomeric regions of males can occur between homologous and homeologous chromosomes increasing the chance for exchange in that region.

Mutation Analysis

Our inheritance dataset revealed a great deal of heterogeneity in mutation rates and patterns among the nine microsatellite loci analyzed (Appendix D). All mutations detected, both in embryos from the 1998 cohort, and returning adults from the 1999 cohort were at two of the nine loci (*OGO1c* and *SSA408*). These two loci are, by far, the

most variable examined in this study, both in number of alleles and in length. It seems likely that the high rates of mutation at these loci are responsible for their high levels of genetic variation. These two loci are also the only tetranucleotide repeats; the other seven loci are dinucleotide repeats. All mutant alleles detected differed from the parental allele by four base pairs which is suggestive of addition or deletion of a single repeat unit. The mutation rate estimates at *OGO1c* and *SSA408* are at the high end of the range of 10^{-3} to 10^{-6} reported for other organisms (Dallas 1992, Weber and Wong 1993, Schug et al. 1997). Investigation of the adult returns from the 1999 cohort yielded similar estimates of mutation rates for both *OGO1c* and *SSA408*.

The variability of these two loci makes them powerful tools for assigning parentage. We were able to unambiguously assign parentage to 35 of the 36 returning fish from the 1998 cohort based on these two loci alone. However, given the high probability of mutation at these loci, our results indicate that it is important to use a combination of low and high variability markers for parentage analysis.

Mutations at *SSA408* were not distributed randomly among families, but rather tended to be clustered within families. This pattern of inheritance suggests that a high proportion of novel alleles resulted from mutations occurring early in gametogenesis. Clustering of mutations within single families has been shown to bias estimates of mutation rates and to influence basic population genetic processes such as fixation probabilities (Woodruff et al. 1996). Another potential source of bias we detected at *SSA408* was the tendency for mutations to increase allele size and for particular alleles to be hypermutable. The variability of mutations within and among loci and among families suggests that mutation should not be ignored when interpreting patterns of genetic differentiation (e.g., when conducting stock structure analysis). Loci with a high mutation rates violate the customary assumption that the effect of mutation is negligible, and may be less useful in estimating gene flow and historical patterns of isolation because these signals will be obscured by the accumulation of mutations. Certainly, if data from both highly polymorphic and less polymorphic loci are being combined, the possibility for locus-specific effects should be evaluated.

2000 returns

In August and September 2000 no fish returned to the ALSC fish pass. We expected most of our returning population to detect and be drawn toward the freshwater signal at the ASLC. However, due to the failure of the fish pass we were forced to survey freshwater streams in upper Resurrection Bay for marked pink salmon using seine nets. We also relied on recreational fisherman to turn in marked pink salmon. Though we were able to collect 36 marked pink salmon this sample is too small to test for correlation between genes and fitness traits. One problem with the fishpass was that the amount of freshwater the facility was releasing was probably inadequate for the returning adults to detect.

2001 returns

Increased outflow from the fish pass at the ASLC in August and September 2001 was likely responsible for successfully attracting fish. However, it did not actually catch any fish which necessitated other means of retrieving them. The 260 fish recovered is slightly greater than one percent of the released fry. We have only recently completed genetic analysis of these fish at nine of ten loci examined in the 1999 parents. This has enabled us to unambiguously assign the fish into parental families, and observe trends of heritability of length for this cohort.

The near random distribution of returns among families indicates that the influence of selection favoring some families over others is limited. The index of variability (mean family size divided by variance) tends toward 1 under random survival, and increases with deviation from random survival (as families tend to survive or perish as a unit; Crow and Morton, 1955). The index of variability value of 1.82 in this cohort indicates some departure from complete random survival but is much lower than the values of 4.03 and 4.97 found by Geiger et al. (1997) in the two cohorts of pink salmon in which they were able to detect a significant sire effect on survival.

The relatively high heritability of length found in both males and females (0.38 and 0.41; Appendix E) is similar to the values (0.4 and 0.2) found in pink salmon from Auke Creek, AK, released into the wild by Smoker et al. (1994). Heritability of length combined with random family returns suggests that, at least under the oceanic conditions this cohort experienced, inherited body length had little effect on marine survival. However, this does not address the effect of length on mating success since these fish were mated in captivity.

CONCLUSIONS

We constructed odd- and even-year linkage maps that can be used to test for effects of regions of the genome on traits that are important to the recovery of pink salmon (e.g., growth and survival) and to evaluate stock structure. We placed a gene of known function on the even-year map, *MHCBA2*. Comparisons between odd and even-year maps have not detected any differences in gene order.

Understanding the genetic basis of phenotypic variation is essential for predicting the direction and rate of phenotypic evolution. We estimated heritabilities and genetic correlations of morphological (fork length, pectoral and pelvic fin ray counts, and gill arch raker counts) and life history (egg number and individual egg weight) traits of pink salmon from Likes Creek, Alaska, in order to characterize the genetic basis of phenotypic variation in this species. Families were created from wild-caught adults, raised to the fry stage in the lab, released into the wild, and caught as returning adults and assigned to family using microsatellite loci and a growth hormone locus.

Morphological traits were all moderately to highly heritable, but egg number and egg weight were not heritable, suggesting that past selection has eliminated additive genetic variation in egg number and egg weight or that there is high environmental

variance in these traits. Genetic correlations were similar for non-adjacent morphological traits and adjacent traits. Genetic correlations predicted phenotypic correlations fairly accurately, but some pairs of traits with low genetic correlations had high phenotypic correlations, and vice versa, emphasizing the need to use caution when using phenotypic correlations as indices of genetic correlations.

ACKNOWLEDGMENTS

We thank the Alaska Department of Fish & Game Gene Conservation Lab for their assistance throughout this project. We thank Brian Mullaly, Susan Inglis, Jon Moreland, and Annette D'Alessandro and the staff and volunteers at the Alaska SeaLife Center for their assistance in capturing, raising, and sampling fish. We thank Barry Stratton and the ADF&G Gene Conservation Lab for their continued assistance in this project. Thanks to David Knudsen, Kevin Sage for assistance in capturing and creating the 1999 cohort. Thanks also to Ellie Steinberg, Megan Skinner, Kristin Bott, Jim Gallea, Joe Meng, Aaron Maxwell, Andrew Whiteley and the rest of the crew at the Wild Trout and Salmon Genetics Lab, University of Montana, Missoula. In addition, we thank the Prince William Sound Aquaculture Corporation and the Valdez Fisheries Development Association for providing adult fish and gametes. We also thank D. Knudsen, K. Sage, B. Mullaly, S. Inglis, J. Moreland, R. Hocking, and A. D'Alessandro and the staff of the Alaska SeaLife Center for their assistance in capturing and raising fish. We thank S. Forbes, M. Skinner, and J. Wenburg for comments on the manuscript, R. Leary, M. Steinberg, and A. Whiteley for helpful discussions, and G. Chambers for his suggestions and references on the process of mutations in microsatellites. E.K. Steinberg was supported by a postdoctoral fellowship from the National Science Foundation.

LITERATURE CITED

- Allendorf, F. W. 1983. Isolation, gene flow, and genetic differentiation among populations. In *Genetics and conservation*. Eds. C. Schonewald-Cox, S. Chambers, B. MacBryde, and L. Thomas, pp. 51-65. Menlo Park, CA. Benjamin/Cummings.
- Allendorf, F.W., and R. G. Danzmann. 1997. Secondary tetrasomic segregation of *MDH-B* and preferential pairing of homeologues in male rainbow trout. *Genetics* 145:1083-1092.
- Allendorf, F. W., W. A. Gellman, and G. H. Thorgaard. 1994. Sex-linkage of two enzyme loci in *Oncorhynchus mykiss* (rainbow trout). *Heredity* 72:498-507.
- Allendorf, F.W., and R.F. Leary. 1986. Heterozygosity and fitness in natural populations of animals. In: *Conservation Biology: The Science of Scarcity and Diversity*. M. Soule', editor. Sinauer Assoc. pp. 57-76.
- Allendorf, F. W., and S. R. Phelps. 1981. Use of allelic frequencies to describe population structure. *Can. J. Fish. Aquat. Sci.* 38: 1507-1514.
- Allendorf, F.W., and L.W. Seeb. 2000. Concordance of genetic divergence among sockeye salmon populations at allozyme, nuclear DNA, and mtDNA markers. *Evolution* 54: 640-651.

- Allendorf, F. W. and G. H. Thorgaard. 1984. Tetraploidy and the evolution of salmonid fishes. Pages 1-53 in B. J. Turner, ed. *Evolutionary Genetics of Fishes*. Plenum Publishing Corp., New York.
- Altukhov, Y. P., E. A. Salmenkova, V. T. Omel'chenko, G. A. Rubtzova, and Y. E. Dubrova. 1987. Balancing selection as a possible factor maintaining uniformity of allele frequencies of enzyme loci in populations of Pacific Ocean pink salmon *Oncorhynchus gorbuscha* (Walbaum). *Genetika SSSR* 23:1884-1896.
- Altukhov, Yu. 1990. *Population genetics: diversity and stability*. London: Harwood Academic Publishers.
- Altukhov, Yu. P., S. V. Mezherin, E. A. Salmenkova, and V. T. Omel'chenko. 1989. Effect of selective hatchery on adaptive genetic and biological structure of pink salmon *Oncorhynchus gorbuscha* (Walb.) population. *Genetika* 25:1843-1853.
- Altukhov, Yu. P., E. A. Salmenkova, and Yu. P. Kartavtsev. 1991. Relation between allozyme heterozygosity and viability and growth rate of pink salmon. *Cytology and Genetics* 25:47-51.
- Amos, W., S.J. Sawcer, R.W. Feakes, and D.C. Rubinsztein. 1996. Microsatellites show mutational bias and heterozygote instability. *Nature Genetics*, 13, 390-1.
- 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.
- Banks, M.A., M.S. Blouin, B.A. 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* 90:281-288.
- Becker, J. P. Vos, M. Kuiper, F. Salamini, and M. Heun, 1995. Combined mapping of AFLP and RFLP markers in barley. *Mol. Gen. Genet.* 249:65-73.
- Beacham, T. D., R. E. Withler, C. B. Murray, and L. W. Barner. 1988. Variation in bodysize, morphology, egg size, and biochemical genetics of pink salmon in British Columbia. *Transactions of the American Fisheries Society* 117:109-126.
- Biggs, E., T. Baker, M. McGurk, J. E. Hose and R. Kocan. 1991. Injury to Prince William Sound Herring. State/Federal Natural Resources Damage Assessment Draft Preliminary Status Report Unpub. rep. Alaska Department of Fish and Game, Cordova, AK.
- Braat, A.K, T. Zandbergen, S. Van de Water, H.J.Th. Goos, and D. Zivkovic. 1999. Characterization of zebrafish primordial germ cells: morphology and early distribution of *vasa* RNA. *Developmental Dynamics*, 216:153-67.
- Bue, B.G., S. Sharr, and J.E. Seeb. 1998. Evidence of damage to pink salmon populations inhabiting Prince William Sound, Alaska, two generations after the *Exxon Valdez* oil spill. *Trans. Amer. Fish. Soc.* 127:35-43.
- Cronin, M. A., and J. W. Bickham. 1998. A population genetic analysis of the potential for a crude oil spill to induce heritable mutations and impact natural populations. *Ecotoxicology* 7:259-278.
- Dallas, J.F. 1992. Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. *Mammalian Genome*, 3, 452-6.
- Doyle, R. W., C. Herbinger, C. T. Taggart, and S. Lochmann. 1995. Use of DNA microsatellite polymorphism to analyze genetic correlations between hatchery and natural fitness. *Amer. Fish. Soc. Symposium* 15:205-211.

- Dubrova, Y. E., V. N. Nesterov, N. G. Krouchinsky, V. A. Ostapenko, R. Neumann, D. L. Neil, and A. J. Jeffreys. 1996. Human minisatellite mutation rate after the Chernobyl accident. *Nature* 380:683-86.
- Dubrova, Y. E., E. A. Salmenkova, Y. P. Altukhov, Y. F. Kartavtsev, E. V. Kalkova, and V. T. Omel'chenko. 1995. Family heterozygosity and progeny body length in pink salmon *Oncorhynchus gorbuscha* (Walbaum). *Heredity* 75:281-289.
- Ferguson, M.M. 1992. Enzyme heterozygosity and growth in rainbow trout - genetic and physiological explanations. *Heredity* 68:115-122.
- Forbes, S. H., K. L. Knudsen, T. W. North, and F. W. Allendorf. 1994. One of two growth hormone genes in coho salmon is sex-linked. *Proc. Nat. Acad. Sci. USA* 91:1628-1631.
- Gharrett, A. J., W. W. Smoker, R. R. Reisenbichler, S. G. Taylor. 1999. Outbreeding depression in hybrids between odd- and even-broodyear pink salmon *Aquaculture*, 173:117-130.
- Gharrett, A. J., and W. W. Smoker. 1994. Introduction to genetics of subarctic fish and shellfish. *Canadian Journal of Fisheries and Aquatic Sciences* 51, Suppl. 1:1-3.
- Goodier, J. L. and W. S. Davidson. 1994. *Tc1* transposon-like sequences are widely distributed in salmonids. *J. Mol. Biol.* 241:26-34.
- Greene, B. A., and J. E. Seeb. 1997. SINE and transposon sequences generate high-resolution DNA fingerprints, "SINE prints", that exhibit faithful Mendelian inheritance in pink salmon (*Oncorhynchus gorbuscha*). *Molecular Marine Biology and Biotechnology* 6:328-338.
- Habicht, C., W.B. Templin, L.W. Seeb, and J.E. Seeb. 1998. Genetics of pink salmon inhabiting Prince William Sound, *Exxon Valdez* Oil Spill Restoration Project Annual Report (Restoration Project 96196), Alaska Department of Fish and Game, Anchorage, Alaska.
- Hunt, R.L. 1969. Overwinter survival of wild fingerling brook trout in Lawrence Creek, Wisconsin. *J. Fish. Res. Board Can.* 26:1473-1483.
- Hurst, L. D. 1999. The evolution of genomic anatomy. *Trends Ecol. Evol.* 14:108-112.
- Jacob, H.J. and 19 other authors. 1995. A genetic linkage map of the laboratory rat, *Rattus norvegicus*. *Nature Genetics* 9:63-69.
- Jackson, T. R., M. M. Ferguson, R. G. Danzmann, A. G. Fishback, P. E. Ihssen, M. O'Connell, and T. J. Crease. 1998. Identification of two QTL influencing upper temperature tolerance in three rainbow trout (*Oncorhynchus mykiss*) half-sib families. *Heredity* 80:143-151.
- Jin, L., and R. Chakraborty. 1995. Population structure, stepwise mutations, heterozygote deficiency and their implications in DNA forensics. *Heredity* 74:274-285.
- Johnson, K. R., J. E. Wright Jr., and B. May. 1987a. Linkage relationships reflecting ancestral tetraploidy in salmonid fish. *Genetics* 116:579-591.
- Johnson, O. W., F. M. Utter, and P. S. Rabinovitch. 1987b. Interspecies differences in salmonid cellular DNA identified by flow cytometry. *Copeia* 1987:1001-1009.
- Johnson, S. L., M. A. Gates, M. Johnson, W. S. Talbot, S. Horne, K. Baik, S. Rude, J. R. Wong, and J. H. Postlethwait. 1996. Centromere-linkage analysis and consolidation of the zebrafish genetic map. *Genetics* 142:1277-1288.
- Jones, A.G., G. Rosenqvist, A. Berglund, and J.C. Avise. 1999. Clustered microsatellite mutations in the pipefish *Syngathus typhle*. *Genetics* 152:1057-1063.

- Karl, S. A. and J.C. Avise. 1992. Balancing selection at allozyme loci in oysters - Implications from nuclear RFLPs. *Science* 256:100-102.
- Kartavtsev, Y. P. 1992. Allozyme heterozygosity and morphological homeostasis in pink salmon, *Oncorhynchus gorbuscha* (Walbaum) - Evidences from family analysis. *Journal of Fish Biology* 40: 17-24.
- Katagiri, T., I. Hirono, T. Aoki, and M. Sakai. 1996. Isolation of major histocompatibility complex class I cDNA from pink salmon (*Oncorhynchus gorbuscha*). *Developmental and Comparative Immunology* 20:217-228.
- Kido, Y., M. Aono, T. Yamaki, K. Matsumoto, S. Murata, M. Saneyoshi, and N. Okada. 1991. Shaping and reshaping of salmonid genomes by amplification of tRNA-derived retroposons during evolution. *Proc. Natl. Acad. Sci. USA* 88: 2326-2330.
- Kocher, T. D., W. J. Lee, H. Sobolewska, D. Penman, and B. McAndrew. 1998. A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). *Genetics* 148:1225-1232.
- Koehn, R.K., W.J. Diehl, and T. M. Scott. 1988. The differential contribution of glycolysis and protein catabolism to the relationship between heterozygosity and growth rate in the coot clam, *Mulina lateralis*. *Genetics* 118:121-130.
- Lander, E.S. and 6 other authors. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174-181.
- Lander, E.S., and Schork, N.J. 1994. Genetic dissection of complex traits. *Science* 265:2037-2048.
- Leary, R.F., F.W. Allendorf, and K.L. Knudsen. 1987. Differences in inbreeding coefficients do not explain the association between heterozygosity at isozyme loci and developmental stability in rainbow trout. *Evolution* 41:1413-1415.
- Leary, R.F., F.W. Allendorf, and K.L. Knudsen. 1989. Genetic differences among rainbow trout spawned on different days within a single spawning season. *Progressive Fish-Culturist* 51:10-19.
- Levinson, G. and G.A. Gutman. 1987. Slipped strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution*. 4:203-221.
- Lewontin, R.C. 1991. 25 Years ago in genetics - electrophoresis in the development of evolutionary genetics - milestone or millstone? *Genetics* 128: 657-62.
- Lie, O., A. Slettan, F. Lingaas, I. Olsaker, I. Hordvik, and T. Refstie. 1994. Haploid gynogenesis: A powerful strategy for linkage analysis in fish. *Anim. Biotech.* 5:33-45.
- Lindner, K.R., J. E. Seeb, C. Habicht, K.L. Knudsen, E. Kretschmer, D. J. Reedy, P. Spruell, and F. W. Allendorf. 2000. Gene-centromere mapping of 312 loci in pink salmon by half-tetrad analysis. *Genome* 43: 538-549.
- Longwell, A. C. 1977. A genetic look at fish eggs and oil. *Oceanus* 20(4):46-58.
- Lynch, M., and B. Walsh. In preparation. *Genetics and Analysis of Quantitative Traits*. Volume 2: Evolution and Selection of Quantitative Traits. Sinauer Assoc. Sunderland, MA. http://nitro.biosci.arizona.edu/zbook/volume_2/vol2.html
- Maheswaran, M., P. K. Subudhi, S. Nandi, J. C. Xu, A. Parco, D. C. Yang, and H. Huang, 1997. Polymorphism, distribution and segregation of AFLP markers in a doubled haploid rice population. *Theor. Appl. Genet.* 94:39-45.

- Malkin, D. 1994. Germline p53 mutations and heritable cancer. *Annual Reviews in Genetics* 28:443-465.
- Matova, N. and L. Cooley. 2001. Comparative aspects of animal oogenesis. *Developmental Biology* 231, 291-320.
- McGregor, A. J., S. Lane, M. A. Thomason, L. A. Zhitovsky, W. W. Smoker, and A. J. Gharrett. 1998. Migration timing, a life history trait important in the genetic structure of pink salmon. *N. Pac. Anadr. Fish. Comm. Bull.* 1:262-273.
- Miller, K. M., and Withler R. E. 1998. The salmonid class I MHC: limited diversity in a primitive teleost. *Immunological Reviews* 166:279-293.
- Mironov, O. G. 1969. The development of some Black Sea fishes in seawater polluted by petroleum products. *Probl. Ichthyol.* 9(6):1136-1139.
- Moles, A., M. M. Babcock and S. D. Rice. 1987. Effects of oil exposure on pink salmon (*Oncorhynchus gorbuscha*) alevins in a simulated intertidal environment. *Marine Environment Research* 21:49-58.
- Moran, P., J. I. Izquierdo, A. M. Pendas, and E. Garciazvazquez. 1997. Fluctuating asymmetry and isozyme variation in Atlantic salmon: Relation to age of wild and hatchery fish. *Transactions of the American Fisheries Society* 126:194-199.
- Morizot, D. C., S. A. Slaugenhaupt, K. D. Kallman, and A. Chakravarti. 1991. Genetic linkage map of fishes of the genus *Xiphophorus* (Teleostei, Poeciliidae). *Genetics* 127:399-410.
- Mousseau, T. A., K. Ritland, and D. D. Heath. 1998. A novel method for estimating heritability using molecular markers. *Heredity* 80:218-224.
- Naruse, K. and 19 additional authors. 2000. A detailed map of Medaka, *Oryzias latipes*: Comparative genomics and genome evolution. *Genetics* 154: 1773-1784.
- Nelson, D.L., S.A. Ledbetter, L. Corbo, M.F. Victoria, R. Ramirez-Solis, T.D. Webster, D.H. Ledbetter, and C.T. Caskey. 1989. Alu polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources. *Proc. Nat. Acad. Sci. USA* 86:6686-6690.
- O'Brien, S. J., J. E. Womack, L. A. Lyons, K. J. Moore, N. A. Jenkins, and N. G. Copeland. 1993. Anchored reference loci for comparative genome mapping in mammals. *Nature Genetics* 3:103-112.
- Pamilo, P., and S. Pálsson. 1998. Associative overdominance, heterozygosity and fitness. *Heredity* 81:381-389.
- Phillips, R. B., and A. R. Kapuscinski. 1988. High frequency of translocation heterozygotes in odd-year populations of pink salmon (*Oncorhynchus gorbuscha*). *Cytogenetics and Cell Genetics* 48:178-182.
- Phillips, R. B., M. P. Matsuoka, W. W. Smoker, and A. J. Gharrett. 1999. Inheritance of a chromosomal polymorphism in odd-year pink salmon from southeastern Alaska. *Genome*. 42:816-820.
- Pogson, G. H., K. A. Mesa, and R. G. Boutilier. 1995. Genetic population structure and gene flow in the Atlantic cod *Gadus morhua*: A comparison of allozyme and nuclear RFLP loci. *Genetics* 139:375-385.
- Postlethwait, J.H. and 14 other authors. 1994. A genetic linkage map for the zebrafish. *Science* 264:699-703.
- Primmer, C.R., H. Ellegren, N. Saino, and A.P. Moller. 1996. Directional evolution in germline microsatellite mutations. *Nature Genetics*, 13, 391-3.

- Roush, W. 1997. A zebrafish genome project? *Science* 275:923.
- Roy, N. K., J. Stabile, J. E. Seeb, C. Habicht, and I. Wirgin. 1999. High frequency of K-ras mutations in pink salmon embryos experimentally exposed to Exxon Valdez oil. *Environmental Toxicology and Chemistry* 18:1521-1528.
- Sakamoto, T., R. G. Danzmann, K. Gharbi, P. Howard, A. Ozaki, S. K. Khoo, R. A. Woram, N. Okamoto, M. M. Ferguson, L. Holm, R. Guyomard, and B. Hoyheim. 2000. A microsatellite linkage map of Rainbow Trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. *Genetics* 155: 1331-1345.
- Sakamoto, T., R. G. Danzmann, N. Okamoto, M. M. Ferguson, and P. E. Ihssen. 1999. Linkage analysis of quantitative trait loci associated with spawning time in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 173:33-43.
- Schug, M.D., T.F.C. Mackay, and C.F. Aquadro. 1997. Low mutation rates of microsatellite loci in *Drosophila melanogaster*. *Nature Genetics*, 15, 99-102.
- 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.* 1:300-318.
- Seeb, J.E., C. Habicht, W.B. Templin, and L.W. Seeb. 1996. Genetics of pink salmon inhabiting Prince William Sound, *Exxon Valdez* Oil Spill Restoration Project Annual Report (Restoration Projects 94320D and 95320D), Alaska Department of Fish and Game, Anchorage, Alaska.
- Shimoda, N., E. W. Knapik, J. Ziniti, C. Sim, E. Yamada, S. Kaplan, D. Jackson, F. deSavage, H. Jacob, M. Fishman. 1999. Zebrafish genetic map with 2000 microsatellite markers. *Genomics* 58:219-232.
- Shum, B. P., K. Azumi, S. Zhang, S. R. Kehrer, R. L. Raison, H. W. Detrich, and P Parham. 1996. Unexpected β 2-microglobulin sequence diversity in individual rainbow trout. *Proc. Natl. Acad. Sci. USA* 93: 2779-84.
- Slatkin, M. 1995. Hitchhiking and associative overdominance at a microsatellite locus. *Mol. Biol. Evol.* 12:473-480.
- Spruell, P. and G. H. Thorgaard, 1996. SINE sequences detect DNA fingerprints in salmonid fishes. *Heredity* 317-324.
- Spruell, P., K. L. Pilgrim, B. B. Greene, C. Habicht, K. Knudsen, K. R. Lindner, J. Olsen, G. K. Sage, J. E. Seeb, and F. W. Allendorf. 1999. Inheritance of nuclear DNA markers in gynogenetic haploid pink salmon (*Oncorhynchus gorbuscha*). *Journal of Heredity* 90:289-296.
- Smoker, W.W., A.J. Gharrett, and M.S. Stekoll. Manuscript. Genetic variation in timing of anadromous migration within a spawning season in a population of pink salmon. Presented at Proc. Int. Symp. on Biol. Interactions of Enhanced and Wild Salmonids, Nanaimo, B.C. 1992.
- Stanley, J. G. 1983. Gene expression in haploid embryos of Atlantic salmon. *J. Hered.* 74:19-22.
- Unwin, M. J. 1997. Fry-to-adult survival of natural and hatchery-produced chinook salmon (*Oncorhynchus tshawytscha*) from a common origin. *Canadian Journal of Fisheries and Aquatic Sciences* 54:1246-1254.

- 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.
- Van der Beek, S., and Van Arendonk, J.A.M. 1993. Criteria to optimize designs for detection and estimation of linkage between marker loci from segregating populations containing several families. *Theor. Appl. Genet.* 86:269-280.
- Wada, H., K. Naruse, A. Shimada, and A. Shima. 1995. Genetic linkage map of a fish, the Japanese medaka *Oryzias latipes*. *Molec. Marine Biol. and Biotech.* 4:269-274.
- Waples, R. S. 1995. Evolutionary significant units and the conservation of biological diversity under the Endangered Species Act. *American Fisheries Society Symposium* 17:8-27.
- Weber, J.L. and C. Wong. 1993. Mutation of human short tandem repeats. *Human Molecular Genetics* 2:1123-1128.
- Wierdl, M., M. Dominska, and T.D. Petes. 1997. Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics*, 146, 769-79.
- Woodruff, R. C, H. Huai, and J. N. Jr. Thompson. 1996. Clusters of identical new mutations in the evolutionary landscape. *Genetica* 98:149-160.
- Young, W. P., P. A. Wheeler, V. H. Coryell, P. Keim, and G. H. Thorgaard. 1998. A detailed linkage map of rainbow trout produced using doubled haploids. *Genetics* 148:1-13.
- Zhivotovsky, L. A., K. I. Afanasiev, and G. A. Rubtzova. 1987. Selection for enzyme loci in the pink salmon (*Oncorhynchus gorbuscha* Walbaum). *Genetika* 23:1876-1883.
- 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 (*Oncorhynchus gorbuscha*). *Canadian Journal of Fisheries and Aquatic Sciences* 51, Suppl. 1:223-232.
- Zouros, E.W., and D. Foltz. 1986. The use of allelic isozyme variation for the study of heterosis. In: M.C. Rattazzi, J.G. Scandalios, and G.S. Whitt (eds.), *Isozymes: Current Topics in Biological and Medical Research, Volume 13*. Alan R. Liss, New York, New York.

OTHER REFERENCES

Personal Communications

Mark Daly, Whitehead Institute, 9 Cambridge Center, Cambridge MA 02142.

