Exxon Valdez Oil Spill
Restoration Project Annual Report

Construction of a Linkage Map for the Pink Salmon Genome

Restoration Project 96190
Annual Report

This annual report has been prepared for peer review as part of the *Exxon Valdez* Oil Spill Trustee Council restoration program for the purpose of assessing project process. Peer review comments have not been addressed in this annual report.

Fred W. Allendorf
Paul Spruell
Kathy L. Knudsen
Kate R. Lindner
Kristine L. Pilgrim

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

April 1997

Construction of a Linkage Map for the Pink Salmon Genome

Restoration Project 96190
Annual Report

Study History: This project began in FY 96. Two oral presentations have been given at professional meetings on the project. Fred Allendorf gave an overview of the project at the national meeting of the American Fisheries Society in Dearborn, Michigan, August 1996. Kate Lindner presented accumulated data during the semi-annual Coast-wide Salmonid Genetics meeting in Seattle, March 1997. A journal article (Spruell, P., B. A. Greene, C. Habicht, K. L. Knudsen, K. R. Lindner, K. L. Pilgrim, J. E. Seeb, and F. W. Allendorf. Inheritance of nuclear DNA markers in haploid pink salmon embryos.) was submitted to Molecular Ecology.

Abstract: We began construction of a detailed genomic linkage map for pink salmon (*Oncorhynchus gorbuscha*). The completed map will allow oil induced lesions to be thoroughly described and will aid other pink salmon recovery efforts. In August 1995 families of haploid pink salmon were produced using gametes collected at the Armin F. Koernig Hatchery in Prince William Sound. The resulting progeny were screened with DNA markers to identify and remove any embryos that were not haploid. We selected one reference family and have analyzed the segregation pattern of 181 DNA fragments in the maternal parent and each of 94 haploid embryos from that family. Of these fragments, 137 have been assigned to one of 41 linkage groups. We have modified our study plan to use techniques that provide more polymorphic marker per unit effort than the technique originally proposed. We are currently using these techniques to accumulate polymorphic markers and analyze segregation ratios in the reference family.

Key Words: Adaptation, fitness, genetics, lesions, linkage map, marine survival, *Oncorhynchus gorbuscha*, pink salmon, stock structure.

Project Data: Data are inheritance of 181 polymorphic DNA fragments in 94 progeny of a single family of haploid pink salmon. The presence or absence of each fragment in each individual is recorded in an Excel spread sheet. Data will be made available to individuals within the reasonable bounds of sharing unpublished data. For information regarding data contact Kate Lindner, Division of Biological Sciences, University of Montana, Missoula, MT 59812. Phone: (406) 243-5503. E-mail: klindner@selway.umt.edu

Citation:

Allendorf, F.W., P. Spruell, K.L. Knudsen, K.R. Lindner, and K.L. Pilgrim. 1997. Construction of a linkage map for the pink salmon genome, *Exxon Valdez* Oil Spill Restoration Project Annual Report (Restoration Project 96190), Alaska Department of Fish and Game, Habitat and Restoration Division, Anchorage.

TABLE OF CONTENTS

LIST OF TABLES 4

LIST OF FIGURES 5

EXECUTIVE SUMMARY 6

INTRODUCTION 8

OBJECTIVES 10

METHODS 11
- Gamete Collection and DNA Isolation 11
- Identification of Useful Markers 12
- RAPDs 12
- AFLPs 12
- PINEs 13
- Map Construction 13

RESULTS 13
- Confirmation of Haploidy 13
- Nomenclature 14
- Individual Segregation 15
- Microsatellites 15
- RAPDs 15
- AFLPs 16
- PINEs 16
- Joint