Roy Danzmann, Department of Zoology, University of Guelph, Guelph, Ontario N1G
2W1 Canada.

Takashi Sakamoto, Department of Fisheries, Tokyo University, Tokyo, Japan.

Table 1. A breakdown of linkage groups based on number of markers in each group and the average size in cM.

Number of Markers	Number of groups	Average size (cM)
1-5	10	23.26
6-10	14	53.51
11-15	9	118.50
16-20	6	181.27
21-25	1	189.20
26-30	1	243.70
31-35	2	263.05
36-40	0	-----
41-45	0	-----
46-50	0	-----
over 50	1	457.40

Table 2. Summary of polymorphic loci detected by four different techniques.

	Number of polymorphic loci	Percent assigned to linkage group	Percent of loci codominant
AFLP	393	91	2.5
PINE	162	96	4.9
Micro	35	100	91.4
Allozymes	13	100	100
Total	603		

Table 3. Summary of linkages in normal diploid families between allozymes and microsatellites.

Loci	Family	Informative Parent	N	r	χ^2 1 df
<i>sAAT3 - FH</i>	A14	F	86	0.337	9.12
<i>sAAT3 - sMDHB1,2</i>	A14	F	89	0.112	53.49
<i>sAAT4 - μ Str60</i>	A104	F	21	0.238	5.76
<i>ADA2 - PGDH</i>	A120	M	56	0.125	31.50
<i>ADA2 - SSA197</i>	A103	F	42	0.024	38.10
	A120	M	18	0.111	10.89
<i>CKC2 - μ Str60</i>	A120	F	46	0.348	4.26
<i>FH - MDHB1,2</i>	A14	F	86	0.291	15.07
<i>bGALA - G3PDH1</i>	V2	M	75	0.346	7.05
<i>GDA - PEPD2</i>	A8	M	82	0.012	78.05
	A20	M	95	0.105	59.21
	A29	M	45	0.000	45.00
<i>G3PDH1 - PEPLT</i>	V5	M	75	0.240	20.28
<i>GPIB1,2 - PEPD2</i>	V2	M	75	0.013	71.05
<i>sIDHP2 - Ost1</i>	A29	M	41	0.366	2.95
	A104	F	33	0.303	5.12
<i>PGDH - Ssa197</i>	A120	M	20	0.050	12.20

Table 4. Comparison of interference estimates based on half-tetrad and haploid data from nine linkage groups. The number in parenthesis corresponds to the relative amount of interference for each method. The half-tetrad interference estimates are grouped into no(0), low (1-3), moderate (4-6), high (7-9) or in some cases complete interference. The haploid data range is grouped into no (0), low (0.01-0.33), moderate (0.34-0.66), high (0.67-0.99), and complete (1.0) interference.

Linkage Group	Estimated degree of Interference	
	Half-tetrad	Haploid
2	High (9)	Moderate (0.46)
5	No (0)	Moderate (0.47)
8	Complete	Complete (1.00)
25	High (9)	No (0.00)
27	Complete	High (0.67)
33	High (8)	Complete (1.00)
34	High (9)	Complete (1.00)
40	High (8)	Low (0.23)
57	?	Complete (1.00)

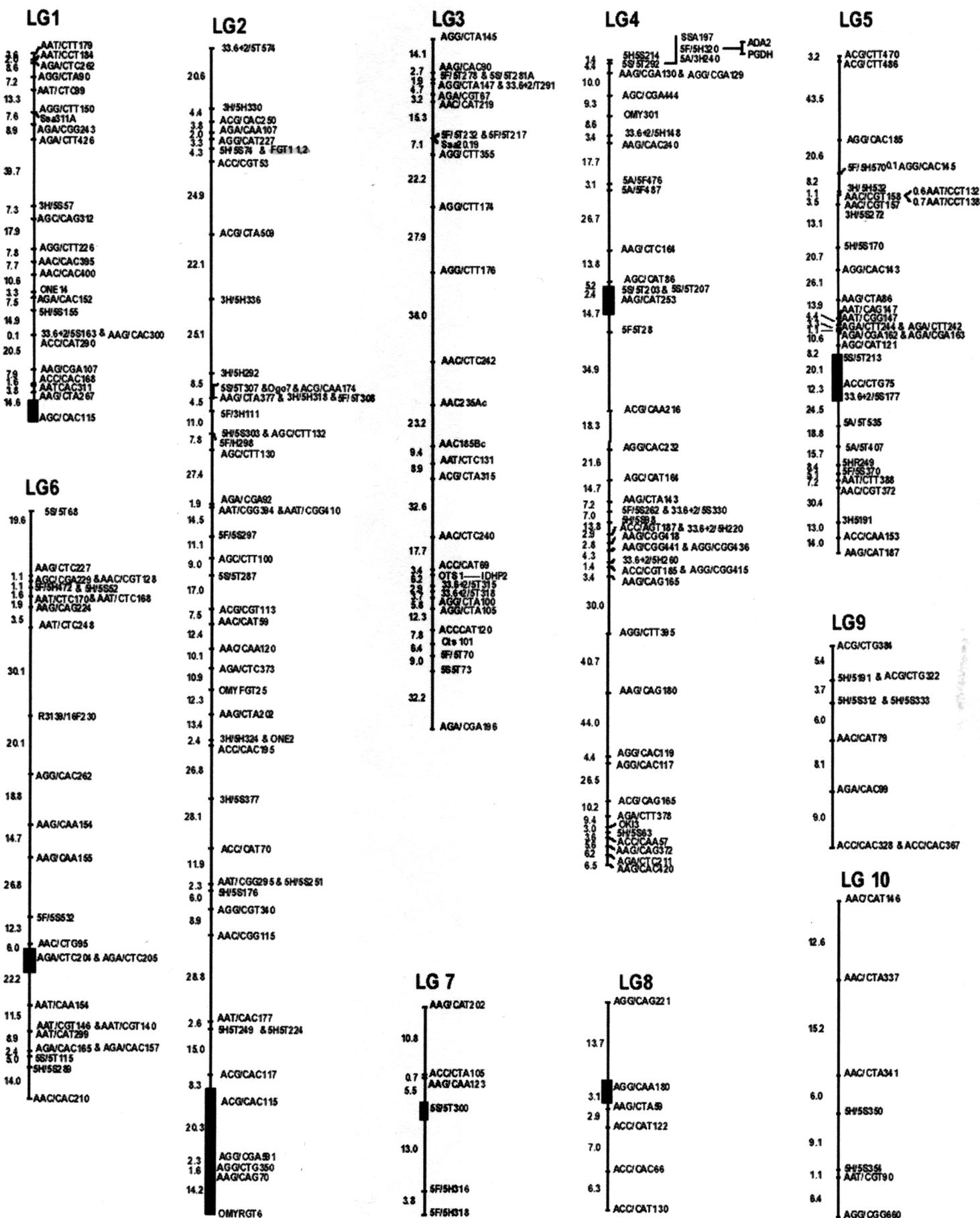
Table 5. Summary of adult progeny recovered from each family and parent, the number of alevins from each parent pooled prior to freshwater rearing, and the percentage of alevins from each parent recovered as adults. See Figure 1 for an explanation of parental cross schemes and family designations.

Fam. No.	Progeny per family		Dam	Progeny	%Survival	Alevins	Sire	Progeny	%Survival	Alevins
	A	B								
1	9	3	1	12	0.84	1424	101	18	1.44	1246
2	9	6	2	15	1.14	1321	102	9	0.60	1499
3	2	1	3	3	0.77	391	103	9	1.44	624
4	7	6	4	13	1.57	828	104	7	1.17	596
5	4	2	5	6	0.80	754	105	9	1.13	797
6	5	2	6	7	0.85	821	106	4	0.51	778
7	0	1	7	1	0.22	450	107	6	0.83	721
8	6	9	8	15	1.48	1017	108	10	1.34	746
9	4	2	9	6	1.61	374	109	7	2.40	291
10	3	6	10	9	1.97	456	110	8	1.48	539
11	0	1	11	1	0.17	597	111	8	1.14	702
12	8	3	12	11	1.21	911	112	4	0.50	805
13	2	11	13	13	2.07	628	113	6	1.22	490
14	4	7	14	11	0.96	1149	114	18	1.40	1287
15	2	4	15	6	0.57	1057	115	6	0.65	929
16	4	7	16	11	1.29	854	116	11	1.12	983
17	4	1	17	5	0.57	877	117	12	1.31	918
18	8	5	18	13	1.65	788	118	6	0.80	747
19	3	2	19	5	0.73	687	119	14	1.25	1124
20	11	5	20	16	1.77	905	120	7	1.50	468
21	3	5	21	8	0.76	1058	121	9	0.77	1171
22	6	0	22	6	0.41	1465	122	5	0.37	1351
23	4	2	23	6	0.55	1096	123	8	0.97	827
24	4	3	24	7	0.95	740	124	5	0.50	1010

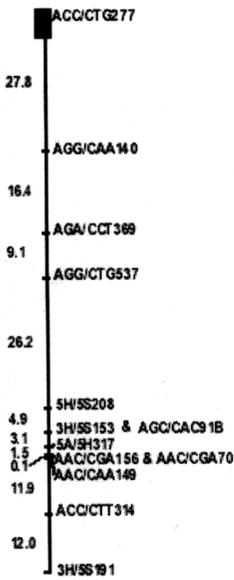
Table 5. (continued)

Fam. No.	Progeny per family		Dam	Progeny	%Survival	Alevins	Sire	Progeny	%Survival	Alevins
	A	B								
25	6	3	25	9	0.97	925	125	9	0.97	927
26	3	5	26	8	0.89	904	126	8	0.89	901
27	2	2	27	4	0.43	941	127	6	0.58	1041
28	4	3	28	7	0.62	1125	128	5	0.49	1025
29	2	0	29	2	0.59	341	129	7	1.07	653
30	5	0	30	5	1.29	388	130	0	0.00	75
31	1	6	31	7	0.93	753	131	2	0.28	706
32	1	4	32	5	0.49	1017	132	10	0.94	1064
33	1	1	33	2	0.86	231	133	4	1.14	352
34	3	1	34	4	0.71	567	134	2	0.45	447
Total	259			259		27841		259		27841
Avg.	3.81			7.62	0.93	819		7.62	0.93	819

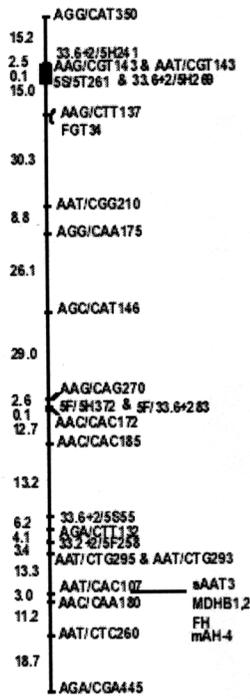
Figure 3. Genetic linkage map of pink salmon based on the inheritance of 602 polymorphic loci. Numbers to the left indicate recombination rates (cM). Locus names are to the right. Centromeres are indicated by black rectangles.



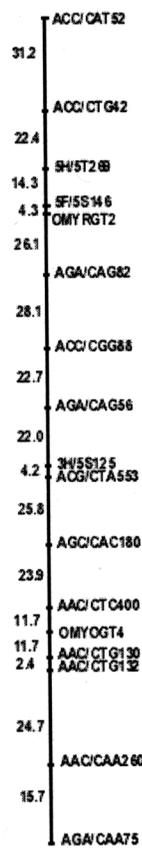
LG11



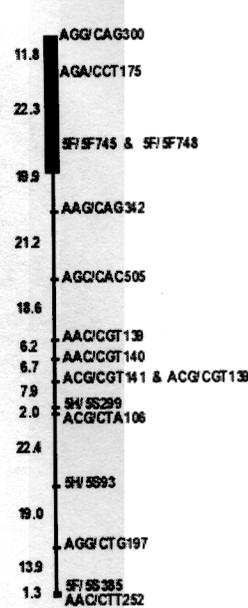
LG12



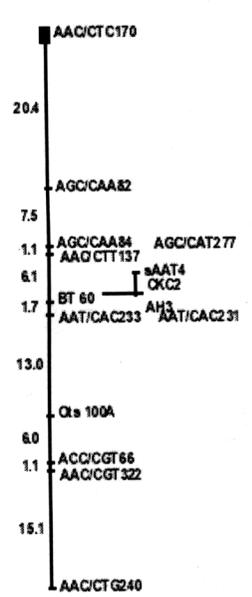
LG13



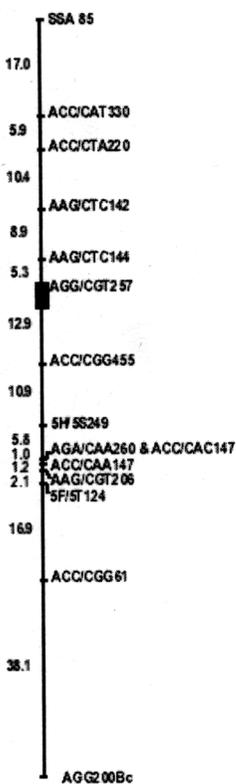
LG14



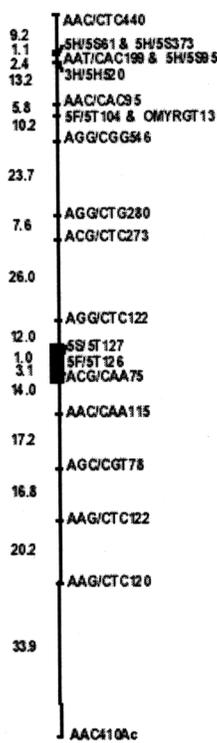
LG15



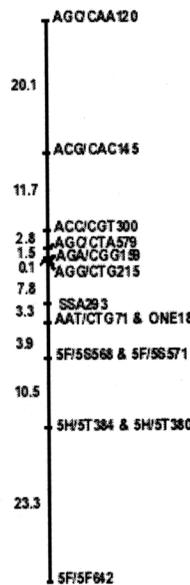
LG16



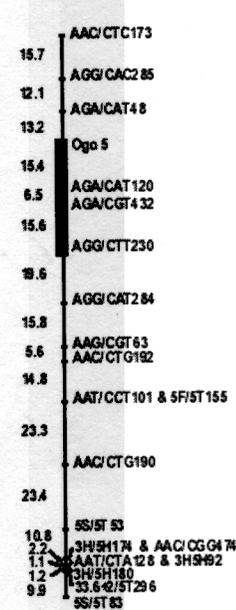
LG17



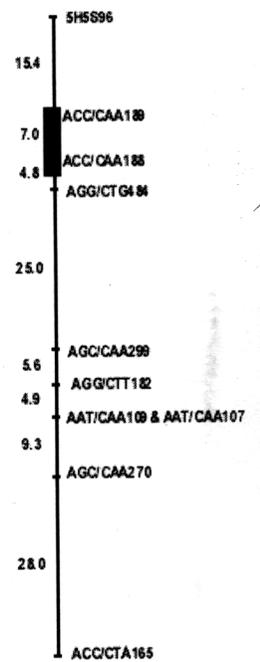
LG18



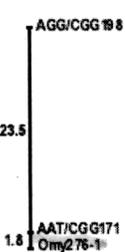
LG19



LG20



LG21



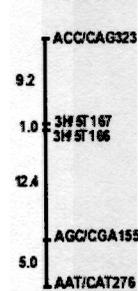
LG22



LG23

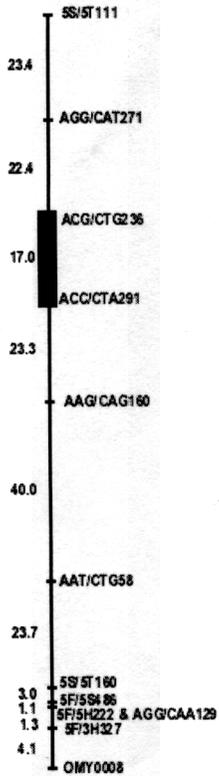
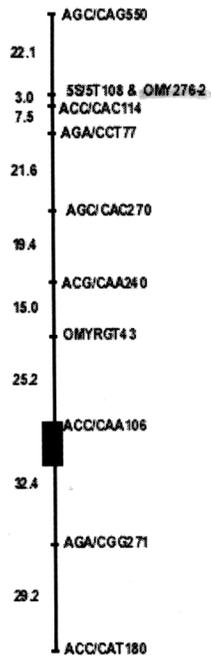
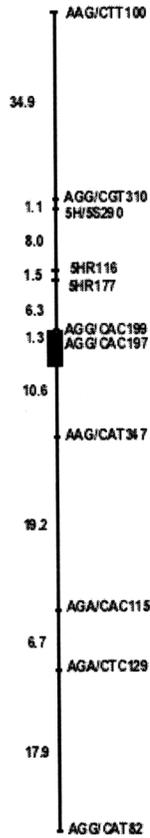
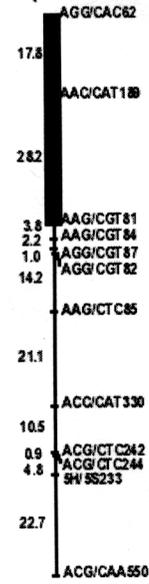
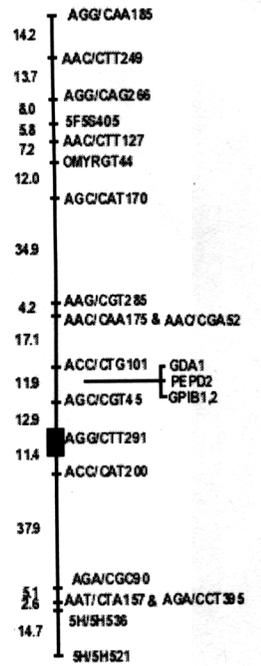
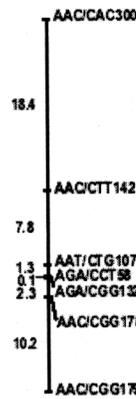
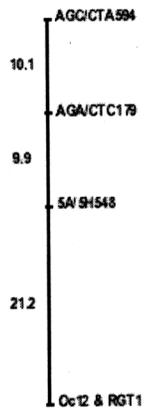
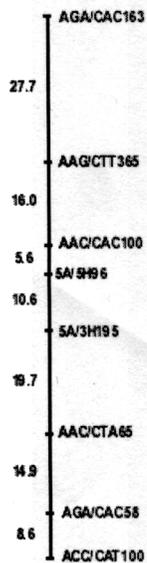
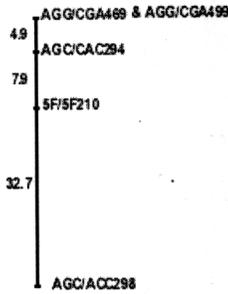
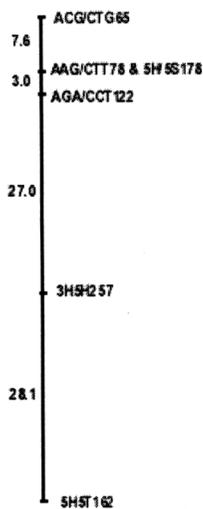
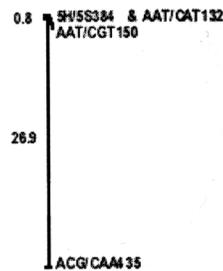
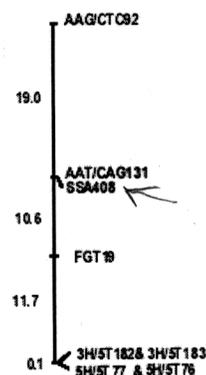
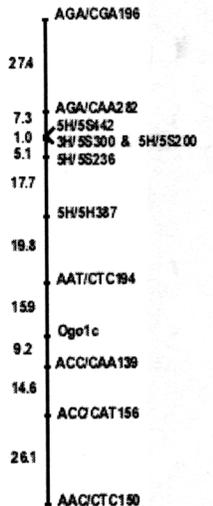
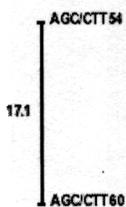
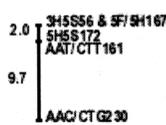
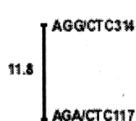


LG24



LG25



LG26**LG27****LG28****LG29****LG30****LG34****LG35****LG31****LG32****LG33****LG36****LG37****LG38****LG39****LG40****LG42****LG41**

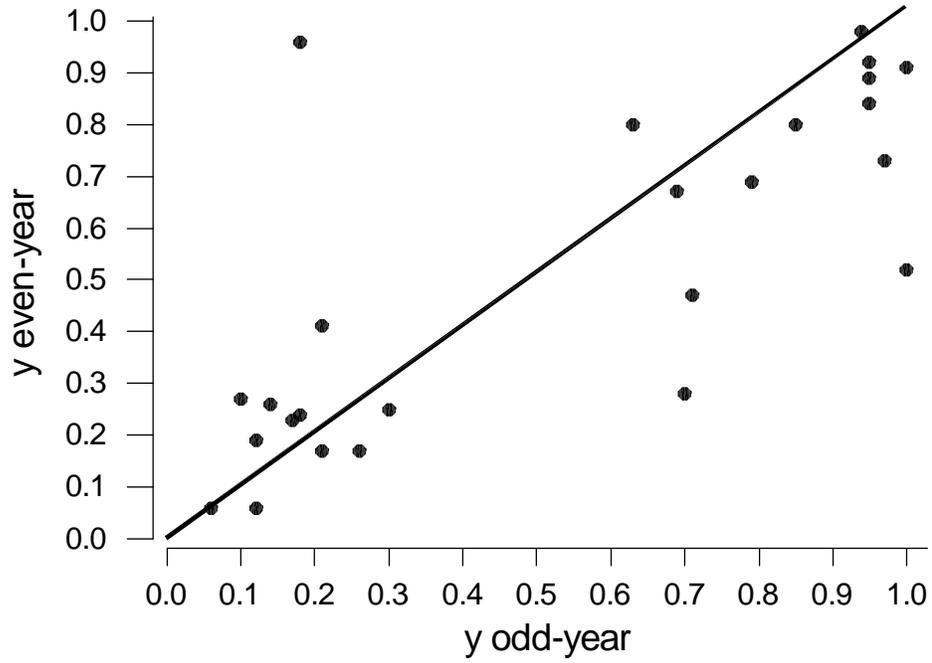
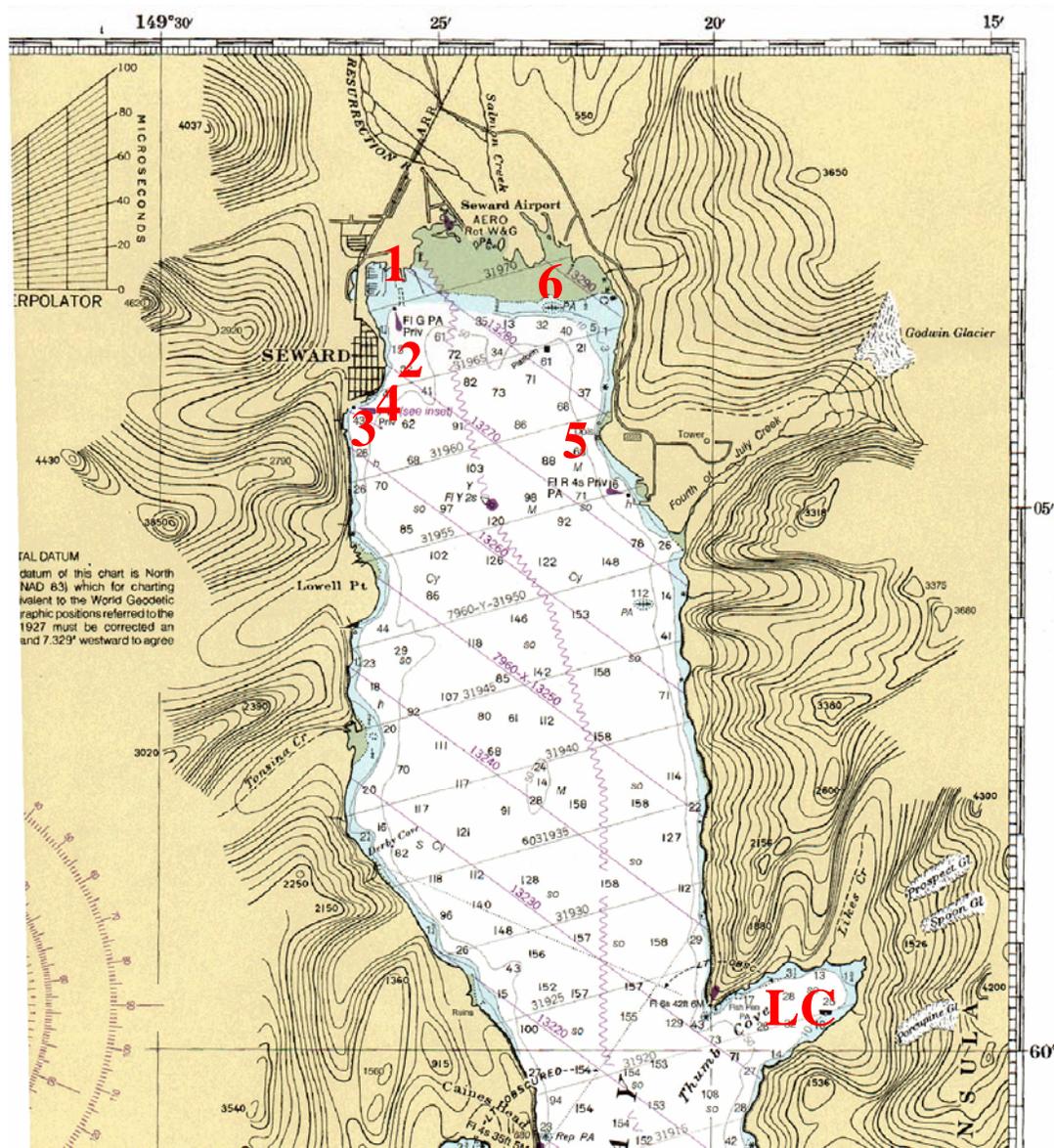


Figure 2. Comparison of gene-centromere distances (y) for odd and even year pink salmon. The gray line indicates the expectation if y values are equivalent in both years.

Figure 3. Diagram of our half-sib family experimental design. Numbers across the top represent females, numbers down the side represent males. The squares contain the family designation and the numbers of individuals used to make each family. The number of the dam is the family number in each half-sibling cross. The letter A or B for each family designates the sire. The parents of family 5A are dam 5 and sire 105, 5B are dam 5 and sire 106, 6A dam 6 and sire 105, and 6B are dam 6 and sire 106, and so forth.

		FEMALES									
		01	02	03	04	05	06	...	33	34	
MALES	101	1A 01x101	2A 02x101								
	102	1B 01x102	2B 02x102								
	103			3A 03x103	4A 04x103						
	104			3B 03x104	4B 04x104						
	105					5A 05x105	6A 06x105				
	106					5B 05x106	6B 06x106				
	.							.			
	.							.			
	.							.			
	133									33A 33x133	34A 34x133
	134									33B 33x134	34B 34x134

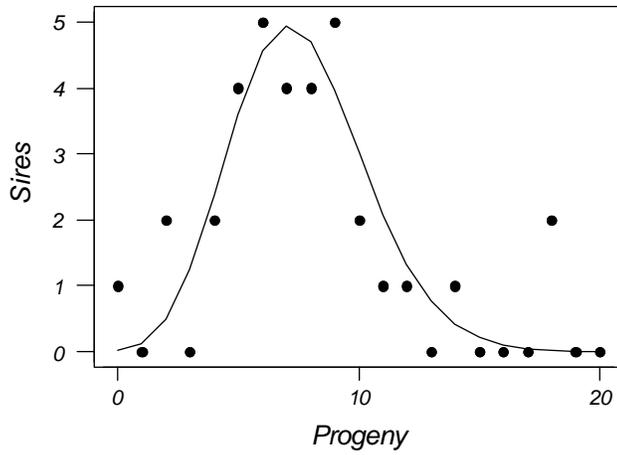
Figure 4. Map of Resurrection Bay. Numbers indicate the location where experimental fish were collected in 2001 as designated below. LC designates Likes Creek, the location where the parents were collected.



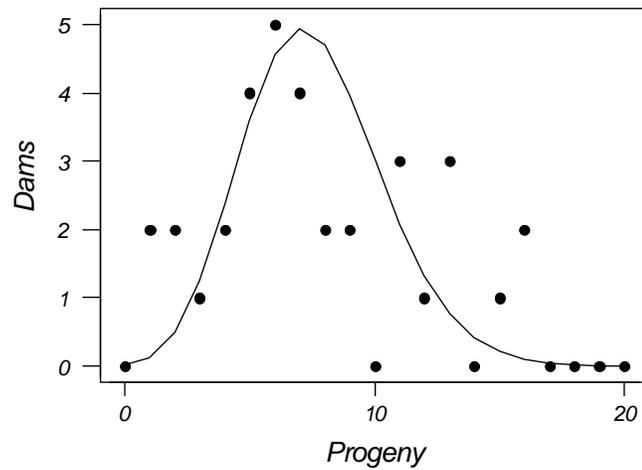
Map Number	Location	Number of fish collected
1	Boat Harbor	2
2	Culvert	31
3	Lowell Cr.	26
4	ASLC	182
5	Spring Cr.	4
6	Resurrection R.	10
	Unknown	4
	Total	259

Figure 5. Scatterplot of the number of progeny returning for sires (a) and dams (b). For example, there were two sires that produced 10 returning progeny. The line is the expected Poisson distribution.

(a)



(b)



A. Pilgrim, K. 1999. Identification of the sex-determining locus in pink salmon. M.A., Organismal Biology and Ecology, University of Montana.

The Search for Sex-Linked Markers in Pink Salmon
(*Oncorhynchus gorbuscha*)

by

Kristine Laurel Pilgrim

B.A., The University of Montana, 1996

presented in partial fulfillment of the requirements
for the degree of Master of Science
The University of Montana
1999

Approved by:

Dr. Fred Allendorf, committee chair

Dean, Graduate School

Date

Pilgrim, Kristine L., M.S., March 1999 Organismal Biology and Ecology

Searching for Sex-Linked Markers in Pink Salmon (*Oncorhynchus gorbuscha*)

Committee Chair: Dr. Fred W. Allendorf

The process of sex chromosome evolution has been studied in several model organisms, and much has been inferred based on detecting and mapping sex-linked markers in other organisms (e.g. mammals). Salmonids, like many other species, are in a primitive state of sex-chromosome evolution, where the X and Y chromosomes are not highly divergent.

This research uses pink salmon (*Oncorhynchus gorbuscha*) as a model salmonid in which to detect sex-linked genomic markers. This study is part of an ongoing project to map the pink salmon genome, in order to assess possible genetic changes resulting from the March 1989 *Exxon Valdez* oil spill.

The first chapter provides a summary of the process of sex chromosome evolution. In addition, it provides background information on what is known about sex-linked markers in salmonids, and why these markers are important.

The second chapter focuses on detecting and determining whether a growth hormone pseudogene is Y-linked in pink salmon. The third intron (intron C) from growth hormone 2, and the growth hormone pseudogene were characterized in pink salmon. Sequence data from other salmonid species is also incorporated in this chapter for phylogenetic analysis to determine when the pseudogene arose within the *Oncorhynchus* lineage.

The third chapter uses the method of bulked segregant analysis to detect a random genetic marker linked to the sex-determining region in pink salmon. Bulked segregant analysis, used in other mapping studies to target a particular region of the genome, is used in conjunction with AFLP and PINE primers to target the sex-determining region of the Y-chromosome in pink salmon.

Table of Contents

Abstract.....	page ii
List of Tables.....	page iv
List of Figures.....	page v
CHAPTER I: The Search for Sex-Linked Markers in Pink Salmon (<i>Oncorhynchus gorbuscha</i>).....	page 1
CHAPTER II: Detection of a Y-Linked Pseudogene in Pink Salmon (<i>Oncorhynchus gorbuscha</i>).....	page 12
CHAPTER III: The Search for a Y-Linked Marker in Pink Salmon (<i>Oncorhynchus gorbuscha</i>) using Bulked Segregant Analysis of AFLP and PINE Markers.....	page 47
Literature Cited.....	page 77

LIST OF TABLES:

CHAPTER II

Page

- 34 **Table 1.** Sequences from nine salmonid species used in phylogenetic analysis of intron C from growth hormone 2 (*GH-2*), and the growth hormone pseudogene (*GH-2p*).
- 35 **Table 2.** Observed and expected genotypes for *GH-2* in pink salmon odd and even-year samples for the two alleles (*GH-2*C446* and *GH-2*C527*) observed at this locus.
- 36 **Table 3.** Genotypes and allele frequencies of known male and female pink salmon at *GH-2*C*
- 37 **Table 4.** Inheritance of *GH-2*C* in gynogenetic haploid pink salmon.

CHAPTER III

- 69 **Table 1.** Inheritance of two microsatellite loci in progeny from A14.
- 70 **Table 2.** Table showing the numbers of individuals from odd and even year families used in the study.
- 71 **Table 3.** Primers and references for PINE primers
- 72 **Table 4.** Results from pooled DNA samples using eight AFLP and one PINE primer combination.
- 73 **Table 5.** Results from individuals using eight AFLP and one PINE primer combination.

LIST OF FIGURES:

CHAPTER II:

Page

- 38 **Fig. 1.** Aligned nucleotide sequence of intron C and a portion of flanking exons from *GH-2* and *GH-2p* from pink salmon.
- 39 **Fig. 2.** Aligned nucleotide sequence of intron C from *GH-2* from pink and sockeye salmon, showing the 81 bp *SmaI* SINE insert and a diagram of the full *SmaI* element.
- 40 **Fig. 3.** Agarose gel showing PCR products from male, female, and gynogenetic haploid pink salmon amplified using the *GH-2p* specific primers.
- 41 **Fig. 4.** Aligned nucleotide sequence of intron C from *GH-2* and *GH-2p* from nine salmonid species used for phylogenetic analysis.
- 45 **Fig. 5.** Parsimony tree generated using intron C sequence data from *GH-2* and *GH-2p* from nine salmonid species.
- 46 **Fig. 6.** Published phylogeny of *Oncorhynchus* species.

CHAPTER III:

Page

- 74 **Fig. 1.** Diagram showing sex chromosome composition in pooled DNA samples.
- 75 **Fig. 2.** Diagram showing procedure for pooling DNA for bulked segregant analysis.
- 76 **Fig. 3.** Acrylamide gel showing AFLP fragments in pooled DNA from males and females from odd and even-year pink salmon.

The Search for Sex-Linked Markers in Pink Salmon
(*Oncorhynchus gorbuscha*)

Kristine L. Pilgrim

*Division of Biological Sciences, University of Montana, Missoula,
MT 59812.*

Several mechanisms controlling sex determination have evolved in plants and animals. Many marine invertebrates and plants are cosexual (Charlesworth 1991). In contrast, most terrestrial animal species are dioecious, where male and female gametes are produced by separate individuals. In such animals, an individual's sex may be determined by environmental cues, the presence or absence of egg fertilization (haplo-diploidy; e.g. ants, bees, and wasps), the presence of a sex determining gene or group of genes (polyfactoral), or by specialized sex chromosomes (Bull 1983, Rice 1996). Sex chromosomes are more common in animals than in plants and are believed to have evolved from autosomes (Rice 1996). Vertebrate sex chromosomes appear to have evolved independently in many lineages (Rice 1996, Saxena et al. 1996, Fridolfsson et al. 1998).

The process of Y-chromosome evolution begins when the progenitor Y-chromosome evolves a dominant, genic factor that determines gender. This factor may be a single, dominant gene or a collection of tightly linked genes. There is subsequent lack of recombination between the X-chromosome and the region containing the sex determining factor on the Y-chromosome (Charlesworth 1991, Rice 1996). Recombination is further suppressed due to the evolution of sexually antagonistic alleles that are tightly linked with the sex determining factor (Fisher 1931, Bull 1983, Rice 1987, Rice 1997).

Once suppressed recombination has occurred, the sex determining region is expected to accumulate mutations and chromosomal changes due to relaxed selection. Presumably, the X-chromosome contains functional copies of genes also found on the Y-chromosome that compensate for Y-linked gene degeneration. Additionally, the Y-chromosome will never be in the homozygous state in an individual for selection to eliminate the deleterious changes. Mutation rates are expected to be higher for Y-linked genes compared to X-linked genes due to the greater number of germ cell divisions required for spermatogenesis relative to oogenesis (Haldane 1947).

In addition to higher mutation rates, the Y-chromosome may also accumulate different types of repetitive elements.

Transposable elements have been found on the Y-chromosome in *Drosophila simulans* and *D. melanogaster* (Junakovic et al. 1998), and a short interspersed nuclear element (SINE) has been reported in *Zfy* (zinc finger of the Y) in seven domestic cat lineages (Slattery and O'Brien 1998). In humans and other higher apes, an *Alu* element has been detected at the boundary between the pseudoautosomal and sex-chromosome specific region of the Y-chromosome (Ellis et al. 1990).

The differential segment on the Y-chromosome containing the sex determining region continues to grow due to the accumulation of mutations, repetitive elements, and sexually antagonistic alleles, and lack of recombination with the X-chromosome. Eventually, a chromosomal sex determining system is reached once the differential Y segment constitutes most of the chromosome except for the pseudoautosomal region that pairs with and undergoes recombination with the X-chromosome during meiosis (Burgoyne 1986, Ellis et al. 1990).

Sex determination in fish species is rather complex. Hermaphroditism is present in species from at least 14 families of teleost fishes (Moyle and Cech 1996). In most cases, these species are transitional or sequential hermaphrodites, most commonly where a female turns into a male (often when the dominant male of a harem dies). Environmental sex determination,

where the temperature of the water determines the sex of the individuals is also present in many species. The first evidence for a genetic basis of sex determination came via the discovery of a sex-linked color factor in the medaka (*Oryzias* (formerly *Aplocheilus) latipes* ; Aida 1921). Later, studies with the guppy (*Poecilia reticulatus*), found that ornamental characteristics in male guppies were controlled by 18 genes, 17 of which were perfectly linked with the sex determining factor (Winge 1927).

The evolution of a chromosomal sex determining system has occurred independently and at different times in several fish species (Kirpichnikov 1981, Price 1984, Nanda et al. 1992, Rice 1996). Hermaphroditism appears to have been the ancestral state in fish, with a genic sex determining system evolving later (Kirpichnikov 1981). Previous work suggests that many of the teleost fishes, including salmonids, are in a primitive state of chromosomal sex determination.

Fish in the family Salmonidae descended from a single tetraploid ancestor 25-100 million years ago (Allendorf and Thorgaard 1984). Salmonids have been found to have a XX female, XY male sex determining system based on chromosomal studies (Thorgaard 1977, Thorgaard and Gall 1979) and breeding studies with sex-reversed fish (Okada et al. 1979, Johnstone et al. 1979). In triploid rainbow trout (*Oncorhynchus mykiss*), XXY fish

are males, providing evidence for a dominant Y sex determining factor (Allendorf and Thorgaard 1984).

Unlike the mammalian system of full chromosomal sex determination with a small pseudo-autosomal region, salmonids show few genetic differences between the sex chromosomes. The viability of YY males in coho salmon (*O. kisutch*; Hunter et al. 1982), and the viability and fecundity of YY hermaphrodites in rainbow trout (Chevassus et al. 1988) support the existence of functional genes on the Y-chromosome, and slight genetic differentiation between the X and Y. Due to the lack of genetic differences observed, recombination between the X and Y chromosome must still occur. Indeed, recombination between the sex chromosomes was observed in the cross used for a linkage map study in rainbow trout (Young et al. 1998).

Sex chromosomes in the family Salmonidae are largely isomorphic. While the karyotypes of many species in the Salmonidae family are currently known, most species show little difference between the X and Y chromosomes in both staining and morphology (Frolov 1993). The sex chromosomes in rainbow trout (Thorgaard 1983), lake trout (*Salvelinus namaycush*; Phillips and Ihssen 1985, Frolov 1993), and sockeye salmon (*O. nerka*; Thorgaard 1978, Frolov 1990) have shown some differentiation cytogenetically. However, the inability to distinguish the X and

Y chromosomes in the majority of species indicates the sex chromosomes are in an early state of evolution.

Little is understood about the sex determining region (*SEX*) in salmonids, and relatively few sex-linked markers are known. Earlier studies in rainbow trout placed *SEX* near the centromere through gene-centromere mapping (Allendorf et al. 1986) and cytogenetic studies (Thorgaard 1977, Lloyd and Thorgaard 1988). A linkage map recently developed for rainbow trout mapped *SEX* to a more distal position on the short arm of a subtelocentric chromosome (Young et al. 1998). *SEX* is also thought to reside on the short arm of a subtelocentric chromosome in lake trout (Reed et al. 1995). Using inheritance data to calculate recombination rates, two sex-linked allozyme loci, *HEX-2* and *sSOD1* were detected in rainbow trout (Allendorf et al. 1994). In some chinook salmon (*O. tshawytscha*) populations, the allozyme *PEPB-1* is sex-linked based on inheritance data (Marshall, et al. in preparation). In addition, joint segregation analysis on sparctics, a hybrid between brook trout (*S. fontinalis*) and Arctic char (*S. alpinus*), detected tight linkage between *SEX* and three allozyme loci, *LDH-1*, *AAT-5*, and *GPI-3* in brook trout (May et al. 1989).

In addition to allozymes, other sex-linked markers have been detected in salmonids. A male-specific DNA probe (OtY1) has been

developed for chinook salmon using subtractive DNA hybridization (Devlin et al. 1991). Further work has shown that the OtY1 probe is located within a tandemly repeated 8 kb unit comprising about 2.4 Mb located on the Y-chromosome (termed OtY8; Devlin et al. 1998). The repeat unit, or similar sequences, are present in all closely related species but are not sex-specific (Devlin et al. 1998).

In brown trout, (*Salmo trutta* L.) tight linkage between *SEX* and the minisatellite probe *Str-A9*, developed from a brown trout DNA phagemid library, was detected using joint segregation analysis (Prodohl et al. 1994). In rainbow trout, fluorescent *in situ* hybridization (FISH) was used to demonstrate an X-linked copy of 5s rDNA (Moran et al. 1996). Using FISH, females exhibited 4 hybridization signals, while males had 3, indicating 5s rDNA is duplicated and located on both an autosomal metacentric chromosome pair and on the subtelocentric X-chromosome (Moran et al. 1996). Finally, a Y-linked growth hormone pseudogene has been detected in some *Oncorhynchus* species (Forbes et al. 1994, Kavsan et al. 1994, Du et al. 1993).

The majority of sex-linked loci observed in salmonid species thus far appear to be largely taxon-specific. The apparent lack of sex-linked orthologous loci suggest that either these loci are monomorphic in closely related taxa and sex-linkage can not be

observed in inheritance studies, or taxon-specific karyotype differences exist and *SEX* evolved in different genomic regions.

Sex-linked markers can provide important genetic information. Sex-linked markers are useful for identifying the gender of immature fish, or fish that have been hormonally treated. Sex-linked genetic markers can be used to investigate the genealogy and phylogeny of species, for gene mapping, and to detect geographic population structures. Hybridization between species can often be detected through Y-introgression of markers.

Allele frequencies at sex-linked loci differ from autosomal loci, and in salmonids, where diploidization is taking place (Allendorf and Danzmann 1997a), tetrasomic versus disomic inheritance can be evaluated with sex-linked loci. Sex-linked markers can also help evaluate rates of recombination between the X and Y chromosomes in males, the X and X chromosomes in females, and can help determine how large the non-combining, sex determining region of the Y-chromosome is.

Many species of Pacific salmon are currently threatened or endangered (Nehlsen et al. 1991, National Research Council 1996, Allendorf et al. 1997b). While pink salmon are not endangered, populations have disappeared from coastal Washington, Oregon, and California (National Research Council 1996), and this species is susceptible to habitat degradation (e.g. the March 1989 *Exxon*

Valdez oil spill). Pink salmon (*O. gorbuscha*) are currently the most abundant anadromous salmon in North America (National Resource Council 1996). They have an extensive range, spawning in Asian and North American streams bordering the Pacific and Arctic oceans, and are the most numerous salmon in the commercial salmon fisheries of this region, important to the fisheries of Japan, Russia, Canada, and the United States (U.S. Fish and Wildlife Service 1989).

Currently, no sex-linked markers have been described in pink salmon. Consequently, the main objective of this thesis research was to detect and describe sex-linked genetic markers in pink salmon. The work presented here discusses the results from a two-year study aimed at detecting sex-linked markers in pink salmon using a variety of molecular methods.

The goal of chapter 2 was to detect and determine if a pseudogene derived from growth hormone 2 is Y-linked in pink salmon. This pseudogene (*GH-2p*) has previously been found to be Y-linked in chinook salmon (Du et al. 1993), coho salmon (Forbes et al. 1994), and chum salmon (*O. keta*; Kavsan et al. 1994). In addition to testing for the presence of the growth hormone pseudogene (*GH-2p*) in pink salmon males, the same region (intron C) of *GH-2* was also investigated in pink salmon to be sure the pseudogene was not being confused with amplification of *GH-2*.

The goal of chapter 3 was to detect a random, genetic, sex-linked marker in pink salmon. DNA was combined from male and female pink salmon respectively in a method known as pooling or bulking. Using individuals from the same family, the pooling method created two pools of DNA that theoretically differ only for the sex determining region of the Y-chromosome. Random genetic markers such as AFLPs (amplified fragment length polymorphic DNAs) and PINES (paired interspersed nuclear elements) were screened on the pooled DNA in order to detect a marker that differs between the pools and is thus linked to the sex determining region.

Detecting a random, sex-linked marker in pink salmon would complement an ongoing genome mapping project of this species (Allendorf et al. 1997c, Allendorf et al. 1998, Spruell et al. 1999). Finding a sex-linked marker such as an AFLP or a PINE may make it possible to identify the sex chromosome linkage group on the map. Yet, even if the marker is not able to be "linked up" to the rest of the markers on the current linkage map, a sex-linked marker detected in pink salmon would be valuable and useful for genetic studies in this species. In addition, such a sex-linked marker could be potentially valuable for application in other Pacific salmon species, or alternatively provide

additional evidence for the lack of sex-linked orthologous loci present within salmonids.

Detection of a Y-Linked Pseudogene in Pink Salmon

(*Oncorhynchus gorbuscha*)

Kristine L. Pilgrim

*Division of Biological Sciences, University of Montana, Missoula,
MT 59812.*

ABSTRACT

I amplified and sequenced a portion of an intron from a Y-linked growth hormone pseudogene in pink salmon (*Oncorhynchus gorbuscha*). This pseudogene provides an unambiguous, PCR-based test for sex in pink salmon, and is the first Y-linked marker described in this species. A duplication event of the functional growth hormone 2 (*GH-2*) gene gave rise to the growth hormone 2 pseudogene (*GH-2p*). The third intron (intron C) of *GH-2* was also amplified and sequenced in pink salmon to be certain that amplification of *GH-2p* was not confounded by *GH-2*. A length polymorphism was detected in intron C of *GH-2* and is due to an 81 bp insertion homologous to the 3' end of the *Sma*I SINE (short interspersed nuclear element). Phylogenetic analysis of intron C sequence data from *GH-2* and *GH-2p* in pink salmon and other salmonids suggests that the gene duplication event giving rise to *GH-2p* occurred after the split of *Salmo* and *Oncorhynchus* but

before rainbow trout and cutthroat trout diverged from the other *Oncorhynchus* species.

INTRODUCTION

Growth hormone (somatotrophin) plays an important role in many physiological and biochemical functions in salmon. Fish growth hormone is a 20-22 kDa single chain polypeptide hormone, expressed in the somatotrophs of the anterior pituitary (Du et al. 1993, Yang et al. 1997). Growth hormone stimulates protein synthesis and promotes lipid and glycogen breakdown (important during migration to spawning sites when the fish do not eat). In addition, growth hormone is the primary regulator of somatic growth and is involved in sexual maturation (Bjornsson 1997). Growth hormone is also important in the anadromous life cycle of salmon, improving hypoosmoregulatory in fresh water and increasing tolerance to sea water (Bjornsson 1997).

Salmonids descended from progenitors in which a single, auto-tetraploidization event occurred between 25-100 million years ago (Allendorf and Thorgaard 1984). As a result of this tetraploid event, salmonids have two functional, autosomal, growth hormone genes, *GH-1* and *GH-2* (Devlin 1993, Forbes et al. 1994, Yang et al. 1997). Both these genes have been sequenced and characterized using cDNA GH clones in a variety of species

including Atlantic salmon (*Salmo salar*; Johansen et al. 1989, Male et al. 1992), coho salmon (*Oncorhynchus kisutch*; Nicoll et al. 1987), chum salmon (*O. keta*; Kawauchi et al. 1986), chinook salmon (*O. tshawytscha*; Du et al. 1993), sockeye salmon (*O. nerka*; Devlin, R. H. 1993), and rainbow trout (*O. mykiss*; Yang et al. 1997).

In comparison with other vertebrates such as mammals and birds, and with other fish species such as carp, flounder, and bass, salmonids are the only group known to have two, separate, functional genes coding for growth hormone (Yang et al. 1997). Both genes are structurally similar to the growth hormone gene of other species, but lack an intron (intron E) that splits the terminal exon of other species, resulting in six exons and five introns in salmonids (Yang et al. 1997). The functions of both growth hormone genes is not precisely known. Two-year old, female rainbow trout have higher levels of *GH-1* mRNA than of *GH-2*, while this difference is not seen in males. Rainbow trout fry also appear to express higher levels of *GH-1* than *GH-2*, although a sex-specific difference can not be determined at this stage (Yang et al. 1997).

In addition to the two, autosomal growth hormone genes, a later gene duplication of *GH-2* gave rise to a growth hormone pseudogene (*GH-2p*; Du et al. 1993). Pseudogenes are DNA segments

with a high degree of homology with functional genes, but contain nucleotide changes that prevent their translation (Li et al. 1981, Li 1997). Pseudogenes have been detected in a wide-range of taxa, providing evidence for the widespread occurrence of both gene duplication, and loss of gene function due to relaxed selective constraints (Li et al. 1981). Pseudogenes can evolve by either a gene duplication event, or can arise by reverse transcription of RNA (termed processed pseudogenes).

The growth hormone pseudogene has been detected in three salmonid species: chinook salmon (Du et al. 1993), coho salmon (Forbes et al. 1994), and chum salmon (Kavsan et al. 1994). The pseudogene has been detected only in the males of these species, and appears to be perfectly linked with the sex determining region on the Y-chromosome. This pseudogene contains introns and therefore evolved through gene duplication rather than reverse transcription. The complete nucleotide sequence of the pseudogene in chinook salmon has been determined (Du et al. 1993). The pseudogene spans 5.5 kb and although it has a proper TATA box, it contains a premature termination codon, a 150-bp deletion in the last half of exon 5, and a wrong splicing signal at the intron A/exon 2 junction which precludes correct splicing and translation (Du et al. 1993).

In chum salmon, the pseudogene has been sequenced up through the third intron (intron C; Kavsan et al. 1994). The pseudogene

shows close homology with the functional *GH-2* from rainbow trout for the first intron, second exon, and second intron, (although nucleotide mutations prevent correct splicing of the first intron, and a premature stop codon is found in the second exon).

The third exon, and beginning of the third intron (intron C) are deleted in this species, and are partially replaced by a 83 bp A-T rich region (Kavsan et al. 1994).

Pink salmon (*O. gorbuscha*) are a model species for genetic study due to their strict two year, semelparous life cycle (Heard 1991) and genetically distinct odd and even year runs of fish throughout their range (Allendorf and Waples 1996). The purpose of this paper was to detect *GH-2p* in pink salmon, and determine if this pseudogene is Y-linked in this species as well and could be useful as a sex-linked marker in this species. Intron C of *GH-2* was also investigated in pink salmon to be sure that detection of the pseudogene was not confounded by amplifying the functional counterpart. The results of amplifying and sequencing a portion of intron C from *GH-2p* and intron C from *GH-2* are presented. Phylogenetic analysis was performed using sequence data from intron C of *GH-2* and *GH-2p* to determine when the pseudogene arose.

MATERIALS AND METHODS

Samples and Haploid Gynogenesis

In August of 1995, gametes and tissues of 31 adult (based on sexual maturity) pink salmon were collected from the Armin F. Koernig hatchery, Prince William Sound, Alaska. Seven families of gynogenetic haploid embryos were produced by sperm inactivation as described by Thorgaard et al. (1983). Sperm from four males were pooled and irradiated with UV light and mixed with the eggs from individual females. Females were numbered and their progeny were designated by year class and the number assigned to the female (e.g. family 95-105). Embryos were incubated until just prior to hatching when they were collected and preserved in ethanol.

Gametes and tissues were also collected from 37 adult pink salmon from the 1996 year class in August 1996, and were used to make normal diploid as well as gynogenetic haploid families. Parents from additional diploid families were also used in analysis of the growth hormone pseudogene. Chum salmon tissue samples taken from throughout their range (10 individuals from each of nine populations) were kindly provided by the Alaska Department of Fish and Game.

DNA was isolated from muscle or liver tissue in the adults, and from the embryos after separation from the yolk sac in the progeny using the PurgeneTM DNA isolation kit (Genta Systems

Inc.) The concentration of DNA was determined using a scanning spectrofluorometer.

GH-2

Primers designed from conserved positions in exons 3 and 4 in coho salmon, chum salmon, and rainbow trout, (Fobes et al. 1994) were used to amplify intron C of *GH-2* in pink salmon. Polymerase Chain Reaction (PCR) was performed in mixtures that contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 5.0 mM MgCl₂, all four dNTPs (each at 0.2 mM), each primer at 0.3 μ M, and 0.08 unit of *Taq* DNA polymerase (Perkin-Elmer/Cetus), and 40 ng of template DNA in a total volume of 15 μ l. The PCR profile was 30 cycles of 92°C for 1 min, 63°C for 1 min, and 72°C for 2 min. PCR products were electrophoresed on 2.0% agarose gels in TAE buffer (Ausubel et al. 1989), containing ethidium bromide at 0.5 μ g/ml.

GH-2 Pseudogene

The forward primer used to amplify a portion of intron C of the Y-linked pseudogene, 5'-TTTCTCTACGTCTACATTCT-3', was kindly provided by R. H. Devlin. This primer was designed from *GH-2* sequence from chinook salmon and lies 324 bp within intron C (Fig. 1). The reverse primer was the same as that used to amplify *GH-2* (Fig. 1). PCR and gel electrophoresis conditions were the same as for *GH-2*, except 5.8 mM MgCl₂ and an annealing temperature of 51°C were used.

Sequencing and Analysis

Four gynogenetic haploid pink salmon derived from female 95-115 were used to sequence intron C of *GH-2*. Two male pink salmon from the 1995 adult sample were used to sequence intron C of *GH-2p*. For sequencing of the pseudogene, PCR products were first run out on a Metaphor agarose gel to separate the pseudogene from the *GH-2* products, and then were subjected to a second round of PCR. PCR products for sequencing were purified from agarose gels using the GENECLAN kit (BIO 101 Inc.) and sequenced by direct automated sequencing (Applied Biosystems Inc.). These sequences, and other intron C sequences from *GH-2* and *GH-2p* from several other salmonids (Table 1) were aligned using Sechencher version 3.1 (Genecodes Corp., Ann Arbor MI).

Phylogeny reconstruction was performed using the methods of parsimony, maximum likelihood, and neighbor joining available within PAUP* version 4.0.0d64 (Swofford 1999). Parsimony analysis was performed using the branch-and-bound algorithm. Bootstrap values represent the percentage of 1000 replicates that supported a particular branch. Branch lengths were analyzed using MacClade version 3 (Maddison and Maddison 1992).

Phylogenetic analysis was performed using nucleotide sequence data from intron C of *GH-2* in nine species, and intron C of *GH-2p* in four species (Table 1). Atlantic salmon was chosen

as the outgroup because the ancient split of *Salmo* and *Oncorhynchus* has been established (Stearley and Smith 1993, McKay et al. 1996, Murata et al. 1996), and this allowed resolution of the *Oncorhynchus* salmonid sequences. Brown trout (*Salmo trutta*) was included to help provide further resolution (excluding this species resulted in 15 most parsimonious trees). Because of the potential problems associated with taxa not having complete sequence information, phylogenetic analysis was also performed after removing the pink salmon and chum salmon *GH-2p* sequences.

RESULTS

GH-2

Intron C of *GH-2* was amplified in 31 pink salmon collected in 1995, and 37 collected in 1996. The *GH-2* intron C primers amplified two products in pink salmon of 540 bp and 621 bp in length. The product lengths correspond to intron C, 17 bp of exon 3, 29 bp of exon 4 (since the priming sites lie within exon 3 and exon 4), and the primers (24 bp each). Sequencing both alleles revealed the intron lengths specifically to be 446 bp and 527 bp in length (Fig. 1). These alleles therefore have been named *GH-2*C446* and *GH-2*C527* to reflect intron length (Spruell et al. 1999; Genbank No. AF075571, AF075572).

For the 1995 year class, seven of 31 individuals were heterozygotes, and the remaining 24 individuals were homozygotes for the *GH-2*C446* allele (Table 2). In the 1996 year class sample, eight of 37 individuals were heterozygotes, and the remaining 29 individuals were homozygotes for the *GH-2*C446* allele (Table 2). There were no individuals homozygous for the *GH-2*C527* allele present in either year class sample. Both odd and even-year class samples were in Hardy Weinberg equilibrium at this locus, (Table 2). A sex-specific difference in genotypes and allele frequencies at *GH-2*C* was not observed (Table 3).

The allelic nature of these products was investigated by looking at segregation patterns in gynogenetic haploids. Two females heterozygous for *GH-2*C* and their gynogenetic haploid progeny were tested for inheritance at *GH-2*C*. Thirty-six haploid progeny from female 95-115 displayed expected 1:1 Mendelian segregation for these alleles (Table 4). Thirty-six haploid progeny from female 95-115 showed deviation from expected 1:1 Mendelian segregation ($p < .05$; Table 4) but this is likely due to sampling.

Aligning sequences revealed that alleles *GH-2*C446* and *GH-2*C527* differed by an 81 bp insert in the *GH-2*C527* allele. This insert is nearly identical to the 3' end of the consensus sequence of the *SmaI* SINE (short interspersed nuclear element;

Fig. 2) found in the genomes of pink salmon and chum salmon (Kido et al. 1991).

Chum Salmon

In addition to pink salmon, 90 chum salmon were tested for the presence of the *Sma*I insert. Using the primers designed to amplify intron C of *GH-2*, a 601 bp product was amplified, which is what is expected according to published *GH-2* chum salmon sequence data (Fig. 4). There was no evidence for a *GH-2* allele containing the *Sma*I insert in any of the chum salmon tested.

***GH-2* Pseudogene**

Known male and female pink salmon (based on sexual maturity) were analyzed using the pseudogene primers. Twelve males and 19 females from the 1995 year class, and 15 males and 22 females from the 1996 year class were tested. Using the pseudogene primers, all known males from both year classes preferentially amplify a 166 bp product from *GH-2p* (Fig. 3). Males also weakly amplify a portion of the *GH-2**C446 and *GH-2**C527 alleles (depending on genotype), and a 213 bp product from intron C of *GH-1* (Fig. 3). Using the pseudogene primers, all known females from both year classes amplify a portion of the *GH-2**C446 and *GH-2**C527 alleles, and a 213 bp product from intron C of *GH-1* (Fig. 3). Gynogenetic haploids generated from a female heterozygous at *GH-2*, segregate for the *GH-2* alleles as

well as amplify a portion of intron C of *GH-1* (Fig. 3) when amplified using these primers.

All 27 known males tested using the pseudogene primers amplified the 166 bp product of *GH-2p*. All 41 known females tested lacked the pseudogene. The 122 bp of intron C and the beginning of exon 4 sequenced from *GH-2p* show close homology with the *GH-2*C446* and *GH-2*C527* alleles (Fig. 1). **Phylogenetic**

Analysis

Phylogenetic analysis was performed using intron C data from *GH-2* and *GH-2p* to determine when the pseudogene arose. All three tree-building methods available within PAUP* version 4.0.0d64 (parsimony, likelihood, and neighbor-joining) produced trees with similar topologies. A single most parsimonious tree was generated using the method of parsimony and is shown in Fig. 5. Analysis performed after removing the pink salmon and chum salmon pseudogene sequences (in case incomplete sequence data was a confounding factor) did not change the tree topology or the relationships of the taxa from the tree in Fig. 5, except that the pseudogene sequences from coho salmon and chinook salmon clustered together. The relationships of coho salmon, chinook salmon, pink salmon, chum salmon, sockeye salmon, rainbow trout and cutthroat trout generated from intron C sequence data of *GH-2*, agree with published trees for these species based on the

combined evidence from intron D of *GH-2*, as well mtDNA NADH dehydrogenase subunit 3 (Fig. 6).

Phylogenetic analysis using intron C of *GH-2* and *GH-2p* results in the pseudogene sequences forming a single clade, branching off before any of the *Oncorhynchus GH-2* sequences. This indicates that the gene duplication event of *GH-2* occurred after the split of *Salmo* and *Oncorhynchus*, but before the divergence of rainbow and cutthroat from the Pacific salmon. While the bootstrap value lends weak support (62) to this branch, the overall tree topology is consistent and suggestive. In comparing the branch lengths for species with sequence data for both *GH-2* and *GH-2p*, branch lengths are longer for the pseudogenes in chinook salmon, pink salmon, and chum salmon (Fig. 5). This is consistent with pseudogenes having a higher rate of nucleotide substitution in comparison with different parts of other genes (Li 1997).

DISCUSSION

GH-2

Growth hormone 1 and 2 have previously been characterized in a number of species including rainbow trout (Yang et al. 1997, Agellon et al. 1988), Atlantic salmon (Johansen et al. 1989, Male et al. 1992), sockeye salmon (Devlin 1993), and chinook salmon

(Du et al. 1993). Using PCR, a single amplicon is produced for intron C at *GH-2* in coho, and chinook salmon. While no allelic differences in intron C of *GH-2* has been detected by length in rainbow trout, SSCP (Single Strand Conformational Polymorphism) analysis has detected up to five alleles in a single population (Bagley and Gall 1998).

Pink salmon have two distinct alleles (based on length) at intron C of *GH-2*. These alleles, *GH-2**C446 and *GH-2**C527 are found in both odd and even-year pink salmon runs (Table 2), and occur in males and females with nearly identical frequencies (Table 3). Both odd and even-year classes showed no significant deviations from expected Hardy-Weinberg proportions (Table 2). Sequencing revealed that the size difference of the alleles is due to a 81 bp insert that is nearly identical to the 3' end of the *SmaI* element.

SINES are a family of repetitive sequences found throughout the eukaryotic genome (Okada 1991). SINES are typically less than 500 bp in length, and are characterized by having a tRNA-related region, and internal RNA polymerase III promoter, and an AT rich region (Okada 1991, Spruell and Thorgaard 1996, Li 1997).

SINES are amplified by reverse transcription, and appear to be inserted irreversibly into various sites throughout the genome (Takasaki et al. 1997).

The two-year life cycle of pink salmon is so strict that odd and even year fish are genetically isolated (Hart 1973, Donnelly 1983), such that pink salmon differ more between odd and even year runs than between fish within odds and evens throughout their range (Phillips and Kapuscinski 1988). The presence and frequency of the *GH-2**C446 and *GH-2**C527 alleles in both year classes suggest the *SmaI* insert occurred before pink salmon year classes evolved.

The *SmaI* element has been found to be restricted to the genomes of pink salmon and chum salmon (Takasaki et al. 1997, Greene and Seeb 1997). It might be expected that the insertion of the 3' end of *SmaI* into *GH-2* occurred in the common ancestor of pink and chum salmon (around 6 million years based on fossil evidence; McKay et al. 1996). However, chum salmon tested from throughout their geographic range provided no evidence for the presence of the *SmaI* insert in *GH-2*. Thus the insertion seems to have occurred after pink salmon diverged from the other *Oncorhynchus* species. A previous study of the interspecific and intraspecific variation of *SmaI* found a lack of shared insertion sites between the genomes of pink salmon and chum salmon (Takasaki et al. 1997). The authors offer several possible explanations for this result such as horizontal gene transfer, introgression, and temporal differences in amplification within lineages. While horizontal gene transfer and introgression can

not be ruled out, it also seems plausible that the *SmaI* family evolved in the common ancestor of pink salmon and chum salmon, and underwent differential amplification and distribution in the genomes after these two species diverged.

It is interesting that only a portion (the 3' end) of the *SmaI* element is found inserted within *GH-2**C527. Previous studies have also observed that some regions of SINEs may be distributed independently of the remainder of the element. Spruell and Thorgaard 1996, detected different DNA fingerprint patterns in 14 species of salmonid fishes when they used probes homologous to different regions of the *HpaI* and *FokI* SINEs. A genetic linkage map of rainbow trout found that the same locus was never detected when probes homologous to alternative ends of the *HpaI* SINE were used (Young et al. 1998). It appears that SINEs are more complex than originally thought. Caution should be taken when using such elements for phylogenetic analysis since their patterns of distribution in the genome and mode of insertion is not fully clear.

***GH-2* Pseudogene**

The growth hormone pseudogene is Y-linked in pink salmon as it is in chinook salmon, coho salmon, and chum salmon. All known pink salmon males tested from both odd and even-year classes strongly amplified a pseudogene-specific product when primers

designed to amplify the pseudogene were used. In addition, males weakly amplify the *GH-2*C* alleles as well as a portion of intron C from *GH-1*, which is expected based on the close homology of *GH-2*, *GH-2p*, and *GH-1* in the priming sites. Female pink salmon amplify one or both *GH-2*C* alleles depending on their genotype, as well as a portion of intron C from *GH-1*.

This Y-linked growth hormone pseudogene is the first known sex-linked marker described in pink salmon. The ability to detect the pseudogene using PCR makes the genetic test for sex in immature pink salmon, and of tissue or DNA samples of unknown origin both cost effective and time efficient. Moreover, the test is extremely reliable and dependable. Since PCR products are amplified in females as well as males, the test for the presence of the pseudogene will not be confounded by a failed PCR reaction.

The nucleotide sequence for the 93 bp from *GH-2p* shows close homology with that of *GH-2*C446* and *GH-2*C527* (Fig. 1). Previous attempts to detect the pseudogene in pink salmon by amplifying the entire intron and using restriction enzymes to differentiate *GH-2*C* from *GH-2p* (as was done in coho salmon; Forbes et al. 1994) were unsuccessful (F. W. Allendorf personal communication; R. H. Devlin personal communication). This indicates that in *GH-2p*, the upstream region of intron C (from where the sequence

begins (Fig. 4), and possibly exon 3 may be deleted in pink salmon as it is in chum salmon.

It is clear that the growth hormone pseudogene is closely linked to the sex determining region on the Y chromosome, although it is unclear how physically close the pseudogene may be. A chromosomal walk initiated from *GH-2p* in chinook salmon produced cosmids containing mostly repetitive DNA that does not show sex-specificity with other salmonid species (Devlin et al. in preparation). Allele frequencies observed for male and female pink salmon at *GH-2*C* are similar despite small sample sizes suggesting *GH-2* is not closely linked with the sex determining region in pink salmon. Previous segregation studies in coho salmon and chinook salmon (Devlin et al. in preparation), demonstrated that *GH-2* is not X-linked in these species, suggesting that the pseudogene may have arisen from a chromosomal rearrangement rather than divergence of one copy of a sex-linked locus.

Phylogenetic Reconstruction

All methods of reconstructing phylogeny based on sequence data from intron C of *GH-2* and *GH-2p* produced the same overall tree topology. The *GH-2* duplication event giving rise to *GH-2p* appears to have occurred after the split of Atlantic and Pacific salmon (at least 19.9 million years ago; McKay et al. 1996), but

before rainbow and cutthroat diverged from the other Pacific salmon. This agrees with the findings of Du et al. 1993 as to the timing of the emergence of the pseudogene. The four pseudogenes form their own clade, suggesting that *GH-2p* likely diverged before the species (i.e. chinook salmon, coho salmon, pink salmon, and chum salmon) diverged from one another.

However, the different rate of nucleotide substitution, and the subsequent long branches in the pseudogene clade may be a potential source of error in reconstructing phylogenies due to long-branch attraction. In long-branch attraction, long branches of a tree may be randomly attracted to each other and confound a phylogenetic signal (Felsenstein 1978). This could be the reason the pseudogene clade clusters closely with the outgroup (Fig. 5), which is also a long branch. However, if there is a problem with long-branch attraction in the data set, the pseudogene clade should also be randomly attracted to the long branch leading to the coho salmon and chinook salmon *GH-2* branch, which is not observed. This suggests a phylogenetic signal is present and the pseudogene arose after *Oncorhynchus* and *Salmo* diverged, but before the seven Pacific salmonids tested diverged from one another.

Attempts to amplify the pseudogene in sockeye salmon, rainbow trout, and cutthroat trout have been unsuccessful to date (Du et al. 1993, Forbes et al. 1994, R.H. Devlin personal

communication). It is possible that the pseudogene is still present in these species, but is unable to be amplified using PCR techniques due to mutations in the priming sites. Indeed, primers used to amplify the pseudogene in coho salmon and chinook salmon failed to amplify the pseudogene in chum salmon due to a deletion in the priming site in exon 3 (Kavsan et al. 1994). Alternatively, the pseudogene may not be detected in these species because the pseudogene has been excised. The apparent absence of *GH-2p* in sockeye salmon may be due to a loss of this region during a Y-autosomal robertsonian translocation (Thorgaard 1978, Frolov 1993).

Pseudogenes are now thought to be extremely common in most taxa (Li 1997). Gene duplication allows a gene to be free of selection, and since a normal gene is still functioning, the loss or change in function of one of the pairs is common. Since pseudogenes are free from selective constraints, all nucleotide substitutions are neutral and will become fixed or lost in a population by genetic drift. Pseudogenes have been shown to evolve more rapidly (i.e. accumulate more mutations) than other regions of the genome (Li 1997). The longer branch lengths (corresponding to the amount of nucleotide substitution) observed for *GH-2p* in chinook salmon, pink salmon, and chum salmon compared with *GH-2*C* in these species (Fig. 5) is consistent with the expectation of higher rates of evolution in pseudogenes.

Results from *GH-2**C** in this study show pink salmon and chum salmon as being sister species, although bootstrap support for this branch is less than 50. The relationships among sockeye salmon, pink salmon, and chum salmon has been debated (see McKay et al. 1996 for review). In the past, sockeye salmon and pink salmon have been placed as sister species based on morphology, and the similarities observed in sequence data between pink salmon and chum salmon were thought to be due to hybridization between these species (Stearley and Smith 1993). Recent analysis of mtDNA (Domanico and Phillips 1995), and analysis of mtDNA combined with nuclear markers (McKay et al. 1996, Domanico et al. 1997, Kitano et al. 1997) provide strong support for a sister relationship between chum and pink salmon (Fig. 6).

Summary

The growth hormone pseudogene is Y-linked in pink salmon. This pseudogene appears to be perfectly linked to the sex determining region in pink salmon (there is no recombination with the X-chromosome) and is therefore an important sex-linked marker in this species. Pink salmon have a length polymorphism in *GH-2**C** that is present in males and females of both odd and even-year classes. The length difference is caused by an insertion of a portion of the *Sma*I SINE into the longer allele, and is specific to pink salmon. The duplication of *GH-2* that led to *GH-*

2p appears to have occurred in the genus *Oncorhynchus* after the spilt from *Salmo*, but before the rainbow trout and cutthroat trouts branched off from the other Pacific salmon.

Table 1. Sequences used in phylogenetic analysis of intron C from growth hormone 2 (*GH-2*), and the growth hormone pseudogene (*GH-2p*).

Species	Common name	Locus	Reference	Genbank No.
<i>Salmo salar</i>	Atlantic salmon	<i>GH-2</i>	Johansen et al. 1989	M21573
<i>Salmo trutta</i>	Brown trout	<i>GH-2</i>		AF005912
<i>O. mykiss</i>	Rainbow trout	<i>GH-2</i>	Agellon et al. 1988	J03797
<i>O. clarki</i>	Cutthroat trout	<i>GH-2</i>		AF005913
<i>O. tshawytscha</i>	Chinook salmon	<i>GH-2</i>	Du et al. 1993	AF005914
<i>O. kisutch</i>	Coho salmon	<i>GH-2</i>	Forbes et al. 1994	U04930
<i>O. keta</i>	Chum salmon	<i>GH-2</i>		D16491
<i>O. gorbuscha</i>	Pink salmon	<i>GH-2*</i>	This study	AF075571
<i>O. nerka</i>	Sockeye salmon	<i>GH-2</i>	Devlin 1993	U14535
<i>O. kisutch</i>	Coho salmon	<i>GH-2p</i>	Forbes et al. 1994	U04931
<i>O. gorbuscha</i>	Pink salmon	<i>GH-2p</i>	This study	
<i>O. tshawytscha</i>	Chinook salmon	<i>GH-2p</i>	Courtesy R.H. Devlin	
<i>O. keta</i>	Chum salmon	<i>GH-2p</i>	Kavsan, et al. 1994	X74157

*Sequence data from allele *GH-2*C446* was used for phylogentic analysis

Table 2. Observed and expected genotypes for *GH-2*C* in both the odd and even year classes of pink salmon. The *GH-2*C446* and *GH-2*C527* alleles have been abbreviated with the length (bp) of the alleles.

Year Class	n	Observed			Expected			x ²	p-val
		<i>446/446</i>	<i>446/527</i>	<i>527/527</i>	<i>446/446</i>	<i>446/527</i>	<i>527/527</i>		
1995	31	24	7	0	24.4	6.2	0.4	.51	p >.1
1996	37	29	8	0	29.4	7.1	0.4	.51	p >.1

Table 3. Genotypes and allele frequencies of known male and known female adult pink salmon from both odd and even-year samples at *GH-2*C*. The *GH-2*C446* and *GH-2*C527* alleles have been abbreviated with the length (bp) of the alleles.

Year Class	Sex	Genotypes			Allele Frequencies	
		446/446	446/527	527/527	f (446)	f (527)
1995	F	14	5	0	.92	.08
	M	10	2	0	.87	.13
1996	F	18	4	0	.91	.09
	M	11	4	0	.87	.13

Table 4. Inheritance of *GH-2*C* in gynogenetic haploid progeny. The *GH-2*C446* and *GH-2*C527* alleles have been abbreviated with the length (bp) of the alleles.

Female	Genotype	Progeny Genotype		χ^2
		446	527	
95-105	446/527	12	24	4.0
95-115	446/527	20	16	0.4

Fig. 1. Aligned nucleotide sequence of intron C and a portion of flanking exons from *GH-2* and *GH-2p* from pink salmon. Unknown sequence is represented by ?; dashes represent gaps introduced to produce optimal sequence alignment. Stars indicate the priming sites for the pseudogene, the walls indicate the transition between exon and intron.

```

GH-2*C446  GAGATCAGAAGAGTTCA |GTAAGTTACCTGGCTGAGACAATCCTCCATGATGCACAATTC  59
GH-2*C527  ..... |.....  59
GH-2p      ????????????????? |????????????????????????????????????????
GH-2*C446  CAACATGAATAATAGGGCATCTCAATTTGAACAATCGATACAACCTTAGTCATTAGTTATT 119
GH-2*C527  .....  119
GH-2p      ?????????????????????????????????????????????????????????????
GH-2*C446  GGGCAAGCAGATCCCCGATTGTCTAAACTCCATGGGTAAATATATACTGTAGATAAGAAG 179
GH-2*C527  .....G.....  179
GH-2p      ?????????????????????????????????????????????????????????????
GH-2*C446  AACCAGCATCATGCATGGTGGAAATTAATCTAGCCATGACAGGGAGTTTTAAATTGTAC 239
GH-2*C527  .....  239
GH-2p      ?????????????????????????????????????????????????????????????
GH-2*C446  ACTTAAAAATCGGCAGGAAAATGTTGCTATAACATCAGTGCCTTCAAAAACAACCACATTT 299
GH-2*C527  .....  299
GH-2p      ?????????????????????????????????????????????????????????????
GH-2*C446  CATAGTCATTGTAAGTAAACCCATCACTCTCTAATTGGCGGTTTCTCTACGTCTACATT 359

```


Fig. 4. Aligned nucleotide sequence of intron C from *GH-2* and *GH-2p*. Unknown sequence is represented by ?; dashes represent gaps introduced to produce optimal sequence alignment.

Atlantic <i>GH-2</i>	GTAAGTTACCTGGCTGAGACAA-TCCTCCATGATGCTAGATTCCAAAA	47
Brown <i>GH-2</i>-.....A.....AT.....G...	47
Rainbow <i>GH-2</i>-.....AC.....C.	47
Cutthroat <i>GH-2</i>-.....AC.....C.	47
Chinook <i>GH-2</i>-.....ACA.....C.	47
Coho <i>GH-2</i>-.....ACA.....C.	47
Chum <i>GH-2</i>-.....ACA.....C.	47
Pink <i>GH-2</i>-.....ACA.....C.	47
Sockeye <i>GH-2</i>G.....-.....ACA.....C.	47
Coho <i>GH-2p</i>G.....-...G-.C....AC.....	46
Pink <i>GH-2p</i>	??	
Chinook <i>GH-2p</i>G.....-...G-.C....AC.....	46
Chum <i>GH-2p</i>	-----	
Atlantic <i>GH-2</i>	TAAATAATAGGGCATCTCAATTTGAACAA-TCGATACAACCTTAGTCAT	94
Brown <i>GH-2</i>-.....	94
Rainbow <i>GH-2</i>	.G.....-.....	94
Cutthroat <i>GH-2</i>	.G.....T.....-.....	94
Chinook <i>GH-2</i>	.G.....G..AT..GTTA..G.G..A	95
Coho <i>GH-2</i>	.G.....G..AT..GTTA..G.G..A	95
Chum <i>GH-2</i>	.G.....-.....	94

Pink GH-2	.G.....-	94
Sockeye GH-2	.G.....G.....-	94
Coho GH-2p	.T.....T.....-	93
Pink GH-2p	??	
Chinook GH-2p	.T.....G.....G.T.....-	93
Chum GH-2p	-----	
Atlantic GH-2	T-AGTTATTGGGCAAGCAGATCCC-GATTGTGTAAACTCCATGGGTAA	140
Brown GH-2	.-.....C.....	141
Rainbow GH-2	.-.....C.....C.....	141
Cutthroat GH-2	.-.....C.....C.....	141
Chinook GH-2	GC...--..C---.---.---.---.....C.....	128
Coho GH-2	GC...--..C---.---.---.---.....C.....	128
Chum GH-2	.-.....C.....C.....C.....	141
Pink GH-2	.-.....C.....C.....	141
Sockeye GH-2	.-.....C.....C.....	141
Coho GH-2p	.-.....CA.....C.....	140
Pink GH-2p	??	
Chinook GH-2p	.-.....C.....C.....	140
Chum GH-2p	----- .TCC.....	18

Chinook GH-2pT.G.....-----	229
Chum GH-2pT.G.....-----	105
Atlantic GH-2	-----	229
Brown GH-2	-----	230
Rainbow GH-2	-----	227
Cutthroat GH-2	-----	230
Chinook GH-2	-----	217
Coho GH-2	-----	217
Chum GH-2	ATCAAATAAAAAATGTAATAAAAAAATAACAAAATAAAATAAATT	285
Pink GH-2	-----	231
Sockeye GH-2	-----	230
Coho GH-2p	-----	229
Pink GH-2p	??	
Chinook GH-2p	-----	229
Chum GH-2p	-----	105

Atlantic GH-2	-----TC-AGCAGTAAAATGTTGCTATACCTCAGTGCTTTCAACT	268
Brown GH-2	-----..-G.....C.....	269
Rainbow GH-2	-----..-G.....AA.....C.....--	264
Cutthroat GH-2	-----..-G.....AA.....C.....--	267
Chinook GH-2	-----..A.....C.....	257
Coho GH-2	-----..A-A.....C.....	256
Chum GH-2	TAAAAAAA.T-G...G.....C.....--	330
Pink GH-2	-----..-G...G.....A.....C.....--	268
Sockeye GH-2	-----..-G...G.....C.....--	267
Coho GH-2p	-----..-G.....C.....	268
Pink GH-2p	??	
Chinook GH-2p	-----..AG.....G.....C.....T.	269
Chum GH-2p	-----.T-G.....T.....-.....C.....T.	143
Atlantic GH-2	AAGGTAGGT-AAAACAACCACATATCACAGTCCTTGTAAGTAAAAC-C	314
Brown GH-2-.....-	315
Rainbow GH-2	.----.----.---.....T.....C.	301
Cutthroat GH-2	.----.----.---.....T.....C.	304
Chinook GH-2A.....C.	305
Coho GH-2-..G.....T.....C.	303
Chum GH-2	.----.----.---.....T..T.....C.	367
Pink GH-2	.----.----.---.....T..T...A.....C.	305
Sockeye GH-2	.----.----.---.....G..T.....C.	304
Coho GH-2pT.-.....C.C..T..G.....G.....T.	315

**The Search for a Y-Linked Marker in Pink Salmon
(*Oncorhynchus gorbuscha*) using Bulked Segregant
Analysis of AFLP and PINE Markers**

Kristine L. Pilgrim

*Division of Biological Sciences, University of Montana,
Missoula, MT 59812.*

ABSTRACT

The objective of this study was to identify molecular markers linked to sex determination in pink salmon (*Oncorhynchus gorbuscha*). Pools of DNA from males, and pools of DNA from females from the same family were created using the method of bulked segregant analysis. Seventy-four amplified fragment length polymorphism (AFLP) primers, and 18 paired interspersed nuclear elements (PINE) primers were used to amplify random, multilocus DNA fragments. A total of ~2400 bands (loci) were examined without detecting a sex-specific marker. The failure to detect a sex-linked marker in this study may be due to the sex-determining region (*SEX*) being quite small on the Y-chromosome, such that detecting a marker in this region is difficult.

INTRODUCTION

Finding a marker specific to the Y-chromosome (linked to the sex determining region) in pink salmon (*Oncorhynchus gorbuscha*) requires a technique that will be able to effectively target that particular region of the genome. Classically, to detect linkage (i.e. between a genomic region of interest and other markers) requires a large number of individuals and is often a labor intensive process (Wang and Paterson 1994). In contrast, a method called "bulked segregant analysis" has been developed to detect genetic markers linked to a target gene or region of the genome that is much less labor intensive than classic linkage studies (Michelmore et al. 1991, Ronin et al. 1996).

In bulked segregant analysis, DNA from individuals that segregate for a trait of interest are pooled together (or bulked), such that the target region is homogeneous within pools, but differs between the pools. These pools of DNA can then be screened with random genetic markers to detect a marker that differs between the two pools and is thus linked to the genomic region of interest.

Pooling DNA has been shown to be an extremely cost effective and time-saving technique to detect markers linked to a particular genomic region (Giovannoni et al. 1991,

Michelmore et al. 1991). Bulked segregant analysis may cut down the number of PCR runs needed to detect a marker linked to a target region by up to 90-97% (Ronin et al. 1996). Bulked segregant analysis can pool DNA from any individuals in a segregating population (e.g. full-sibs, back crossed individuals), and can be used to target specific genes, or chromosomal regions of interest (Giovanonni et al. 1991).

Theoretically, bulked segregant analysis can be used to detect a Y-linked marker in pink salmon where the sex determining region (*SEX*) is the genomic region of interest. Pooling DNA from males and females generated from the same, single pair mating will create two pools of DNA that differ for *SEX* and the surrounding, non-recombining region of the Y-chromosome, yet are genetically similar at all loci unlinked to the target region. However, in order for bulked segregant analysis to be an effective method, it is necessary that a large number of polymorphic loci are screened on the two pools of DNA.

Since the introduction of this method by Michelmore et al. 1991 and simultaneously by Giovannoni et al. 1991, randomly amplified polymorphic DNAs (RAPDs) and restriction fragment length polymorphisms (RFLPs) have most commonly been used with bulked segregant analysis. Although RFLPs and RAPDs can be useful as molecular markers and in mapping

studies, they also have associated problems. RFLPs are often time consuming and laborious, and is the reason RAPDs have often been used solely in bulked segregant analysis (Chalmers et al. 1993, Paran and Michelmore 1993, Chagué et al. 1996). RAPDs generate several DNA fragments, potentially amplify throughout the entire genome, and have been used widely to assess genetic variation. However, in an evaluation of RAPDs in red deer, wild boar, and fruit fly, the reproducibility of banding patterns in the same individual was poor, and the percentage of repeatable RAPD bands ranged from 23-36% (Pérez et al. 1998).

In recent years, new types of molecular marker that amplify multiple loci simultaneously have been developed. Amplified fragment length polymorphisms (AFLPs), take advantage of random priming throughout the genome, but with a degree of specificity. The AFLP method uses the polymerase chain reaction (PCR) to amplify restriction fragments generated by digesting the genome using two restriction enzymes. Special adapter sequences are then ligated to the sticky ends created by the restriction enzymes and PCR is performed using primers specific to the adapter sequence, the restriction site, and an additional base in order to reduce the number of genomic DNA fragments

originally generated by the restriction enzymes. A second round of PCR is performed using primers specific to the adapters, restriction site, the base used during the first round of PCR, and finally, two additional bases to further reduce the number of fragments so that they can be visualized by gel electrophoresis. The AFLP procedure is able to use multiple primer combinations during the second round of PCR; and an entirely new array of primer combinations is possible by choosing a different original selective base during the first round of PCR.

AFLP primers have proven extremely useful in plant mapping studies (Thomas et al. 1995, Vos et al. 1995) and are highly polymorphic and reliably reproducible in comparison to RAPDs. Reproducibility is high because the annealing temperatures used during PCR are much higher (~60°C) compared with temperatures used for RAPD primers. In a preliminary survey of RAPD and AFLP markers in pink salmon (*Oncorhynchus gorbuscha*), RAPDs produced an average of 2-8 bands per primer, while each AFLP primer pair amplified at least 30 bands, with an average of 7.5 being polymorphic (Allendorf et al. 1997c).

Another class of variable molecular markers takes advantage of short interspersed nuclear elements (SINEs) found throughout the genome. SINEs are characterized by an

internal POL III promoter, an A-rich 3' end, flanking direct repeats, and may be related to 7SL RNA and tRNA (Okada 1991, Li 1997). SINEs have been found in several mammal species, and various repetitive DNA elements have been discovered throughout the genome of salmonid fishes. Kido et al. (1991) documented the presence of two such elements, *HpaI* and *SmaI* in pink salmon. Spruell and Thorgaard (1996) subsequently reported the presence of the 5' end of the third element, *FokI*, in pink salmon. Goodier and Davidson (1994) confirmed that salmonids also contain the transposon *Tc1*, a member of another class of repetitive elements. Finally, Bois et al. 1998 documented the presence of a family of VNTR repeats (named 33.6+2, in the family of Jeffrey's repeats) in salmonids.

DNA sequences homologous to salmonid-specific SINEs and the transposon *Tc1*, have been used as primers to generate multiple DNA fragments from a single PCR reaction (Spruell et al. 1999). This technique has been called PINES (paired interspersed nuclear elements). Primers homologous to one end of the element are oriented such that they initiate DNA synthesis from the end of the element, progressing into the surrounding genomic DNA. A single primer or combinations of primers may be used to generate multilocus patterns.

Pooling DNA from males and females (respectively) from a single-pair mating, creates two pools of DNA that differ only in their sex chromosome composition (Fig. 1). A polymorphism detected in the male pool that is absent in the female pool could indicate a marker that is on the Y-chromosome, and presumably linked to the sex determining region. It may also be possible to detect a marker specific to the X-chromosome. If a marker is present in the female pool and absent in the male pool, this could indicate a marker specific to the X-chromosome from the male parent (chromosome X3, Fig. 1). Using AFLP and PINE markers will greatly increase the chances of finding a sex-linked marker due to the greater number of polymorphisms produced per unit effort. I report here on the use of AFLP and PINE markers to find a sex-linked marker in pink salmon using bulked segregant analysis.

MATERIALS and METHODS

Pink Salmon Families and DNA Extraction

Two, full-sibling families were used in this study. Pink salmon families from a controlled, single-pair mating were generated at the Armin F. Koering Hatchery in Prince William Sound and at the Genetics Lab facilities of the Alaska Department of Fish and Game in Alaska. The first

family, known as "family A14" was generated from 2, even year sexually mature fish collected in 1996, and represents an even year run sample. The second family, designated "family 7B" was generated from 2, odd-year sexually mature fish, and represent an odd-year run sample.

Embryos were incubated until just prior to hatching when they were collected and preserved in ethanol. Ninety-six embryos from family A14, and 58 embryos from family 7B were collected. DNA was isolated from the embryos after separation from the yolk sac using the Purgene^(TM) DNA isolation kit (Gentra Systems Inc.) The concentration of DNA was determined using a scanning spectrofluorometer.

Confirmation of Families

It was important to test that the individuals chosen for analysis belonged to the correct family in case embryos from another family mistakenly ended up in the wrong rearing trays during hatchery incubations. Confirmation of family identity of 58 embryos from family 7B was performed by DNA fingerprinting analysis using primers matching the *Sma*I SINE and *Tc*1 transposon (Greene and Seeb 1997).

Family identity of embryos from family A14 was tested using two microsatellite loci and one DNA fingerprint using primers matching the *Fok*1 SINE and the *Tc*1 transposon.

Microsatellite analysis was performed in 10 μ l PCR mixtures that contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 4.0 mM MgCl₂, all four dNTPs (each at 0.2 mM), each fluorescently labeled primer at 0.45 μ M, 0.5 U *Taq* DNA polymerase (Perkin-Elmer/Cetus), and 20 ng of purified genomic DNA. Primers and annealing temperatures are as follows: μ Sat60, 55°C (Estoup et al. 1993); and *Ots1*, 55°C (Hedgecock D, personal communication). PCR products were electrophoresed on a 7% denaturing polyacrylamide gel, and visualized with a Hitachi FMBIO-100 fluorescent imager. The individuals tested from each family showed Mendelian inheritance and allelic patterns consistent with the parents used to create the family at both microsatellite loci used for screening (Table 1).

PCR amplification of DNA fragments flanked by the *FokI* SINE and the *Tc1* transposon were conducted in a total reaction volume of 10 μ l. PCR mixtures contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 5.0 mM MgCl₂, all four dNTPs (each at 0.2 mM), each fluorescently labeled primer at 0.4 μ M, 0.1 unit of *Taq* DNA polymerase (Stoffel fragment; Perkin-Elmer/Cetus), and 20 ng of DNA. PCR products were electrophoresed on a 4% denaturing polyacrylamide gel, and visualized with a Hitachi FMBIO-100 fluorescent imager.

Ninety-three individuals from family A14 demonstrated banding patterns consistent with the parents used to create the family.

Sexing Embryos

DNA from all embryos from family A14, and family 7B was amplified using primers designed from chinook salmon that target the growth hormone pseudogene (*GH-2p*; R. H. Devlin personal communication). This Y-linked pseudogene allows the sex of unknown individuals to be determined. PCR was performed in mixtures that contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 5.8 mM MgCl₂, all four dNTPs (each at 0.2 mM), each primer at 0.3 μ M, and 0.08 unit of *Taq* DNA polymerase (Perkin-Elmer/Cetus), and 40 ng of template DNA in a total volume of 15 μ l. The PCR profile was 30 cycles of 92°C for 1 min, 51°C for 1 min, and 72°C for 2 min. PCR products were electrophoresed on 2.0% agarose gels in TAE buffer (Ausubel et al. 1989), containing ethidium bromide at 0.5 μ g/ml. Males were identified by the presence of a 166 bp fragment that is absent in females (previously discussed, chapter 2). Twenty individuals identified as males, and 20 individuals identified as females from each family were chosen for subsequent DNA analysis (Table 2).

DNA Pooling

The number of individuals to include in a pool is a balance between a pool that is large enough to avoid detecting false positives (unlinked loci that appears polymorphic between the pools) but small enough to maximize the length of chromosome that remains polymorphic between the pools (Giovannoni et al. 1991, Churchill et al. 1993, Ronin et al. 1996). It has been recommended that more than 5 individuals per pool be used (Michelmore et al. 1991) and that 10 individuals is sufficient to avoid most false positives (Wang and Paterson 1994). The size of the non-recombining, sex determining region on the Y-chromosomes of salmonids is unknown. Therefore, using smaller pools of individuals will maximize the chance of detecting markers within the sex-determining region of the Y-chromosome.

Individuals to be included in the pools of DNA were chosen based on their sex and the quality of DNA (as measure by quantity with the spectrofluorometer). Two male pools of DNA (labeled "pool 1" and "pool 3"), and two female pools of DNA (labeled "pool 2" and "pool 4") were created from family A14 using different individuals in each. One male pool ("pool 5") and one female pool ("pool 6") were created from family 7B. Each pool of DNA was created by combining .5 μ g of DNA from each of 10 individuals (Fig. 2).

AFLP and PINE Analysis of DNA Pools

Ten ng/ μ l of the pooled DNA was used in the AFLP restriction/ligation and pre-selective amplification steps according to the Perkin-Elmer/Applied Biosystems AFLP plant mapping protocol with the modifications outlined below. Ten μ l selective amplification reactions contained 1.5 μ l pre-amplification products as DNA template, 0.5 μ l EcoRI selective primer, 0.5 μ l MseI selective primer, 2.0 mM MgCl₂, all four dNTPs (each at 0.1 mM), 2X Amplitaq PCR buffer, and 0.5 U of *Taq* DNA polymerase (Perkin-Elmer/Cetus). The PCR profile for selective amplification was: initial denaturation at 96°C for 2 minutes, followed by a series of 7 cycles with denaturation at 96°C for 1 second, annealing at 65°C for 30 seconds, extension at 72°C for 2 minutes. The annealing temperature was decreased by 1 degree per cycle for 6 cycles; and an additional 30 cycles with an annealing temperature of 59°C for 30 seconds followed. Products were electrophoresed on a 7% denaturing polyacrylamide gel and visualized using a Hitachi FMBIO-100 fluorescent imager.

PCR amplification of DNA fragments flanked by interspersed elements (PINEs) was performed using primers described in Table 3. PCR mixtures were performed in a

total reaction volume of 10 μ l and contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 5.0 mM MgCl₂, all four dNTPs (each at 0.2 mM), each fluorescently labeled primer at 0.4 μ M, 0.1 unit of *Taq* DNA polymerase (Perkin-Elmer/Cetus) and 20 ng of the pooled DNA. PCR products were electrophoresed and visualized in the same way as AFLPs.

Primers that generated putative polymorphisms seen between male and female pools using AFLP and PINE primers were subsequently confirmed by segregation analysis of individuals. These individuals included those used to create the pools of DNA.

Robustness of Pooling

Five AFLP primer pair combinations, and 3 PINE combinations were initially tested for priming consistency on both individuals and their respective DNA pools. Pool 1 and pool 2 from family A14 and the individuals comprising the pools were tested. Testing both AFLP primers, and PINE primers showed that all bands present in the individuals were also present in the pooled DNA samples. The test gave an estimate of the robustness of the pooling technique, showing that with both AFLPs and PINES, a band present in only one of the ten individuals comprising the pool was also present in the pooled DNA sample.

The sensitivity of detection estimate for AFLP and PINE markers is consistent with results observed using RAPD and RFLP markers. Using a single-copy DNA probe in Southern blot analysis, Churchill et al. 1993 was able to detect a rare allele in a mixture as low as 40:1 for RFLP markers. The limit of detection for RAPD markers was around a ratio of 10:1, although band intensity seemed to be band (locus) specific (Michelmore et al. 1991).

RESULTS

AFLP primer pairs produced scorable, multi-locus bands in the pooled DNA samples with sizes ranging from 75-400 bp. Seventy-four AFLP primer pair combinations were screened on the six pools of DNA. DNA fragments flanked by SINEs, the transposon Tc1, and a VNTR repeat 33.6+2 produced scorable, multilocus fragments in the pooled DNA samples. Using two different PINE primers (e.g. HpaI5'/SmaI5', HpaI3'/Tc15') produced around 20 bands, with products ranging in size from 75-450 bp. Using the same PINE primer (e.g. SmaI5') to generate multiple DNA fragments produced fewer bands with products ranging from around 200-500 bp. Fifteen PINEs using primers from different elements, and three PINEs using a single primer were screened on the pooled DNA samples.

A total of eight AFLP primer pairs produced potentially sex-specific products in the pooled DNA samples (Table 4). Three AFLP primer combinations produced bands present in the male pools of family A14 (pools 1 & 3) and absent in both the female pools from this family and the male and female pools from family 7B. Four AFLP primer combinations produced bands present in the male pool (pool 5) from family 7B that were absent from all other pools. One AFLP primer combination produced a product present in the female pools of family A14 that was absent from all other pools. No AFLP primer combination produced apparent sex-specific bands in males or females from both families.

A sex-specific AFLP fragment was not detected in this study. The eight AFLP primer pairs that produced sex-specific products based on the pooled DNA study were subsequently screened on the individuals from the family in which the polymorphism was observed (Table 5). In all cases, the potentially sex-specific band seen in the pooled DNA sample was not present in all individuals of the same sex from that family. Moreover, the band of interest was present in few individuals (usually 5 or 6) of both sexes.

One PINE primer combination (Fok1 5'/Tcl 5') produced a 260 bp fragment present in the female pools (pools 2 & 4) that was absent in all other pools (Table 4). These primer

combinations were then screened in the 40 individuals from family A14. The 260 bp fragment was present in only 12 females and 8 males and was not sex-linked.

Clearly, using AFLP markers, and DNA markers flanked by SINEs or the transposon Tc1 did not detect a sex-linked marker in this study. All potential polymorphisms seen between pools that differed by sex were false positives of markers occurring in low frequencies. Based on a sample of 25 AFLP selective primer pairs used in this study, an average of 33.6 bands are produced per primer pair. Therefore, on the order of 2000 bands (loci) were screened with the 74 AFLP selective primer pairs without detecting a sex-linked marker. An average of 21.8 bands per primer pair are amplified in pink salmon using the method of PINES (based on a sample of 10 PINE primer combinations used in the is study) suggesting that around 400 bands (loci) were examined with the 18 PINE primer combinations used.

Both AFLP and PINE primer combinations produced fragments that were specific to either the odd or even year family. However, family differences were not able to be distinguished from odd and even year-class differences in this study.

DISCUSSION

In previous studies, bulked segregant analysis has been a successful method for detecting markers linked to a particular region of the genome. Random markers such as RAPDs and RFLPs have been used successfully with bulked segregant analysis to find markers linked to disease resistance genes in several plants (Michelmore et al. 1991, Ballvora et al. 1995, Borovkova et al. 1995, Meksem et al. 1995). Additionally, bulked segregant analysis has been successful in detecting genetic markers linked to sex determination. One RAPD marker linked to the sex determining region in pistachio was detected after screening 700 RAPD primers (Hormaza et al. 1994). Similarly, one RAPD marker linked to sex determination in the basket willow after screening 380 RAPD primers (Aistrom-Rapaport et al. 1998). Bulked segregant analysis using AFLPs detected nine markers (out of 253 primer pairs screened) linked to the sex locus in *Asparagus officinalis* L. (Reamon-Büttner et al. 1998).

Detecting sex-linked markers in salmonids has been difficult. The sex chromosomes are not highly differentiated from one another cytogenetically, and genetically are thought to be homologous except for the sex determining region found on the Y-chromosome. Previously,

60 RAPD primers that were tested on individual (15 male and 30 female) rainbow trout failed to produce sex-specific markers (Moran et al. 1996). However, bulked segregant analysis used in conjunction with fluorescent *in situ* hybridization (FISH) was able to identify two Y-linked RAPD markers in rainbow trout (Iturra et al. 1998).

In this study, bulked segregant analysis was used to try and detect a sex-linked marker in pink salmon. In order to have a representative sample of pink salmon, families from both odd and even year classes were used because of the large genetic differences observed between odd and even year classes (Phillips and Kapuscinski 1988). Embryos from each family were sexed using the Y-linked, growth hormone pseudogene (Chapter 2). While sexually mature pink salmon show great morphological differentiation (U.S. Fish and Wildlife Service 1989), it is impossible to sex embryos and immature fish based on morphology. The description of a Y-linked growth hormone pseudogene (Chapter 2) now makes it possible to identify males and females genetically. Finally, to ensure that all embryos used were from the correct family, three microsatellite loci and one PINE were initially tested on the embryos and the parents used to create the families.

The methods of pooling DNA for bulked segregant analysis have been explored in other studies (Wang and Paterson 1994, Churchill et al. 1993). For effective pooling, it is important that DNA from all individuals is represented equally. This was ensured by determining the concentration of DNA using a scanning spectrofluorometer, and adding an equal amount of DNA from each individual to the pool.

A second concern of pooling DNA for use in bulked segregant analysis is the ability to detect a marker linked to the target region. Other studies use recombinant lines that differ for a trait or gene of interest, and are concerned with the pools containing enough recombinant chromosomes for informative mapping information. In the case here, the concern is less for recombinant chromosomes since the targeted region is the sex-determining region of the Y-chromosome where no recombination with the X-chromosome take place. In addition, when phenotypes are used to define the trait of interest (instead of using known mapped genomic markers) the size of the targeted region is unknown. Therefore, when the size of the region is unknown and when marker densities are low, smaller pools are desired (Churchill et al. 1993). Therefore, ten pink salmon were chosen for use in each pool because it has been determined

that ten individuals is the optimal pool size in order to detect closely linked markers (Michelmore et al. 1991, Wang and Paterson 1994)

A third concern is the ability to confirm that a marker detected in the pools of DNA is actually linked to the region of interest, and is not a false positive. False positives (i.e. a rare allele unlinked to the target region present in one of the pools), is more of a problem when pools of DNA are composed of fewer individuals. In this study, the ability to detect a linked marker was more of a concern than the risk of false positives since any marker that showed a polymorphism between the male and female pools was subsequently screened in individuals from the family.

For bulked segregant analysis to be an effective method of detecting markers linked to a region of interest, it is important to use genetic markers that are randomly distributed throughout the genome. Traditionally, RFLP and RAPD markers have been often used for these types of studies (Michelmore et al. 1991, Ballvora et al. 1995). However, RFLP analysis is often time consuming, and RAPD fragments may not be reliably reproducible (Pérez et al. 1998). In addition, limited polymorphisms were detected in pink salmon haploid families (Allendorf et al. 1997).

Recently, two new methods to amplify multilocus fragments from a single PCR have been developed. AFLP and PINE analysis are advantageous due to the large number of polymorphisms produced per primer pair and the high reproducibility of the fragments. AFLP and PINE markers have been shown to be fairly randomly distributed along the length of chromosomes in pink salmon (Allendorf et al. 1998), making these markers useful for targeting the sex-determining region.

Unfortunately, bulked segregant analysis of pink salmon males and females using AFLP and PINE techniques failed to detect a sex-linked marker in this study. It is possible that a greater number of primer combinations were needed to be screened on the pooled samples of male and female DNA from pink salmon to detect a Y-linked random marker in this species. Results from other studies demonstrate that bulked segregant analysis used with highly polymorphic markers such as AFLPs is an extremely effective technique for finding genomic markers linked to a genomic region of interest. Using RAPD primers, it was estimated that all segregating markers within 10% recombination of the target region should be detected, and markers within 30% recombination would often be detected; the limit of detection seemed to be approximately 25 cM (Michelmore et al. 1991).

The reason a sex-linked marker was not detected in this study could also be a factor of the targeted region (the sex determining region) itself. It is unknown how much of the Y-chromosome this region occupies in salmonids. Since it is difficult if not impossible to differentiate the sex chromosomes in most species of salmonids, it is likely that the sex-determining, non-recombining region comprises a small physical portion of the Y-chromosome, making detection of a linked marker, even when using bulked segregant analysis difficult. In humans, the sex determining region occupies about 300 kb on the short arm of the Y-chromosome (Sinclair 1994).

The sex determining region in salmon may also be fairly complex. Of the relatively few sex-linked loci known in salmonids, the locus *GH-2p* is Y-linked in chinook salmon, coho salmon, chum salmon, and pink salmon. However, cosmids developed from the surrounding DNA of *GH-2p* in chinook failed to detect Y-linked homologous DNA in the other three species (Devlin et al. in preparation). Unlike the *SRY* gene (sex-determining region Y) in mammals, no widely conserved sex-specific marker has been detected in non-mammals (Griffiths and Tiwari 1993). In salmonids, the lack of conserved, sex-linked markers between closely related species (see Chp. 1) indicates the evolution of sex

chromosomes has occurred in different lineages and at different times. Additionally, the few-sex linked loci known in salmonids, and the failure to detect a random sex-linked marker using bulked segregant analysis in pink salmon, suggests the sex determining region is extremely small, and the X and Y chromosomes are mostly homologous with one another.

Table 1. Inheritance of individuals from family A14 used in the study at two microsatellite loci. The parents genotypes are given under MP and FP (for "male parent" and "female parent").

LOCUS	MP	FP	Progeny Genotypes				X ²
			110/110	110/116			
uSAT60	110/110	110/116	49	43			0.39
			222/224	222/230	224/224	224/230	
OTS1	222/224	224/230	13	22	32	16	4.59

Table 2. Primer Sequences used for paired interspersed nuclear element (PINE) PCR and reference.

Primer Name	Sequence (5'-3')	Reference
HpaI 5'	AACCACTAGGCTACCCTGCC	Kido et al. 1991
HpaI 3'	ACAGGCAGTTAACCCACTGTTCC	Kido et al. 1991
FokI 5'	CTACCAACTGAGCCACACG	Kido et al. 1991
SmaI 5'	AACTGAGCTACAGAAGGACC	Kido et al. 1991
TcI 5'	GTATGTAAACTTCTGACCCACTGG	Greene and Seeb 1997
33.6+2	GGAGGAGGGCTGGAGGAGGGCGC	Bois et al. 1998

Table 3. Eight AFLP primer combinations, and one PINE that produced potentially sex-specific bands seen in analysis of pooled DNA samples. Pools 1 & 3 are male pools, and pools 2 & 4 are female pools from family A14. Pool 5 is comprised of males, and pool 6 is comprised of females from family 7B. An "X" indicates the presence of the particular fragment.

Primer Combination	Fragment Size (bp)	DNA Pools					
		M 1	F 2	M 3	F 4	M 5	F 6
ACG/CTA	420	X		X			
ACG/CTG	380	X		X			
AAC/CAG	180	X		X			
AGG/CTG	230		X		X		
ACG/CTT	280						X
AAG/CGA	140						X
AGG/CAT	120						X
AGG/CGA	260						X
Fok1 5'/Tc1 5'	250		X		X		

Table 4. Segregation analysis of 20 males and 20 females from each family tested using eight AFLP primer combinations, and one PINE that produced potentially sex-specific bands seen in analysis of pooled DNA samples.

Primer Combination	Fragment Size (bp)	Family			
		A14		7B	
		M	F	M	F
ACG/CTA	420	4	3		
ACG/CTG	380	6	4		
AAC/CAG	180	7	5		
AGG/CTG	230	3	5		
ACG/CTT	280			7	6
AAG/CGA	140			1	3
AGG/CAT	120			5	8
AGG/CGA	260			6	7
Fok1 5'/Tc1 5'	250	5	6		

Literature Cited

- Agellon L. B., S. L. Davies, T. T. Chen, and D. A. Powers. 1988. Structure of a fish (rainbow trout) growth hormone gene and its evolutionary implications. *Proceedings of the National Academy of Sciences of the United States of America* **85**:5136-5140.
- Aida, T. 1921. On the inheritance of color in a fresh-water fish, *Aplocheilus latipes*, with special reference to sex-linked inheritance. *Genetics* **6**:554-573.
- Allendorf, F. W. and G. H. Thorgaard. 1984. Tetraploidy and the evolution of salmonid fishes. Pages 1-53 in B. J. Turner ed. *Evolutionary genetics of fishes*. Plenum Press, New York.
- Allendorf, F. W., J. E. Seeb, K. L. Knudsen, G. H. Thorgaard, and R. F. Leary. 1986. Gene-centromere mapping of 25 loci in rainbow trout. *Journal of Heredity* **77**:307-312.
- Allendorf, F. W., W. A. Gellman, and G. H. Thorgaard. 1994. Sex-linkage of two enzyme loci in *Oncorhynchus mykiss* (rainbow trout). *Heredity* **72**:498-507.
- Allendorf, F. W. and R. S. Waples. 1996. Conservation and genetics of salmonid fishes. Pages 238-280 in J. C. Avise and J. L. Harrick eds. *Conservation genetics case histories from nature*. Chapman and Hall, New York.

- Allendorf, F. W., and R. G. Danzmann. 1997a. Secondary tetrasomic segregation of *MDH-B* and preferential pairing of homeologs in rainbow trout. *Genetics* **145**:1083-1092.
- Allendorf, F. W., D. Bayles, D. L. Bottom, K. P. Currens, C. A. Frissell, D. Hankin, J. A. Lichatowich, W. Hehlsen, P. C. Trotter, and T. H. Williams. 1997b. Prioritizing pacific salmon stocks for conservation. *Conservation Biology* **11**:140-152.
- Allendorf, F. W., P. Spruell, K. L. Knudsen, K. R. Lindner, and K. L. Pilgrim. 1997c. *Exxon Valdez* oil spill restoration project annual report.
- Allendorf, F. W., P. Spruell, K. L. Knudsen, K. R. Lindner, and K. L. Pilgrim. 1998. *Exxon Valdez* oil spill restoration project annual report.
- Alstrom-Rapaport, C., M. Lascoux, Y. C. Wang, G. Roberts, and G. A. Tuskan. 1998. Identification of a RAPD marker linked to sex determination in the basket willow (*Salix viminalis* L.). *Journal of Heredity* **89**:44-49.
- Ausbel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1989. *Current protocols in molecular biology*. John Wiley & Sons, New York.
- Bagley, M. J., and G. A. E. Gall. 1988. Mitochondrial and nuclear DNA sequence variability among populations of rainbow trout (*Oncorhynchus mykiss*). *Molecular Ecology* **7**:945-961.

- Ballvora, A., J. Hesselbach, J. Niewöhner, D. Leister, F. Salamini, and C. Gebhardt. 1995. Marker enrichment and high-resolution map of the segment of potato chromosome VII harboring the nematode resistance gene *Gro1*. *Molecular and General Genetics* **249**:82-90.
- Bjornsson, B.T. 1997. The biology of salmon growth hormone: From daylight to dominance. *Fish Physiology and Biochemistry* **17**:9-24.
- Borovkova, I. G., B. J. Steffenson, Y. Jin, J. B. Rasmussen, A. Kilian, A. Kleinhofs, B. G. Rossnagel, and K. N. Kao. 1995. Identification of molecular markers linked to the stem rust resistance gene *rpg4* in barley. *Phytopathology* **85**:181-185.
- Brooks, L. D., and W. Marks. 1986. The organization of genetic variation for recombination in *Drosophila melanogaster*. *Genetics* **114**:525-547.
- Bull, J. J. 1983. Evolution of sex determining mechanisms. Benjamin Cummings, Menlo Park, CA.
- Burgoyne, P. S. 1986. Mammalian X and Y crossover. *Nature* **319**:258-259.
- Chagué, V., J. C. Mercier, M. Guénard, A. de Courcel, and F. Vedel. 1996. Identification and mapping on chromosome 9 of RAPD markers linked to *Sw-5* in tomato by bulked segregant analysis. *Theoretical and Applied Genetics* **92**:1045-1051.
- Chalmers, K. J., U. M. Barua, C. A. Hackett, W. T. B. Thomas, R. Waugh, and W. Powell. 1993. Identification of RAPD markers linked to genetic factors controlling the milling energy requirement of barley. *Theoretical and Applied Genetics* **87**:314-320.

- Charlesworth, B. 1991. The evolution of sex chromosomes. *Science* **251**:1030-1033.
- Charlesworth, B., and D. Charlesworth. 1985. Genetic variation in recombination in *Drosophila melanogaster*. I. Response to selection and preliminary genetic analysis. *Heredity* **54**:71-83.
- Chevassus, B., A. Devaux, D. Chourrout, and B. Jalabert. 1988. Production of YY rainbow males by self-fertilization of induced hermaphrodites. *Journal of Heredity* **79**:89-92.
- Chinnici, J. P. 1971. Modification of recombination frequency in *Drosophila*. I. Selection for increased and decreased recombination. *Genetics* **69**:71-83.
- Churchill, G. A., J. J. Giovannoni, and S. D. Tanksley. 1993. Pooled-sampling makes high-resolution mapping practical with DNA markers. *Proceedings of the National Academy of Sciences of the United States of America* **90**:16-20.
- Devlin, R. H. 1991. Isolation of a Y-chromosomal DNA probe capable of determining genetic sex in chinook salmon (*Oncorhynchus tshawytscha*). *Canadian Journal of Fisheries and Aquatic Sciences* **48**:1606-1612.
- Devlin, R. H. 1993. Sequence of sockeye salmon type 1 and 2 growth hormone genes and the relationship of rainbow trout with Atlantic and Pacific salmon. *Canadian Journal of Fisheries and Aquatic Sciences* **50**:1738-1748.

- Devlin, R. H., G. W. Stone, and D. E. Smailus. 1998. Extensive direct-tandem organization of a long repeat DNA sequence on the Y chromosome of chinook salmon (*Oncorhynchus tshawytscha*). *Journal of Molecular Evolution* **46**:277-287.
- Domanico, M. J., and R. B. Phillips. 1995. Phylogenetic analysis of Pacific salmon (genus *Oncorhynchus*) bases on mitochondrial DNA sequence data. *Molecular Phylogenetics and Evolution* **4**:366-371.
- Domanico, M. J., R. B. Phillips, and T. H. Oakley. 1997. Phylogenetic analysis of Pacific salmon (genus *Oncorhynchus*) using nuclear and mitochondrial DNA sequences. *Canadian Journal of Fisheries and Aquatic Sciences* **54**:1865-1872.
- Du, S. J., R. H. Devlin and C. L. Hew. 1993. Genomic structure of growth hormone genes in chinook salmon (*Oncorhynchus tshawytscha*): Presence of two functional genes, GH-I and GH-II, and a male-specific pseudogene, GH-i. *DNA and Cell Biology* **12**:739-751.
- Ellis, N., A. Taylor, B. O. Bengtsson, J. Kidd, J. Rogers, and P. Goodfellow. 1990. Population structure of the human pseudoautosomal boundary. *Nature* **344**:663-665.
- Estoup, A., P. Presna, F. Krieg, D. Valman, 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.
- Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* **27**: 401-410.

- Fisher, R. A. 1931. The evolution of dominance. *Biological Review* **6**:446-455.
- Forbes, S. H., K. L. Knudsen, T. W. North and F. W. Allendorf. 1994. One of two growth hormone genes in coho salmon is sex-linked. *Proceedings of the National Academy of Sciences of the United States of America* **91**:1628-1631.
- Frolov, S. V. 1990. Differentiation of sex chromosomes in Salmonidae. Karyotype and sex chromosomes of sockeye salmon *Oncorhynchus nerka*. *Tsitologiya*. 32, No. 5, 509-514.
- Frolov, S. V. 1993. Mechanisms of sex chromosome differentiation in salmonidae. Translated from *Biologiya Morya*. No. 5 pp. 3-17.
- Giovannoni, J. J., R. A. Wing, M. W. Ganal, and S. D. Tanksley. 1991. Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Research* **19**:6553-6558.
- Greene, B. A. and J. E. Seeb. 1997. SINE and transposon sequences can be used to generate high-resolution DNA fingerprints, "SINE-prints", which exhibit faithful Mendelian inheritance in pink salmon (*Oncorhynchus gorbuscha*). *Molecular Marine Biology and Biotechnology* **6**(4):328-338.
- Griffiths, R., and B. Tiwari. 1993. The isolation of molecular genetic markers for the identification of sex. *Proceedings of the National Academy of Sciences of the United States of America* **90**:8324-8326.

- Hampikian, G. K., D. W. Cooper, and J. A. Marshall-Graves. 1994. Sex determination in marsupials and monotremes. Pages 143-170 in S. S. Wachtel ed. Molecular genetics of sex determination. Academic Press, San Diego, California USA.
- Heard, W. R. 1991. Life history of pink salmon (*Oncorhynchus gorbuscha*). Pages 119-230 in C. Groot and L. Margolis eds. Pacific salmon life histories. University of British Columbia Press, Vancouver.
- Hormaza, J. I., L. Dollo, and V. S. Polito. 1994. Identification of a RAPD marker linked to sex determination in *Pistacia vera* using bulked segregant analysis. Theoretical and Applied Genetics **89**:9-13.
- Hunter, G. A., E. M. Donaldson, F. W. Goets, and P. R. Edgell. 1982. Production of all female and sterile groups of coho salmon *Oncorhynchus kisutch* and experimental evidence for male heterogamety. Transactions of the American Fisheries Society **111**:367-372.
- Iturra, P., J. F. Medrano, M. Bagley, N. Lam, N. Vergara, and J. C. Marin. 1998. Identification of sex chromosome molecular markers using RAPDs and fluorescent in situ hybridization in rainbow trout. Genetica **101**:209-213.
- Johansen, B., O. C. Johnson, and S. Valla. 1989. The complete nucleotide sequence of the growth-hormone gene from Atlantic salmon (*Salmo salar*). Gene **77**:317-324.

- Johnstone, R., T. H. Simpson, A. F. Youngson, and C. Whitehead. 1979. Sex reversal in salmonid culture. Part III, The progeny of sex-reversed rainbow trout. *Aquaculture* **18**:13-19.
- Junakovic, N., A., Terrinoni, C., Di-Franco, C. Vieira, and C. Loevenbruck. 1998. Accumulation of transposable elements in the heterochromatin and on the Y chromosome of *Drosophila simulans* and *Drosophila melanogaster*. *Journal of Molecular Evolution* **46**:661-668.
- Kavsan, V. M., A. P. Koval, and A. Ju. Palamarchuk. 1994. A growth hormone pseudogene in the salmon genome. *Gene* **141**:301-302.
- Kawauchi, H., S. Moriyama, A. Yasuda, K. Yamaguchi, K. Shirahata, J. Kubota, and T. Hirano. 1986. Isolation and characterization of chum salmon growth hormone. *Archives of Biochemistry and Biophysics* **244**:542-552.
- Kido, Y., M Aono, Y. Toshifumi, K. Matsumoto, S. Murata, M. Saneyoshi, and N. Okada. 1991. Shaping and reshaping of salmonid genomes by amplification of tRNA-derived retroposons during evolution. *Proceedings of the National Academy of Sciences of the United States of America* **88**:2326-2330.
- Kirpichnikov VS. 1981. Genetic basis of fish selection. Springer-Verlag, New York.
- Kitano, T., N. Matsuoka, and N. Saitou. 1997. Phylogenetic relationship of the genus *Oncorhynchus* species inferred from nuclear and mitochondrial markers. *Genes and Genetic Systems* **72**:25-34.

- Li, W.-H. 1997. *Molecular Evolution*. Sinauer Associates, Inc., Publishers. Sunderland MA, USA.
- Li, W.-H., T. Gojobori, and M. Nei. 1981. Pseudogenes as a paradigm of neutral evolution. *Nature* **292**:237-239.
- Lloyd, M. A., and G. H. Thorgaard. 1988. Restriction endonuclease banding of rainbow trout chromosomes. *Chromosoma* **96**:171-177.
- Maddison W. P, and D. R. Maddison. 1992. *MacClade; analysis of phylogeny and character evolution*. Sinauer Associates Inc., Sunderland Massachusetts, USA.
- Male, R., J. A. H. Nerland, J. B. Lorens, W. Telle, I. Lossius, and G. K. Totland. 1992. The complete nucleotide sequence of the Atlantic salmon growth hormone I gene. *Biochemistry and Biophysics* **1130**:345-348.
- Marshall, A. R., K. L. Knudsen, and F. W. Allendorf. In preparation. Sex linkage, inheritance, and geographical distribution of variation at *PEPB-1* in chinook salmon (*Oncorhynchus tshawytscha*).
- May, B., K. R. Johnson, and J. E. Wright Jr. 1989. Sex linkage in salmonids: Evidence from a hybridized genome of brook trout and Arctic charr. *Biochemical Genetics* **27**:291-301.
- McKay, S. J., R. H. Devlin and M. J. Smith. 1996. Phylogeny of Pacific salmon and trout based on growth hormone type-2 and mitochondrial NADH dehydrogenase subunit 3 DNA sequences. *Canadian Journal of Fisheries and Aquatic Sciences* **53**:1165-1176.

- Meksem, K., D. Leister, J. Peleman, M. Zabeau, F. Salamini, and C. Gebhardt. 1995. A high-resolution map of the vicinity of the *R1* locus on chromosome V of potato based on RFLP and AFLP markers. *Molecular and General Genetics* **249**:74-81.
- Michelmore R. W., I. Paran, and R. V. Kesseli. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences of the United States of America* **88**:9828-9832.
- Moran, P., J. L. Martinez, E. Garcia-Vazquez, and A. M. Pendas. 1996. Sex chromosome linkage of 5s rDNA in rainbow trout (*Oncorhynchus mykiss*). *Cytogenetics and Cell Genetics* **75**:145-150.
- Moyle, P. B. and J. J. Cech, Jr. 1996. *Fishes: An introduction to ichthyology*. Simon and Schuster, Upper Saddle River, New Jersey, USA.
- Murata, S., N. Takasaki, M. Saitoh, H. Tachida, and N. Okada. 1996. Details of retropositional genome dynamics that provide a rationale for a generic division: The distinct branching of all the Pacific salmon and trout (*Oncorhynchus*) from the Atlantic salmon and trout (*Salmo*). *Genetics* **142**:915-926.
- Nanda, I., M. Schartl, W. Feichtinger, J. T. Epplen, and M. Schmid. 1992. Early stages of sex chromosome differentiation in fish as analyzed by simple repetitive DNA sequences. *Chromosoma* **101**:301-310.

- National Research Council, 1996. Upstream: Salmon and society in the Pacific Northwest. National Academy Press, Washington D.C.
- 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**:4-21.
- Nicoll, C. S., S. S. Steiny, D. S. King, R. S. Nishioka, G. L. Mayer, N. L. Eberhardt, J. D. Baxter, M. K. Yamanaka, J. A. Miller, J. J. Seilhamer, J. W. Schilling, and L. K. Johnson. 1987. The primary structure of coho salmon growth hormone and its cDNA. *General Comparative Endocrinology* **68**:387-399.
- Ohno, S. 1994. Conservation of the X-linkage group *in Toto* by all eutherian mammals. Pages 143-170 in S. S. Wachtel ed. *Molecular genetics of sex determination*. Academic Press, San Diego, California USA.
- Okada, N. 1991. SINEs: Short interspersed repeated elements of the eukaryotic genome. *Trends in Ecology and Evolution* **6**:358-361.
- Okada, H., H. Maramoto, and F. Yamazaki. 1979. Functional masculinization of genetic females in rainbow trout. *Bull. Jpn. Soc. Sci. Fish.* **45**:413-419.
- Pérez, T., J. Albornoz, and A. Domínguez. 1998. An evaluation of RAPD fragment reproducibility and nature. *Molecular Ecology* **7**:1347-1357.
- Phillips, R. B., and P. E. Ihssen. 1985. Identification of sex chromosomes in lake trout (*Salvelinus namaycush*). *Cytogenetics and Cell Genetics* **39**:14-18.

- Phillips, R. B., and A. R. Kapuscinski. 1988. High frequency of translocation heterozygotes in odd year populations of pink salmon (*Oncorhynchus gorbuscha*). *Cytogenetics and Cell Genetics* **48**:178-182.
- Price, D. J. 1984. Genetics of sex determination in fishes - a brief review. *In* Fish reproduction: strategies and tactics. Potts, G. G. and Wootton, R. J. (eds) pp. 77-89. Academic Press, London, New York.
- Prodohl, P. A., J. B. Taggart, and A. Ferguson. 1994. Single locus inheritance and joint segregation analysis of minisatellite (VNTR) DNA loci in brown trout (*Salmo trutta* L.). *Heredity* **73**:556-566.
- Reamon-Büttner, S. M., J. Schondelmaier, and C. Jung. 1998. AFLP markers tightly linked to the sex locus in *Asparagus officinalis* L. *Molecular Breeding* **4**:91-98.
- Reed, K. M., S. K. Bohlander, and R. B. Phillips. 1995. Microdissection of the Y chromosome and fluorescence in situ hybridization analysis of the sex chromosomes of lake trout, *Salvelinus namaycush*. *Chromosome Research* **3**:221-226.
- Rice, W. R. 1987. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution* **41**:911-914.
- Rice, W. R. 1996. Evolution of the y sex chromosome in animals. *BioScience* **46**(5):331-343.
- Rice, W. R. 1998. Male fitness increases when females are eliminated from gene pool: Implications for the Y chromosome. *Proceedings of the*

- National Academy of Sciences of the United States of America **95**:6217-6221.
- Ronin, Y. I., A. B. Korol, T. Fahima, V. M. Kirzhner, and E. Nevo. 1996. Sequential estimation of linkage between PCR-generated markers and a target gene employing stepwise bulked analysis. *Biometrics* **52**:1428-1439.
- Saxena, R., L. G. Brown, T. Hawkins, R. K. Alagappan, H. Skaletsky, M. P. Reeve, R. Reijo, S. Rozen, M. B. Dinulos, C. M. Disteche, and D. C. Page. 1996. The *DAZ* gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned. *Nature Genetics* **14**:292-299.
- Scribner K. T., J. R. Gust, and R. L. Fields. 1996. Isolation and characterization of novel salmon microsatellite loci; cross-species amplification and population genetic applications. *Canadian Journal of Fisheries and Aquatic Sciences* **53**:833-841.
- Sinclair, A. H. 1994. The cloning of *SRY*. Pages 23-41 in S. S. Wachtel ed. *Molecular genetics of sex determination*. Academic Press, San Diego, California USA.
- Slattery, J., and S. J. O'Brien. 1998. Patterns of Y and X chromosome DNA sequence divergence during the Felidae radiation. *Genetics* **148**:1245-1255.
- Spruell, P. and G. H. Thorgaard, 1996. SINE sequences detect DNA fingerprints in salmonid fishes. *Heredity* **76**:317-324.

- Spruell, P., K. L. Pilgrim, B. A. Greene, C. Habicht, K.L. Knudsen, K. R. Lindner, J.E. Seeb, and F. W. Allendorf. 1999. Inheritance of nuclear DNA markers in haploid pink salmon embryos. *Journal of Heredity* 90:00-00.
- Stearley, R. F., and G. R. Smith. 1993. Phylogeny of the Pacific trouts and salmon (*Oncorhynchus*) and genera of the family salmonidae. *Transactions of the American Fisheries Society* **122**(1):1-33.
- Swofford, D. L. 1999. PAUP*: phylogenetic analysis using parsimony and other methods. Version 4.0.0d64. Smithsonian Institution, Washington D.C.
- Takasaki, N., T. Yamaki, M. Hamada, L. Park, and N. Okada. 1997. The salmon *SmaI* family of short interspersed repetitive elements (SINEs): Interspecific and intraspecific variation of the insertion of SINEs in the genomes of chum and pink salmon. *Genetics* **146**:369-380.
- Thomas, C. M., P. Vos, M. Zabeau, D. A. Jones, K. A. Norcott, B. P. Chadwick, and D. G. Jones. 1995. Identification of amplified restriction fragment polymorphism (AFLP) markers tightly linked to the tomato *Cf-9* gene for resistance to *Cladosporium fulvum*. *Plant Journal* **8**:785-794.
- Thorgaard, G. H. 1977. Heteromorphic sex chromosomes in male rainbow trout. *Science* **196**:900-902.
- Thorgaard, G. H. 1978. Sex chromosomes in the sockeye salmon: A Y-autosome fusion. *Canadian Journal of Genetics and Cytology* **20**:349-354.

- Thorgaard, G. H. 1983. Chromosomal differences among rainbow trout populations. *Copeia* **3**:656-662.
- Thorgaard, G. H. and G. E. Gall. 1979. Adult triploids in a rainbow trout family. *Genetics* **93**:961-973.
- U.S. Fish and Wildlife Service. 1989. Species profiles: Life histories and environmental requirements of coastal fishes and invertebrates (Pacific Northwest); Pink salmon. Biological Report **82**(11.88). U.S. Fish and Wildlife Service, Washington D.C. USA.
- Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**:4407-4414.
- Wang, G-L., and A. H. Paterson. 1994. Assessment of DNA pooling strategies for mapping of QTLs. *Theoretical and Applied Genetics* **88**:355-361.
- Wienberg, J., R. Stanyon, W. G. Nash, P. C. M. O'Brien, F. Yang, S. J. O'Brien, and M. A. Ferguson-Smith. 1997. Conservation of human vs. feline genome organization revealed by reciprocal chromosome painting. *Cytogenetics and Cell Genetics* **77**:211-217.
- Winge, O. 1927. The location of eighteen genes in *Lebistes reticulata*. *Journal of Genetics* **18**:1-43.
- Yang, Bih-Ying, King-Ming, Chan, Chun-Mean Lin, and T. Chen. 1997. Characterization of rainbow trout (*Oncorhynchus mykiss*) growth hormone 1 gene and the promoter region of growth hormone 2 gene. *Archives of Biochemistry and Biophysics* **340**:359-368.

Young, W. P., P. A. Wheeler, V. H. Coryell, P. Keim, and G. H. Thorgaard.
1998. A detailed linkage map of rainbow trout produced using double
haploids. *Genetics* **148**:1-13.

D. Steinberg, E. K., K. R. Lindner, J. Gallea, A. Maxwell, J. Meng, and F. W. Allendorf.
2002. Rates and patterns of microsatellite mutations in pink salmon. *Molecular
Biology and Evolution* 19:1198-1202.

Rates and Patterns of Microsatellite Mutations in Pink Salmon

E. K. Steinberg, K. R. Lindner, J. Gallea, A. Maxwell, J. Meng, and F. W. Allendorf

Division of Biological Sciences, University of Montana, Missoula

The tendency of microsatellites to be highly polymorphic is a major factor responsible for their popularity as markers for ecological and evolutionary studies. Microsatellite polymorphism is generally attributed to slip-strand mispairing errors, causing the addition or deletion of repeat units during replication (Levinson and Gutman 1987). Valdes, Slatkin, and Freimer (1993) found that human microsatellite evolution appears to follow a stepwise mutation model (SMM) (Ohta and Kimura 1973). Following the SMM, single repeat units are added or deleted with equal and constant probability across all alleles. Several statistical methods to evaluate patterns of microsatellite variability and differentiation that assume variants of the SMM have since been developed (e.g., Goldstein et al. 1995a, 1995b; Slatkin 1995; Rousset 1996) and incorporated into widely used software programs such as GENEPOP (Raymond and Rousset 1995).

Although slippage during replication clearly plays a key role in the overall instability of the microsatellites (reviewed by Eisen 1999), mounting evidence indicates that microsatellite mutation dynamics are more complex than is reflected by the SMM. Numerous examples of mutations that do not constitute single repeat unit changes or those that reflect heterogeneity or bias in the mutational processes of particular loci or alleles have been documented (see reviews by Ellegren 2000a, 2000b and Schlotterer 2000). In particular, the length, type, and number of repeat units have been identified as important factors contributing to the complexity of microsatellite evolution.

Mutation events occurring early in gametogenesis can further complicate mutation dynamics. An underlying assumption of many population genetics models is that mutations occur and enter the gene pool independently. However, Woodruff and Thompson (1992) found that as many as 20% of new mutations detected in large-scale *Drosophila* screens did not occur as independent events but rather represented clusters of identical mutant alleles sharing a common premeiotic origin. Subsequently, Woodruff, Huai, and Thompson (1996) have shown that the occurrence of premeiotic cluster mutations can not only bias estimates of mutation rates but can also influence basic population genetic processes such as fixation probabilities. Cluster mutations have been documented at microsatellite loci in only two species: pipefishes (*Syngnathus typhle*, Jones et al. 1999) and green turtles (*Chelonia mydas*, FitzSimmons 1998). However, as pointed out by Ellegren (2000a, 2000b),

this may be because of the difficulty in detecting cluster mutations in organisms that produce small numbers of offspring per generation.

As stressed by Chambers and MacAvoy (2000), "a clear knowledge of the process of mutational change at microsatellite loci is imperative for the correct selection of theoretical models upon which statistical methods can be based" however, a knowledge of mutation dynamics requires more information than is typically available. In particular, a lack of inheritance data often precludes direct evaluation of the markers used in population studies. However, this is not the case for many salmonid fishes that are commonly raised in hatcheries and for which many microsatellite markers have been developed (e.g., Scribner, Gust, and Fields 1996; Olsen, Bentzen, and Seeb 1998; Banks et al. 1999). In this paper, we examine the transmission of nine microsatellite loci in 50 families to evaluate the dynamics of microsatellite mutations in pink salmon (*Oncorhynchus gorbuscha*).

We raised families of pink salmon by randomly pairing mature adults collected in Resurrection Bay, Alaska. We collected embryos from each of the families after eye pigment became apparent in the embryo. We extracted DNA from the embryos and fin clips taken from adults using the Puregene[®] DNA isolation kit (Gentra Systems Inc., Minneapolis, Minn.). We amplified microsatellites using primers for salmonid fishes developed in other laboratories: *OGO1c* and *OGO8* (*Ogo1c* and *Ogo8*, Olsen, Bentzen, and Seeb 1998); *OMY301* (*Omy301UoG*, R. Danzmann, personal communication); *OMYRGT6-1,2* (*OmyTRG6/1,iiTUF*, N. Okamoto, personal communication); *ONE μ 3* (*One μ 3*, Scribner, Gust, and Fields 1996); *OTS1* (*Ots1*, Banks et al. 1999); *SSA20.19-1,2* (*μ -20.19**, Sanchez et al. 1996); and *SSA408* (*Ssa408*, M. Cairney, personal communication). We used fluorescent primers and followed the PCR conditions recommended by the original authors. We visualized PCR products with a Hitachi FMBIO-100 or FMBIO II fluorescent imager after electrophoresis in 4.5% denaturing polyacrylamide gels. We scored alleles relative to commercial size standards (BioVentures, Inc.).

For the seven loci isolated from species other than pink salmon, we determined the repeat arrays in pink salmon by sequencing at least one allele. If multiple products were produced in the PCR reaction, we isolated bands in 3%–4% agarose gels. We then either purified and sequenced the bands or reamplified, purified, and sequenced them with both the forward and reverse primers. We purified PCR products with Qiaquick[®] columns (Qiagen) following the supplier's protocol. Direct sequencing of PCR products was performed by a commercial laboratory. We did not sequence alleles from the two loci that were developed from pink salmon (*OGO1c* and *OGO8*).

We initially genotyped parents and 10 progeny from each of the 50 families. Alleles present in progeny

Key words: microsatellite, mutation, salmon, *Oncorhynchus gorbuscha*.

Address for correspondence and reprints: Fred W. Allendorf, Division of Biological Sciences, University of Montana, Missoula, Montana 59812. E-mail: darwin@selway.umn.edu.

Mol. Biol. Evol. 19(7):1198–1202, 2002

© 2002 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

that were not present in either of that individual's parents were considered to be mutants after confirming that correct parentage could be assigned at the other loci. The progenitor of the mutant allele was assumed to be the parental allele that was closest in size to the mutant allele. Because we detected multiple mutations at *SSA408* in progeny from two of the families (98–23 and 98–26) in the initial analysis, we analyzed all the remaining embryos available from these families (38 and 40 embryos, respectively). To increase our sample size, we also genotyped 36 additional progeny from each of the five randomly selected families.

Our sequence analysis of alleles at six of the seven loci originally developed from other salmonids confirmed that these markers comprise similar microsatellites in pink salmon (table 1). The one exception (*OMYRGT6-1,2*) is a duplicated locus that was reported as a CA repeat in rainbow trout (*Oncorhynchus mykiss*, Sakamoto et al. 2000, table 1). In pink salmon, PCR amplifications using *OMYRGT6-1,2* primers consistently produced a smaller band that is present in all individuals as well as larger polymorphic bands that segregated after Mendelian inheritance. The polymorphic *OMYRGT6-1,2* locus in pink salmon contains a CACT repeat array, and the smaller monomorphic locus contains the short interrupted repeat (CA)₇-TA-(CA)₂ (table 1). The flanking sequences of the rainbow trout CA repeat and the pink salmon CA and CACT repeats can be readily aligned, suggesting that *OMYRGT6-1* and *OMYRGT6-2* have evolved two different microsatellite arrays. In the rainbow trout, the two copies of *OMYRGT6-1,2* map to different linkage groups (Sakamoto et al. 2000). However, whether these loci comprise different microsatellite arrays in the rainbow trout has not been determined (T. Sakamoto, personal communication).

We detected mutations only at the two loci that have the greatest number of alleles in the parental population (*OGO1c* and *SSA408*, table 1). Both these loci are composed of tetranucleotide repeats, whereas the other seven loci comprise a mixture of simple and interrupted dinucleotide repeats (table 1). These data are consistent with the findings of other studies, suggesting that tetranucleotide repeats have higher mutation rates than dinucleotide repeats (Weber and Wong 1993; Lee et al. 1999; Ellegren 2000a, 2000b). However, mutation rates can also depend on allele size (e.g., Schlotterer et al. 1998); the relative importance of repeat motif versus allele size on mutation rates at these nine loci is not known. In total, 24 of the 758 progeny analyzed had genotypes best explained as resulting from mutation events. Five of the mutations were at *OGO1c* and 19 were at *SSA408* (table 2). All these progeny had genotypes at all the other loci that were consistent with their parents. In addition, the second allele at the locus containing the putative mutation was shared with one parent (table 2).

At *SSA408* in the family 98-26, the 320-bp mutant allele was transmitted to 9 of the 50 embryos, and in the family 98-23, the 370-bp mutant allele was transmitted to 4 of the 48 embryos (table 2). These clusters of identical mutations suggest that many novel alleles

Table 1
Characterization of Microsatellite Loci in 50 Pink Salmon Families

Locus ^a	Reported Repeat Sequence ^a	Repeat Sequence in Pink Salmon	Size of Sequenced Allele ^b	Size Range of Alleles ^b	Number of Alleles	Estimated Mutation Rate ^c
<i>OGO1c</i>	(GTCT) ₂₆ -N ₇₀ -(AT) ₅ -(GT) ₁₁	(GTCT) ₂₆ -N ₇₀ -(AT) ₅ -(GT) ₁₁	Not sequenced	275–584	77	3.9 × 10 ⁻³
<i>OGO8</i>	(GA) ₂₂	(GA) ₂₂	Not sequenced	88–171	17	0
<i>OMY301</i>	(CA) ₂₀	(CA) ₃ -A-(CA) ₁₅	83	75–113	21	0
<i>OMYRGT6-1,2</i>	(CA) ₃₀	(CA) ₂ -(CACT) ₁₅ /(CA) ₇ -TA-(CA) ₂	188/170	172–225/170–170	18/1	0/0
<i>ONEμ3</i>	(GA) ₁₈	(GA) ₅ -N ₇ -(GA) ₄	167	160–167	3	0
<i>OTSI</i>	(TG) ₃ -N ₈ -(TG) ₃ -N ₄ -(TG) ₃ -N ₂ -(TG) ₇ -N ₁₆ -(TG) ₉	(TG) ₈ -N ₅ -(TG) ₁₆ -N ₉ -(TG) ₁₅	228	218–248	15	0
<i>SSA20.19-1,2</i>	(AC) ₁₃ -G-(CA) ₄	(AC) ₉ -(AC) ₄ -AG-(AC) ₅	79/74	77–79/62–74	2/3	0/0
<i>SSA408</i>	(GACA) ₂₇	(GACA) ₃₉	370	300–500	49	8.5 × 10 ⁻³

^a See text for references. *OMYRGT6-1,2* and *SSA20.19-1,2* are duplicated loci in pink salmon; data for the paralogs at these loci (when available) are separated by “/”.

^b Length of PCR product including both repeat and flanking sequence.

^c For confidence limits see text; zeros represent loci at which no mutations were detected in approximately 1,300 transmissions.

Table 2
Microsatellite Mutations Observed in Pink Salmon Families

LOCUS	FAMILY	PARENTAL GENOTYPES ^a		PROGENY GENOTYPES ^b						SIZE CHANGE OF MUTATION ^c		
		Dam <i>alb</i>	Sire <i>cl/d</i>	<i>alc</i>	<i>ald</i>	<i>b/c</i>	<i>b/d</i>	Genotypes	Not Matching Parents (No.)	No. Increase	No. Decrease	
<i>SSA408</i> ...	98-15	334/382	<u>322/370</u>	14	9	9	13	<u>334/374</u> (1)		1 (1)	—	
	98-19	<u>338/350</u>	378/404	15	12	8	10	342/404 (1)		1 (1)	—	
	98-22	334/404	350/366	4	1	1	2	<u>334/354</u> (1)		1 (1)	—	
	98-23	326/382	<u>366/386</u>	7	18	8	9	326/370 (1)	382/370 (3)	326/362 (2)	4 (1)	2 (1)
	98-26	316/404	312/450	9	12	12	8	320/450 (4)	320/312 (5)		9 (1)	—
	98-34	<u>354/386</u>	<u>370/440</u>	12	10	6	17	345/444 (1)			1 (1)	—
Total										17 (6)	2 (1)	
<i>OGO1c</i> ...	98-44	342/350	<u>408/474</u>	1	1	3	3	<u>342/478</u> (1)			1 (1)	—
	98-51	295/366	<u>303/362</u>	1	2	4	2	<u>295/366</u> (1)			1 (1)	—
	98-71	269/420	<u>346/450</u>	8	16	10	8	420/446 (2)			—	2 (1)
	98-64	<i>ala</i> ^c 348/348	<i>cl/d</i> <u>309/448</u>	5	4	0	0	348/444 (1)			—	1 (1)
Total										2 (2)	3 (2)	

^a The most likely progenitor of the mutant allele (selected on the basis of size similarity) is underlined.

^b The mutant allele is indicated by bold-face type, and the number of each mutant genotype observed is indicated (*n*).

^c The number in parenthesis is the number of mutation events inferred, assuming that multiple copies of the same mutant allele within a family resulted from a single mutation event.

that appear in a population may reflect mutations that occurred quite early in development. In sexually reproducing animals, gametes develop from primordial germ cells (PGCs) that differentiate from the somatic cells during the first several divisions in the developing zygote. These cells eventually migrate to the area of gonadal development where the germ cells are produced. The number of PGCs produced varies among organisms (see review by Matova and Cooley 2001).

In zebrafish (*Danio rerio*), researchers combining morphological and mRNA expression studies using germ line markers have recently determined that by the 5-somite (32-cell) stage and until about the 1,000-cell stage, there are four PGCs (Braat et al. 1999). During their migration toward the gonads, the four PGCs give rise to a total of 20–30 cells that populate the gonad and differentiate into germ cells (Braat et al. 1999). If a mutation occurs in one of the original four PGCs (and there is no attrition of cell lines), approximately one out of eight (12.5%) of the progeny should inherit the mutant allele. If gametogenesis is similar in pink salmon, our findings of 9 identical mutant alleles out of the 50 transmitted maternally (18%) in family 98-26 and 4 of the 46 identical mutant alleles (8.7%) transmitted paternally in family 98-23 suggest that each of these mutations likely occurred either in one of the four PGCs or in the subsequent one or two generations of cells that populated the gonad.

The occurrence of clustered mutations results in nonuniform distributions of novel alleles in a population which could influence interpretations of mutation rates and patterns as well as estimates of genetic population structure. For example, Woodruff, Huai, and Thompson (1996) have shown that mutant alleles that are a part of clusters are more likely to persist and be fixed in a population than mutant alleles entering the population independently. In the present study, 15 of the 24 mutant alleles detected at *SSA408* (54%) apparently resulted

from premeiotic mutations. Jones et al. (1999) similarly found that a high proportion (40%) of new mutants observed in pipefish occurred in clusters.

We estimated mutation rates by counting each mutant allele detected as one mutation, regardless of whether the allele appeared to be part of a mutational cluster. However, we only included randomly selected individuals in this analysis (i.e., we eliminated the 78 additional progeny from families 98-23 and 98-26 that we analyzed because we had detected multiple mutations in our initial analysis). The remaining 11 mutations in 1,300 transmissions at *SSA408* and 5 mutations in 1,278 transmissions at *OGO1c* yield mutation rate estimates of 0.0085 (0.0042–0.0151) and 0.0039 (0.0013–0.0091) mutations per gamete, respectively, with the numbers in parentheses being the 95% Poisson confidence intervals. We did not detect mutations at any of the other seven loci; the upper 95% confidence limit using a Poisson distribution for detecting zero mutations in 1,300 transmissions is 0.0028. The proportion of mutations observed varied significantly among the nine loci (contingency chi-square, $P < 0.001$). However, our mutation rate estimates for all the loci are within the range reported for other organisms (see Ellegren 2000b, table 1).

All the mutations detected were size changes of four bases (table 2) which is consistent with single-step addition or deletion mutations at both *SSA408* and *OGO1c*. Determining to what degree our data reflect a tendency for mutations to result in size increases or decreases depends on how mutations are counted. If all mutations are treated as single events, 17 mutations at *SSA408* reflected size increases and two reflected size decreases (table 2). Similarly, two mutations detected at *OGO1c* resulted in size increases and three resulted in size decreases (table 2). Alternatively, treating all within-family clusters of the same mutant allele as single mutations that were propagated during gametogenesis

reduces the number of size increases at *SSA408* to six and size decreases to one (table 2). Similarly, because two of the progeny in family 98-71 share the same mutation at *OGO1c*, the number of size decreases would be two rather than three (table 2).

Is the SMM appropriate for pink salmon microsatellites? All 11 of the unique mutations detected in this study were consistent with the addition or deletion of a single repeat unit, which is in accordance with the SMM. Furthermore, our data reflect a high incidence of homoplasy, as 7 of the 11 different mutant alleles detected were alleles already present in other families in this study (E. K. Steinberg et al., unpublished data). Because SMM-based estimators assume that alleles of similar sizes are related, these estimators are expected to be more accurate in the presence of size homoplasy than estimators that assume all mutations are independent and result in novel alleles (Estoup and Angers 1998). These findings suggest that genetic differentiation estimators based on the SMM would be appropriate. However, we also detected a tendency toward size increases in mutant alleles. Estoup and Angers (1998) recommend comparing results from estimators based on different underlying mutational models for concordance because it is not clear how different types of mutational biases affect SMM-based estimators. Given our findings, we agree with this recommendation.

The duplicated locus *OMYRGT6-1,2* apparently comprises two different microsatellite arrays in pink salmon. In rainbow trout, the two copies of this duplicated locus map to different linkage groups (Sakamoto et al. 2000). However, whether these loci comprise different microsatellite arrays in rainbow trout has not been analyzed (T. Sakamoto, personal communication). It would be informative to compare the sequences of *OMYRGT6-1* and *OMYRGT6-2* in rainbow trout, as well as other salmonid fishes, to study the evolution of these two divergent paralogous microsatellite loci. Salmonid fishes have undergone extensive gene duplication compared with other organisms, having diverged from a tetraploid ancestor approximately 25–50 MYA (Allendorf and Thorgaard 1984). Given the prevalence of duplicated loci in salmonid fishes, these organisms may provide an exceptional opportunity to use comparative approaches to study the molecular evolution of microsatellites.

In conclusion, we found evidence for heterogeneity in the rates and patterns of mutation among loci, suggesting that no single model will likely represent the complexity underlying the evolutionary dynamics of microsatellites. Our findings add to the evidence in support of the argument made by Woodruff, Huai, and Thompson (1996) that the occurrence of premeiotic cluster mutations may play an important role in the evolutionary dynamics of microsatellites. Finally, our identification of duplicated microsatellite loci comprising different repeat arrays indicates that the use of comparative analysis to study mutation dynamics (e.g., Amos 1999) could be misleading if the duplication is not detected.

Acknowledgments

We thank S. Forbes, K. Knudsen, M. Skinner, P. Spruell, and J. Wenburg for comments on the manu-

script, R. Leary, M. Steinberg, and A. Whiteley for helpful discussions, and G. Chambers for his suggestions and references on the process of mutations in microsatellites. E.K.S. was supported by a postdoctoral fellowship from the National Science Foundation. This research was funded by the Exxon Valdez Oil Spill Trustee Council (Project/190). However, the findings and the conclusions of the authors are their own and do not necessarily reflect the views or positions of the Trustee Council.

LITERATURE CITED

- ALLENDORF, F. W., and G. H. THORGAARD. 1984. Tetraploidy and the evolution of salmonid fishes. Pp. 1–53 in B. J. TURNER, ed. *Evolutionary genetics of fishes*. Plenum Publishing Corporation, New York.
- AMOS, W. 1999. A comparative approach to the study of microsatellite evolution. Pp. 66–79 in D. B. GOLDSTEIN and C. SCHLOTTERER, eds. *Microsatellites: evolution and applications*. Oxford University Press, Oxford.
- BANKS, M. A., M. S. BLOUIN, B. A. 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*). *J. Hered.* **90**:281–288.
- BRAAT, A. K., T. ZANDBERGEN, S. VAN DE WATER, H. J. T. H. GOOS, and D. ZIVKOVIC. 1999. Characterization of zebrafish primordial germ cells: morphology and early distribution of vasa RNA. *Dev. Dynam.* **216**:153–167.
- CHAMBERS, G. K., and E. S. MACAVOY. 2000. Microsatellites: consensus and controversy. *Comp. Biochem. Physiol.* **126**:455–476.
- EISEN, J. A. 1999. Mechanistic basis for microsatellite instability. Pp. 34–48 in D. B. GOLDSTEIN and C. SCHLOTTERER, eds. *Microsatellites: evolution and applications*. Oxford University Press, Oxford.
- ELLEGREN, H. 2000a. Heterogeneous mutation processes in human microsatellite DNA sequences. *Nat. Genet.* **24**:400–402.
- . 2000b. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends Genet.* **16**:551–558.
- ESTOUP, A., and B. ANGERS. 1998. Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. Pp. 55–86 in G. R. CARVALHO, ed. *Advances in molecular ecology*. Nato Sciences Series, IOS Press, Amsterdam.
- FITZSIMMONS, N. N. 1998. Single paternity of clutches and sperm storage in the promiscuous green turtle (*Chelonia mydas*). *Mol. Ecol.* **7**:575–584.
- GOLDSTEIN, D. B., A. RUIZ LINARES, L. L. CAVALLI-SFORZA, and M. W. FELDMAN. 1995a. Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl. Acad. Sci. USA* **92**:6723–6727.
- . 1995b. An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**:463–471.
- JONES, A. G., G. ROSENQVIST, A. BERGLUND, and J. C. AVISE. 1999. Clustered microsatellite mutations in the pipefish *Syngathus typhle*. *Genetics* **152**:1057–1063.
- LEE, J. S., M. G. HANFORD, J. L. GENOVA, and R. A. FARBER. 1999. Relative stabilities of dinucleotide and tetranucleotide repeats in cultured mammalian cells. *Hum. Mol. Genet.* **8**:2567–2572.

- LEVINSON, G., and G. A. GUTMAN. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4**:203–221.
- MATOVA, N., and L. COOLEY. 2001. Comparative aspects of animal oogenesis. *Dev. Biol.* **231**:291–320.
- OHTA, T., and M. KIMURA. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genet. Res.* **22**:201–204.
- OLSEN, J. B., P. BENTZEN, and J. E. SEEB. 1998. Characterization of seven microsatellite loci derived from pink salmon. *Mol. Ecol.* **7**:1087–1089.
- RAYMOND, M., and F. ROUSSET. 1995. GENEPOP: population genetics software for exact tests and ecumenicism. *J. Hered.* **86**:248–249.
- ROUSSET, F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics* **142**:1357–1362.
- SAKAMOTO, T., R. G. DANZMANN, K. GHARBI, et al. (12 co-authors). 2000. A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. *Genetics* **155**:1331–1345.
- SANCHEZ, J. A., C. CLABBY, D. RAMOS, G. BLANCO, F. FLAVIN, E. VASQUEZ, and R. POWELL. 1996. Protein and microsatellite single locus variability in *Salmo salar* L. (Atlantic salmon). *Heredity* **77**:423–432.
- SCHLOTTERER, C. 2000. Evolutionary dynamics of microsatellite DNA. *Chromosoma* **109**:365–371.
- SCHLOTTERER, C., R. RITTER, B. HARR, and G. BREM. 1998. High mutation rates of long microsatellite alleles in *Drosophila melanogaster* provide evidence for allele specific mutation rates. *Mol. Biol. Evol.* **15**:1269–1274.
- 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. *Can. J. Fish. Aquat. Sci.* **53**:685–693.
- SLATKIN, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**:457–462.
- VALDES, A. M., M. SLATKIN, and N. B. FREIMER. 1993. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics* **133**:737–749.
- WEBER, J. L., and C. WONG. 1993. Mutation of human short tandem repeats. *Hum. Mol. Genet.* **2**:1123–1128.
- WOODRUFF, R. C., H. HUAI, and J. N. THOMPSON JR. 1996. Clusters of identical new mutations in the evolutionary landscape. *Genetica* **98**:149–160.
- WOODRUFF, R. C., and J. N. THOMPSON JR. 1992. Have premeiotic clusters of mutation been overlooked in evolutionary theory? *J. Evol. Biol.* **5**:457–464.

AXEL MEYER, reviewing editor

Accepted March 12, 2002

E. Funk, W.C., Tyburczy, J.A., Knudsen, K.L., Lindner, K.R., and F. W. Allendorf.
2005. Genetic basis of variation in morphological and life history traits of pink
salmon (*Oncorhynchus gorbuscha*). J. Heredity 96:24-31.

Genetic Basis of Variation in Morphological and Life-History Traits of a Wild Population of Pink Salmon

W. C. FUNK, J. A. TYBURCZY, K. L. KNUDSEN, K. R. LINDNER, AND F. W. ALLENDORF

From the Division of Biological Sciences, University of Montana, Missoula, MT 59812-4824 (Funk, Knudsen, Lindner, and Allendorf); and Department of Zoology, Cordley Hall 3029, Oregon State University, Corvallis, OR 97331-2914 (Tyburczy). W. Chris Funk is currently at the Section of Integrative Biology, University of Texas, 1 University Station C0930, Austin, TX 78712.

Address correspondence to W. C. Funk at the address above, or e-mail: wcfunk@mail.utexas.edu.

Abstract

Understanding the genetic basis of phenotypic variation is essential for predicting the direction and rate of phenotypic evolution. We estimated heritabilities and genetic correlations of morphological (fork length, pectoral and pelvic fin ray counts, and gill arch raker counts) and life-history (egg number and individual egg weight) traits of pink salmon (*Oncorhynchus gorbuscha*) from Likes Creek, Alaska, in order to characterize the genetic basis of phenotypic variation in this species. Families were created from wild-caught adults, raised to the fry stage in the lab, released into the wild, and caught as returning adults and assigned to families using microsatellite loci and a growth hormone locus. Morphological traits were all moderately to highly heritable, but egg number and egg weight were not heritable, suggesting that past selection has eliminated additive genetic variation in egg number and egg weight or that there is high environmental variance in these traits. Genetic correlations were similar for nonadjacent morphological traits and adjacent traits. Genetic correlations predicted phenotypic correlations fairly accurately, but some pairs of traits with low genetic correlations had high phenotypic correlations, and vice versa, emphasizing the need to use caution when using phenotypic correlations as indices of genetic correlations. This is one of only a handful of studies to estimate heritabilities and genetic correlations for a wild population.

The genetic architecture of phenotypic variation can be described by two primary parameters: heritability and genetic correlations. Heritability in the narrow sense (b^2) is a dimensionless index of heritable variation, termed the additive genetic variance (V_A). Heritability is defined as the proportion of within-population phenotypic variance (V_P) that is additive genetic (i.e., $b^2 = V_A/V_P$; $0 \leq b^2 \leq 1$). The importance of b^2 for predicting a phenotypic response to selection can be seen with the breeder's equation $R = b^2S$, where S is the selection differential (the difference in the mean of a trait before and after selection within a generation) and R is the response to selection (the difference in the mean trait value between generations) (Roff 1997). Therefore the response to selection is a linear function of b^2 . Moreover, because selection reduces V_A , the magnitude of b^2 reflects the extent to which a trait has been under net directional or stabilizing selection in the past (Endler 2000).

Genetic correlations between pairs of traits also have important effects on the direction and rate of phenotypic evolution. The genetic correlation (ρ_A) between two traits

describes the proportion of the phenotypic correlation between the traits that is caused by genetic variation that affects both traits simultaneously. Genetic correlations between characters can arise by two mechanisms, pleiotropy or gametic phase disequilibrium (Lynch and Walsh 1998). Pleiotropy occurs when a single gene affects multiple traits due to complex biochemical and developmental pathways (Wright 1968). Gametic phase disequilibrium is the tendency of genes affecting different traits to be positively or negatively associated in the same individuals. Both pleiotropy and gametic phase disequilibrium can cause positive or negative genetic correlations. Genetic correlations can constrain or enhance phenotypic evolution, depending on whether the correlations are positive or negative and on whether selection acts in the same or opposite directions with respect to fitness on the traits in question. For example, a positive genetic correlation will speed the rate of evolution of two traits if selection acts on both traits in the same direction with respect to fitness, but will slow the rate of change if selection acts in opposite directions.

Despite the importance of understanding the genetic basis of phenotypic variation in evolutionary biology, many basic questions about heritabilities and genetic correlations remain due to a paucity of estimates of these parameters for natural populations (for examples of studies that have estimated these parameters for natural populations, see Boag and Grant 1978; Conner and Via 1993; Gustafsson 1986; Kruuk et al. 2000; Young et al. 1994). One important question is whether certain categories of traits, including morphological, life-history, behavioral, and physiological traits, have higher heritabilities than others. Generalizations about the heritabilities of different categories of traits will improve our understanding of the potential for evolution of different traits as well as provide insight into the relative intensity of past selection. Several studies have shown that heritability estimates tend to be highest for morphological traits, lowest for life-history traits, and intermediate for behavioral and physiological traits (Gustafsson 1986; Kruuk et al. 2000; Mousseau and Roff 1987). This result has been interpreted as support for Fisher's fundamental theorem of natural selection, which predicts that additive genetic variance for traits with strong effects on fitness, such as life-history traits, will approach zero at equilibrium (Fisher 1930).

A second important question is whether adjacent morphological traits have higher positive genetic correlations than nonadjacent traits. A positive relationship between the physical proximity of morphological traits and genetic correlations might be expected if adjacent traits are encoded or regulated by the same genes. Alternatively, more similar types of morphological traits that are nonadjacent, for example, structures such as limbs, may have higher genetic correlations than adjacent traits. Selection experiments on *Drosophila* wings demonstrate that adjacent morphological traits can evolve independently, suggesting that genetic correlations among adjacent structures are not always high enough to constrain independent evolution (Weber 1992).

A final question is whether genetic correlations accurately predict phenotypic correlations. In general, genetic correlations should predict phenotypic correlations fairly accurately since genetic correlations, along with environmental correlations, determine phenotypic correlations. However, the degree to which genetic correlations predict phenotypic correlations will depend on how often environmental correlations oppose genetic correlations. If, for example, environmental variation tends to cause negative correlations between pairs of traits that have positive genetic correlations, then genetic correlations will do a poor job of predicting phenotypic correlations. The relationship between genetic correlations and phenotypic correlations has important practical implications. Because genetic correlations are difficult to estimate, it would be convenient to use phenotypic correlations as indices of genetic correlations. If genetic correlations accurately predict phenotypic correlations, and vice versa, using phenotypic correlations as indices of genetic correlations would be justified. Reviews of the literature suggest that phenotypic and genetic correlations tend to have the same sign and magnitude (Cheverud 1988, 1995; Roff 1995, 1996).

We addressed the above three questions for a natural population of pink salmon (*Oncorhynchus gorbuscha*) from Likes Creek, Alaska, by estimating heritabilities and genetic correlations for morphological and life-history traits. Pink salmon have a strict, two-year life cycle in which reproductive adults return to their natal streams 2 years after they begin development. We created families from wild-caught adults, raised fish to the fry stage in the lab, released fry into the wild, and caught returning adults and assigned them to families using microsatellite loci and one growth hormone locus. Heritabilities and genetic correlations were then estimated using parent-offspring regressions and covariances, respectively. The morphological traits examined were fork length, the number of pectoral and pelvic fin rays, and the number of upper and lower gill arch rakers. The life-history traits were egg number and individual egg weight. Our specific objectives were to test (1) whether heritabilities were higher for morphological traits than life-history traits; (2) whether genetic correlations of morphological traits depend on the physical proximity of the traits; and (3) whether genetic correlations accurately predict phenotypic correlations.

Materials and Methods

Production of Families and Recapture of Returning Progeny

We collected gametes from 34 female and 34 male pink salmon from the mouth of Likes Creek, Resurrection Bay, Alaska, in August 1999. Sixty-eight full-sib families were created from the gametes and reared at the Alaska SeaLife Center (ASLC) in Seward, Alaska. Eggs collected from each female were divided into two equal groups and each group was fertilized with sperm from one male. Each family was placed in a separate tray of a Heath rack and incubated in fresh water at 4–5°C. Embryos were raised following the procedures of Lindner et al. (2000).

In February 2000, approximately 25,000 fry (young fish) from 67 full-sib families were pooled. In April 2000 we collected a sample of 500 fry, marked the remaining fry by clipping the adipose fins, and subsequently released 24,216 marked individuals into Resurrection Bay. In August and September 2001, we collected 260 marked adult progeny by snag-hooking near the freshwater outlet at the ASLC where the fry were released, seining the rivers in upper Resurrection Bay, and holding a lottery to encourage recreational fishermen to turn in marked fish. Progeny therefore spent approximately 8 of 24 months (one-third of their lives) in the lab and 16 of 24 months (two-thirds of their lives) in the ocean. Progeny were caught from 63 of the 67 released families. The mean number of progeny caught from each of these families was 4.1 (range 1–11).

Assignment of Parentage and Sex

DNA was extracted from fin clips or other collected tissues with the Puregene DNA isolation kit (Gentra Systems Inc., Minneapolis, MN). We analyzed each of the 68 parents

at nine microsatellite loci, including one duplicated locus (*SSA20.19-1,2*) and a growth hormone locus (*GH2*). Loci were amplified according to the original authors with minor modifications. Primers and annealing temperatures were as follows: *OmyRgt6*, 58–52°C (Sakamoto et al. 2000); *Ots1*, 56°C (Banks et al. 1999); *Ssa408*, 60°C (Cairney M, personal communication, 1997); *Ogo1c* and *Omy301*, 60°C and *Ogo8*, 58–52°C (Olsen et al. 1998); *Oneμ3*, 52°C (Scribner et al. 1996); *Ssa20.19-1,2*, 58–52°C (Sanchez et al. 1996); and *GH-2* intron C, 51°C (Spruell et al. 1999). Products from unlabeled primers were fluorescently tagged with TAMRA-labeled dUTP. Polymerase chain reaction (PCR) products were resolved via electrophoresis on denaturing 4.5% polyacrylamide gels and visualized and scored using a Hitachi FMBIO II fluorescent imager.

The returning adult progeny were analyzed at all 10 of these loci. Each fish was initially scored at three microsatellite loci (*OTS1*, *OMYRGT6*, and *SSA408*) and a list of possible families was made for each fish at each locus based on which alleles were carried by the parents of each family. Families not possible at all loci were eliminated, which reduced the number of possible families for most fish to between one and three. Fish were then run on gels next to possible parents for the remaining loci. After each locus was run, the alleles of each fish were compared with the alleles of potential parents, the list of possible families for each individual was revised, and the order of samples was changed so that fish were run beside parents and siblings. This allowed unambiguous family assignment of all fish and detection of any progeny with alleles whose lengths differed from those of their parents (mutants). Fish identified as mutants had allele combinations that placed them positively and uniquely in one family with the exception of a single allele that differed from the parental allele by one or two repeat units.

We also analyzed 240 of the 500 fry sampled prior to release at four loci (*OTS1*, *OMYRGT6*, and *SSA408*, and *OGO1C*), which allowed unambiguous assignment to families of all individuals as described above for adult progeny. Sex was assigned to fry using a Y chromosome-specific growth hormone pseudogene (Spruell et al. 1999). Presence of a 163 bp fragment indicated that a fish was male. However, eight sires lacked this diagnostic band, so the sex of the 49 fry in these families could not be determined.

Morphological and Life-History Data Collection

We measured the fork length of adult parents from 1999, adult offspring from 2001, and progeny released as fry in 2000. Fork length was measured as the distance from the middle of the eye to the fork of the caudal fin. Pectoral and pelvic fin rays and upper and lower gill arch rakers were counted for parents and adult offspring using a dissecting microscope. Average individual egg weight was estimated by dividing the weight of 100 eggs by 100. Egg number was estimated by dividing total egg weight by the weight of 100 eggs and multiplying by 100. Egg weight was estimated after water absorption in the 1999 mothers, but immediately after catching fish in the 2001 daughters. Egg weight is there-

fore not comparable between mothers and daughters, but this inconsistency does not affect egg number estimates.

Data Analysis

We compared the means of morphological and life-history traits between females and males and between parents and progeny using two-sample *t* tests. Variances were compared using *F* tests. Phenotypic correlations were calculated using Pearson correlations. All analyses were performed using MINITAB version 13.

We estimated heritabilities as the regression coefficients of regressions of family means against midparent values (Lynch and Walsh 1998). The significance of regression coefficients was assessed using *F* tests. Genetic correlations among traits were estimated from pairwise comparisons of parents and progeny (Lynch and Walsh 1998). Specifically, genetic correlations (ρ_A) were estimated from the equation

$$\rho_A \cong \frac{\sigma(\bar{z}_{1x}, \bar{z}_{2y}) + \sigma(\bar{z}_{2x}, \bar{z}_{1y})}{2\sqrt{\sigma(\bar{z}_{1x}, \bar{z}_{1y}) \cdot \sigma(\bar{z}_{2x}, \bar{z}_{2y})}}, \quad (1)$$

where $\sigma(\bar{z}_{1x}, \bar{z}_{2y})$ is the phenotypic covariance between midparents for trait 1 and offspring means for trait 2, $\sigma(\bar{z}_{2x}, \bar{z}_{1y})$ is the covariance between midparents for trait 2 and offspring means for trait 1, $\sigma(\bar{z}_{1x}, \bar{z}_{1y})$ is the covariance between midparents and offspring means for trait 1, and $\sigma(\bar{z}_{2x}, \bar{z}_{2y})$ is the covariance between midparents and offspring means for trait 2. The significance of genetic correlations was determined as recommended by Lynch and Walsh (1998) from the regression of offspring means of trait 1 against midparent values of trait 2 and from the regression of family means of trait 2 against midparent values of trait 1, giving two *P* values for each genetic correlation. Finally, regressions of parental phenotypic correlations on genetic correlations and progeny phenotypic correlations on genetic correlations were used to test whether genetic correlations accurately predict phenotypic correlations.

Results

Phenotypic Variation and Phenotypic Correlations

Morphological and life-history trait means and variances differed between females and males and between parents and progeny (Table 1). Mean length of fathers was greater than mothers ($t_{54} = 2.38$, $P = .02$). Variance in length was also greater in males than females in both parents ($n = 68$, $F = 0.36$, $P = .005$) and offspring ($n = 257$, $F = 0.42$, $P < .001$). Mean length ($t_{52} = 5.23$, $P < .001$) and egg number ($t_{57} = 11.29$, $P < .001$) of daughters was greater than mothers. Similarly, the mean number of pectoral fin rays ($t_{89} = 2.45$, $P = .02$) and lower gill arch rakers ($t_{98} = 5.81$, $P < .001$) was greater in offspring than in parents. The variance in pectoral fin ray counts ($n = 325$, $F = 1.63$, $P = .008$) and pelvic fin ray counts ($n = 326$, $F = 1.45$, $P = .05$) was greater in parents than progeny, but variance in counts of upper gill arch rakers was greater in progeny than parents ($n = 323$, $F = 0.66$, $P = .05$).

Table 1. Means and standard deviations for fork length, egg number, individual egg weight, and meristic traits of pink salmon parents and progeny from Likes Creek, Alaska

Trait	Parents		Progeny	
	Female	Male	Female	Male
Fry length	—	—	37.1 ± 3.6 (94)	36.2 ± 4.0 (93)
Length	456 ± 23 (34)	474 ± 39 (34)	480 ± 23 (122)	487 ± 36 (135)
Egg number	1191 ± 287 (34)	—	1836 ± 296 (104)	—
Egg weight	0.18 ± 0.01 (34)	—	0.15 ± 0.02 (110)	—
Pect	15.5 ± 0.6 (34)	15.8 ± 0.7 (34)	15.9 ± 0.5 (122)	15.8 ± 0.5 (135)
Pelv	10.8 ± 0.5 (34)	10.9 ± 0.5 (34)	10.8 ± 0.4 (122)	10.7 ± 0.4 (136)
UGA	13.1 ± 0.5 (34)	13.2 ± 0.5 (34)	13.1 ± 0.6 (121)	13.2 ± 0.7 (134)
LGA	17.4 ± 0.9 (34)	17.4 ± 0.7 (34)	18.0 ± 0.7 (121)	18.0 ± 0.7 (134)

Pect, pectoral fin rays; Pelv, pelvic fin rays; UGA, upper gill arch rakers; LGA, lower gill arch rakers.

All progeny values are for returning adults except for fry length, which was measured prior to release.

Sample sizes are shown in parentheses.

Length is in millimeters and egg weight is in grams; all other values are counts.

Eggs were weighed after water absorption for mothers and prior to water absorption for daughters, so egg weights are not comparable.

Phenotypic correlations are shown in Table 2. Many more pairs of traits had significant phenotypic correlations in progeny than in parents. In general, this is likely because of lower statistical power due to smaller sample sizes in parents. However, the phenotypic correlation between length and egg number was much lower in mothers (0.05) than in daughters (0.58) and likely reflects an actual difference between mothers and daughters.

Heritabilities and Genetic Correlations

Heritabilities of all six morphological traits were significantly different from zero and ranged from 0.33 to 0.63 (Figures 1 and 2). Heritabilities (b^2) were 0.34 for female length ($F_{1,48} = 10.26$, $P = .002$), 0.45 for male length ($F_{1,53} = 8.85$, $P = .004$), 0.55 for pectoral fin ray counts ($F_{1,61} = 47.48$, $P < .001$), 0.33 for pelvic fin ray counts ($F_{1,61} = 8.47$, $P = .005$), 0.63 for upper gill arch raker counts ($F_{1,61} = 17.67$, $P < .001$), and 0.45 for lower gill arch raker counts ($F_{1,61} = 17.92$, $P < .001$). In contrast, the two life-history traits examined—egg number ($b^2 = -0.08$, $F_{1,47} = 0.30$, $P = .59$) and egg weight ($b^2 = 0.22$, $F_{1,48} = 2.62$, $P = .11$)—were not

significantly heritable. Moreover, neither female length ($b^2 = 0.02$, $F_{1,36} = 1.14$, $P = .29$) nor male length ($b^2 = 0.02$, $F_{1,38} = 0.59$, $P = .45$) were heritable when estimated using fry rather than adult offspring.

Genetic correlations were significant for four pairs of traits: pectoral fin rays and pelvic fin rays; pectoral fin rays and lower gill arch rakers; pelvic fin rays and upper gill arch rakers; and upper and lower gill arch rakers (Table 3). For the first and last of the above pairs of traits, both possible regressions of progeny family means on midparent values were significant. For the other two pairs of traits, only one of the two regressions was significant. Genetic correlations could not be estimated for pairs of traits including egg number because the covariance of mother and daughter egg number was negative, resulting in the square root of a negative product in equation 1. Similarly, genetic correlations could not be estimated for pairs of traits including egg weight because the covariance of mother and daughter egg weight was zero, resulting in a product of zero in the denominator of equation 1. However, regression analysis could still be used to test whether genetic correlations between egg number and other traits and between egg weight and other traits were

Table 2. Phenotypic correlations among traits of pink salmon from Likes Creek, Alaska

	Length	Egg no	Egg wt	Pect	Pelv	UGA	LGA
Length	—	0.05 (ns)	0.37 (0.03)	0.01 (ns)	0.010 (ns)	-0.02 (ns)	0.08 (ns)
Egg number	0.58 (<0.001)	—	-0.22 (ns)	-0.01 (ns)	0.23 (ns)	0.18 (ns)	0.18 (ns)
Egg weight	0.33 (<0.001)	0.07 (ns)	—	0.12 (ns)	0.10 (ns)	0.19 (ns)	0.16 (ns)
Pect	-0.15 (0.02)	0.01 (ns)	-0.23 (0.01)	—	0.02 (ns)	0.16 (ns)	0.12 (ns)
Pelv	-0.01 (ns)	0.03 (ns)	-0.06 (ns)	0.30 (<0.001)	—	0.16 (ns)	0.02 (ns)
UGA	0.05 (ns)	0.18 (ns)	0.26 (0.01)	0.04 (ns)	0.08 (ns)	—	0.35 (0.003)
LGA	0.14 (0.02)	0.10 (ns)	0.18 (ns)	0.02 (ns)	0.15 (0.02)	0.50 (<0.001)	—

Correlations above the diagonal are for parents and those below the diagonal are for progeny.

P values are shown in parentheses for correlations significant at the $\alpha = 0.05$ level.

“ns” indicates that a correlation was not significant.

Length, fork length; Pect, pectoral fin rays; Pelv, pelvic fin rays; UGA, upper gill arch rakers; LGA, lower gill arch rakers.

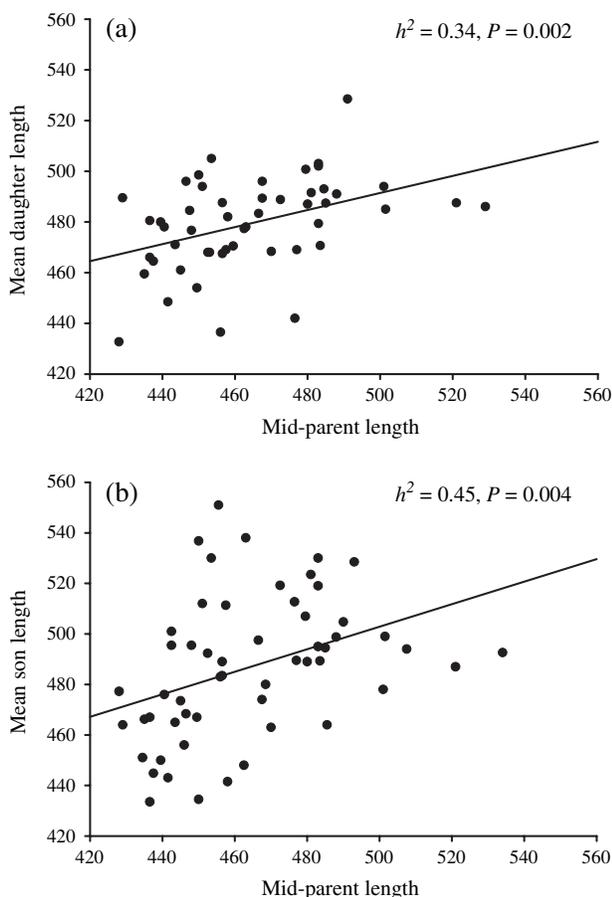


Figure 1. Regressions of (a) mean daughter fork length and (b) mean son fork length on midparent fork length. The heritability of length is the slope of the regression line.

significant. No genetic correlations were significant between egg number and any other trait nor between egg weight and any other trait.

Regression of Phenotypic Correlations on Genetic Correlations

The regression of phenotypic correlations on genetic correlations is positive using parent phenotypic correlations and progeny phenotypic correlations, but only the regression using progeny phenotypic correlations is significant (Figure 3). A few noteworthy outlier data points were observed. Some pairs of traits had high phenotypic correlations, but low genetic correlations. The genetic correlation between length and egg weight was indistinguishable from zero, but the phenotypic correlation between these traits was 0.37 in parents and 0.33 in progeny (Table 2). Similarly the genetic correlation between length and egg number was not significantly different from zero, but the phenotypic correlation between these traits was 0.58 in progeny (Table 2). Conversely, other pairs of traits had low phenotypic correlations, but high genetic correlations. Pelvic and pectoral fin ray

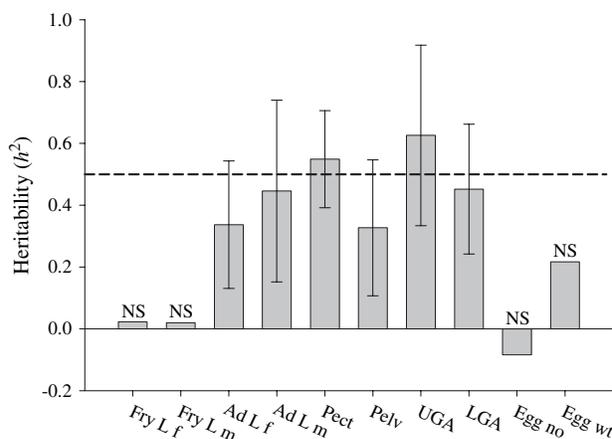


Figure 2. Heritability estimates and 95% confidence intervals for female and male fry fork length (Fry L f and Fry L m, respectively), female and male adult fork length (Ad L f and Ad L m, respectively), pectoral (Pect) and pelvic (Pelv) fin ray counts, upper (UGA) and lower (LGA) gill raker counts, egg number (Egg no), and egg weight (Egg wt) in pink salmon from Likes Creek, Alaska. NS indicates heritabilities that were not significantly different from zero.

counts had a high genetic correlation of 0.64 (Table 3), but a phenotypic correlation of only 0.02 in parents (Table 2). Likewise, pelvic fin rays and upper gill arch rakers had a genetic correlation of 0.76 (Table 3), but a phenotypic correlation of only 0.08 in progeny (Table 2).

Discussion

Heritabilities of Morphological and Life-History Traits

Our heritability estimates for pink salmon corroborate previous studies demonstrating higher heritabilities for morphological traits than life-history traits (Gustafsson 1986; Kruuk et al. 2000; Mousseau and Roff 1987). Heritabilities of the six morphological traits we examined ranged from 0.33 to 0.63, whereas the heritabilities of egg number and egg

Table 3. Genetic correlations among traits of pink salmon from Likes Creek, Alaska

	Length	Pect	Pelv	UGA
Pect	-0.20 (ns)	—	—	—
Pelv	0.08 (ns)	0.64 (0.02, 0.01)	—	—
UGA	-0.30 (ns)	0.04 (ns)	0.76 (0.57, <0.001)	—
LGA	0.19 (ns)	0.30 (0.91, 0.02)	0.34 (ns)	0.71 (0.04, 0.001)

P values are shown in parentheses for both possible regressions of progeny family means against midparent values for each pair of traits for cases in which at least one regression was significant at the $\alpha = 0.05$ level.

“ns” indicates that neither regression was significant.

Length, fork length; Pect, pectoral fin rays; Pelv, pelvic fin rays; UGA, upper gill arch rakers; LGA, lower gill arch rakers.

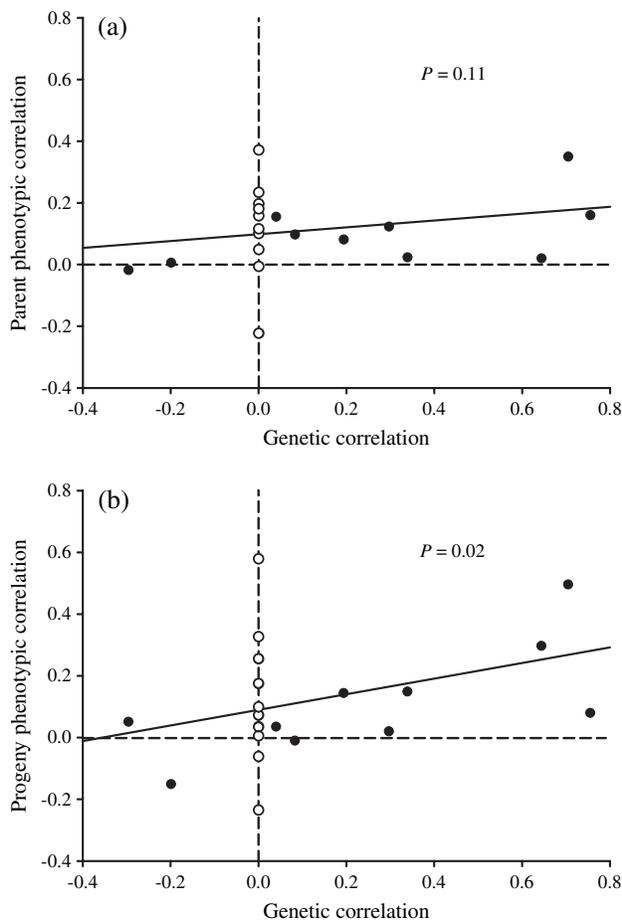


Figure 3. Regressions of (a) parent phenotypic correlations and (b) progeny phenotypic correlations on genetic correlations among phenotypic traits in pink salmon from Likes Creek, Alaska. Genetic correlations that could not be estimated because of negative or zero covariances between parental and offspring traits are represented by open circles (see text for details).

weight were indistinguishable from zero (Figure 2). Our heritability estimates for the four meristic traits examined (pectoral fin rays, 0.55; pelvic fin rays, 0.33; upper gill arch rakers, 0.63; and lower gill arch rakers, 0.45) were similar to heritability estimates for these same traits in rainbow trout (0.52, 0.84, 0.67, and 0.37, respectively) (Leary et al. 1985). Our heritability estimate for egg weight in Likes Creek pink salmon (0.22) was also similar to a heritability estimate for egg weight in captive Chinook salmon of 0.26 (Heath et al. 2003), although egg weight was not significantly heritable for pink salmon, but was significant for Chinook salmon.

The observation of higher heritabilities for morphological traits than for life-history traits suggests that in this population of pink salmon, the morphological traits examined are expected to respond more strongly to selection than egg number and egg weight. It also suggests that egg number and egg weight have been more strongly selected in the past than length, fin ray counts, and gill arch raker counts.

Because pink salmon are semelparous, like all Pacific salmon, there is likely very strong directional selection for increased egg number. There also may be directional selection for increased egg weight, since egg weight is positively related to embryo and fry survival in captive Chinook salmon (Heath et al. 2003). Alternatively, low heritabilities of life-history traits may reflect high levels of environmental variance in these traits (Price and Schluter 1991).

Although length was significantly heritable when estimated using adult progeny, it was not significant when estimated using progeny in the fry stage (Figure 2). This suggests that the inherited variation in physiological and behavioral traits that influence the length of returning spawners is not expressed until the oceanic life-history stages. This result is in agreement with previous work showing that male progeny sired by large males do not begin growing faster than progeny sired by small males until the spring of the year of maturity (Beacham and Murray 1988). In contrast, heritabilities of meristic traits can be estimated from regressions of juvenile offspring on adult parents because final meristic counts are usually determined early in development. For example, Leary et al. (1984) found no significant differences in meristic counts of rainbow trout analyzed 182 days after fertilization and fish from the same families analyzed 394 days after fertilization.

High Genetic Correlations of Nonadjacent Morphological Traits

We found that pairs of nonadjacent morphological traits had genetic correlations as high as pairs of adjacent traits (Table 3). Genetic correlations between pectoral and pelvic fin rays, pectoral fin rays and lower gill arch rakers, and pelvic fin rays and upper gill arch rakers ranged from 0.30 to 0.76, similar to the genetic correlation of 0.71 between upper and lower gill arch rakers, which are adjacent. The high genetic correlation between pectoral and pelvic fin ray counts is not surprising since they are similar structures and therefore might be expected to be under similar gene control. However, the observation of high genetic correlations between pectoral fin rays and lower gill arch rakers and between pelvic fin rays and upper gill arch rakers is somewhat surprising since they are both nonadjacent and different “types” of structures. This result suggests that different types of morphological traits may be regulated by similar genes regardless of the distance among the traits. This explanation is supported by the work of Leary et al. (1984), showing that expression of *Pgm1* by rainbow trout heterozygous or homozygous for a rare allele at the regulatory *Pgm1-t* locus developed faster and had lower counts of all meristic traits than individuals homozygous for the common allele.

Prediction of Phenotypic Correlations from Genetic Correlations

In general, genetic correlations accurately predicted phenotypic correlations for this population of pink salmon, as expected (Figure 3). The regression of parental phenotypic correlations on genetic correlations was not significant,

but this was likely due to relatively imprecise estimates of phenotypic correlations for parents due to fairly small sample sizes. The positive relationship between genetic correlations and phenotypic correlations in this population of pink salmon is in line with other studies showing a similar trend (Cheverud 1988, 1995; Roff 1995, 1996). Another similarity between this study and others is that the regression of phenotypic correlations on genetic correlations is less than unity (Kohn and Atchley 1988; Koots et al. 1994; Searle 1961). Some have argued that the regression coefficient converges on one when sampling variance is removed (Cheverud 1988, 1995; Roff 1995, 1996), but Willis et al. (1991) challenged the results of Cheverud (1988). Because phenotypic correlations are a function of genetic correlations as well as environmental correlations, there should be some correspondence between phenotypic and genetic correlations. However, further study is needed to understand whether the observation of smaller phenotypic correlations than genetic correlations is due entirely to sampling error or whether it has a biological explanation.

The outlier data points of the regressions of phenotypic correlations on genetic correlations provide insights into the interactive effects of genetic and environmental variation on phenotypic correlations. Pairs of traits with high phenotypic correlations despite low genetic correlations are likely phenotypically correlated because of a common effect of environmental variation on the two traits in question. For Likes Creek pink salmon, such pairs of traits included length and egg weight and length and egg number. Therefore, whatever environmental factors cause an increase in fish length, perhaps food resources, apparently cause an increase in egg weight and egg number.

In contrast, pairs of traits with high genetic correlations, but low phenotypic correlations, have low phenotypic correlations because of environmental variation that acts in opposition to the genetic correlations. The phenotypic correlation between pelvic and pectoral fin ray counts was only 0.02 in parents despite a genetic correlation of 0.64 in Likes Creek pink salmon, suggesting that common environmental variation has opposing effects on these two traits. Pelvic fin rays and upper gill arch rakers also had a phenotypic correlation of 0.08 in progeny despite a genetic correlation of 0.76, once again pointing to a negative environmental correlation between these two traits.

High Variability of Male Length

Length was significantly more variable among males than among females, as has been observed previously for pink salmon (Table 1) (Beacham and Murray 1985). Beacham and Murray (1985, 1988) suggest that the high variability of male length in pink salmon may stem from alternative large- and small-male breeding strategies. They note that small males resemble females, suggesting that small males may mimic females to reduce aggression from larger males.

A prediction of Beacham and Murray's hypothesis for high variability of male length in pink salmon is that length should not be more variable in males than in females within

year classes in other salmonid species in which there are multiple year classes of males on the spawning grounds. In species with multiple year classes on the spawning grounds, younger males are so much smaller than older males that there would be no advantage of being a small, older male. Therefore all males that return to spawning grounds later should be selected to grow large and there should be little variability in length within year classes. We are unaware of any published studies showing a greater variability of male length than female length within year classes for any salmonid species other than pink salmon, supporting Beacham and Murray's hypothesis. However, within-year class variability may be overlooked in salmon species with multiple year classes present on the spawning grounds because the difference in size among year classes is so much greater than length variability within year classes.

Generational Variation in the Phenotypic Correlation Between Length and Egg Number

The high phenotypic correlation observed between length and egg number observed in daughters was not seen in mothers. This highlights the fact that phenotypic correlations, like phenotypic variation, are highly dependent on environmental variation. Whatever common environmental factor caused the strong phenotypic correlation between length and egg number in daughters was not present in the parental environment.

Moreover, this observation has important implications for the relationship between fitness, defined as the number of returning progeny per parent, and length in pink salmon. If fitness is proportional to egg number and egg number is phenotypically correlated with length only in some years, one should expect to see temporal variation in the relationship between length and fitness. Our data suggest that fitness should be correlated with egg number because no negative phenotypic or genetic correlation was observed between egg number and egg weight, suggesting that there is no tradeoff between fecundity and egg size in this population. The lack of a phenotypic correlation between length and egg number in the parents suggests that length should not be correlated with fitness in this generation, assuming that other factors do not influence the relationship between fitness and length. In fact, no relationship was seen between fitness and length during the parental generation (unpublished data). In contrast, the high positive correlation between length and egg number in the offspring generation suggests that there should be a strong correlation between length and fitness in this generation.

Acknowledgments

We thank the Alaska Department of Fish & Game Gene Conservation Lab for their assistance in this project. The research described in this article was supported by the Exxon Valdez Oil Spill Trustee Council (project/190). However, the findings and conclusions presented by the authors are their own and do not necessarily reflect the views of the Trustee Council. We also thank D. Knudsen, K. Sage, B. Mullaly, S. Inglis, J. Moreland, R. Hocking, A. D'Alessandro, and the staff of the Alaska SeaLife Center for their assistance in capturing and raising fish. Thanks also to E. Steinberg,

M. Skinner, A. Maxwell, P. Spruell, K. Bott, J. Gallea, J. Meng, and J. Herndon at the Wild Trout and Salmon Genetics Lab, University of Montana, Missoula, for help with lab work and analyses. This project was approved by the University of Montana Institutional Animal Care and Use Committee (ACC 021-98).

References

- Banks MA, Blouin MS, Baldwin BA, Rashbrook VK, Fitzgerald HA, Blankenship SM, and Hedgecock D, 1999. Isolation and inheritance of novel microsatellites in chinook salmon (*Oncorhynchus tshawytscha*). *J Hered* 90:281–288.
- Beacham TD and Murray CB, 1985. Variation in length and body depth of pink salmon (*Oncorhynchus gorbuscha*) and chum salmon (*Oncorhynchus keta*) in southern British Columbia, Canada. *Can J Fish Aquat Sci* 42:312–319.
- Beacham TD and Murray CB, 1988. A genetic analysis of body size in pink salmon (*Oncorhynchus gorbuscha*). *Genome* 30:31–35.
- Boag PT and Grant PR, 1978. Heritability of external morphology in Darwin's finches. *Nature* 274:793–794.
- Cheverud JM, 1988. A comparison of genetic and phenotypic correlations. *Evolution* 42:958–968.
- Cheverud JM, 1995. Morphological integration in the saddle-back tamarin (*Saguinus fuscicollis*) cranium. *Am Nat* 145:63–89.
- Conner J and Via S, 1993. Patterns of phenotypic and genetic correlations among morphological and life-history traits in wild radish, *Raphanus raphanistrum*. *Evolution* 47:704–711.
- Endler JA, 2000. Adaptive genetic variation in the wild. In: *Adaptive genetic variation in the wild*, vol 1 (Mousseau TA, Sinervo B, and Endler J, eds). New York: Oxford University Press; 251–260.
- Fisher RA, 1930. *The genetical theory of natural selection*. New York: Dover.
- Gustafsson L, 1986. Lifetime reproductive success and heritability: empirical support for Fisher's fundamental theorem. *Am Nat* 128:761–764.
- Heath DD, Heath JW, Bryden CA, Johnson RM, and Fox CW, 2003. Rapid evolution of egg size in captive salmon. *Science* 299:1738–1740.
- Kohn LA and Atchley WR, 1988. How similar are genetic correlation structures? Data from mice and rats. *Evolution* 42:467–481.
- Koots KR, Gibson JP, and Wilson JW, 1994. Analyses of published genetic parameter estimates for beef production traits. 2. Phenotypic and genetic correlations. *Anim Breed Abstr* 62:825–853.
- Kruuk LE, Clutton-Brock TH, Slate J, Pemberton JM, Brotherstone S, and Guinness FE, 2000. Heritability of fitness in a wild mammal population. *Proc Natl Acad Sci USA* 97:698–703.
- Leary RF, Allendorf FW, and Knudsen KL, 1984. Major morphological effects of a regulatory gene: *Pgm1-t* in rainbow trout. *Mol Biol Evol* 1: 183–194.
- Leary RF, Allendorf FW, and Knudsen KL, 1985. Inheritance of meristic variation and the evolution of developmental stability in rainbow trout. *Evolution* 39:308–314.
- Lindner KR, Seeb JE, Habicht C, Knudsen KL, Kretschmer E, Reedy DJ, Spruell P, and Allendorf FW, 2000. Gene-centromere mapping of 312 loci in pink salmon by half-tetrad analysis. *Genome* 43:538–549.
- Lynch M and Walsh B, 1998. *Genetics and analysis of quantitative traits*, vol 1. Sunderland, MA: Sinauer Associates.
- Mousseau TA and Roff DA, 1987. Natural selection and the heritability of fitness components. *Heredity* 59:181–197.
- Olsen JB, Bentzen P, and Seeb JE, 1998. Characterization of seven microsatellite loci derived from pink salmon. *Mol Ecol* 7:1087–1089.
- Price T and Schluter D, 1991. On the low heritability of life-history traits. *Evolution* 45:853–861.
- Roff DA, 1995. The estimation of genetic correlations from phenotypic correlations: a test of Cheverud's conjecture. *Heredity* 74:481–490.
- Roff DA, 1996. The evolution of genetic correlations: an analysis of patterns. *Evolution* 50:1392–1403.
- Roff DA, 1997. *Evolutionary quantitative genetics*, vol 1. Montreal: Chapman & Hall.
- Sakamoto T, Danzmann RG, Gharb K, Howard P, Ozaki A, Khoo SK, Worman RA, Okamoto N, Ferguson MM, Holm L-E, Guyomard R, and Hoyheim B, 2000. A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. *Genetics* 155:1331–1345.
- Sanchez JA, Clabby C, Ramos D, Blanco G, Flavin F, Vazquez E, and Powell R, 1996. Protein and microsatellite single locus variability in *Salmo salar* L. (Atlantic salmon). *Heredity* 77:423–432.
- Scribner KT, Gust JR, and Fields RL, 1996. Isolation and characterization of novel microsatellite loci: cross-species amplification and population genetic applications. *Can J Fish Aquat Sci* 53:685–693.
- Searle SR, 1961. Phenotypic, genetic and environmental correlations. *Biometrics* 17:474–480.
- Spruell P, Pilgrim KL, Greene BA, Habicht C, Knudsen KL, Lindner KR, Olsen JB, Sage GK, Seeb JE, and Allendorf FW, 1999. Inheritance of nuclear DNA markers in gynogenetic haploid pink salmon. *J Hered* 90: 289–296.
- Weber KE, 1992. How small are the smallest selectable domains of form? *Genetics* 130:345–353.
- Willis JH, Coyne JA, and Kirkpatrick M, 1991. Can one predict the evolution of quantitative characters without genetics? *Evolution* 45:441–444.
- Wright S, 1968. *Evolution and the genetics of populations*. I. Genetic and biometric foundations. Chicago: University of Chicago Press.
- Young HJ, Stanton ML, Ellstrand NC, and Clegg JM, 1994. Temporal and spatial variation in heritability and genetic correlations among floral traits in *Raphanus sativus*, wild radish. *Heredity* 73:298–308.

Received January 2, 2004

Accepted July 31, 2004

Corresponding Editor: Brian Bowen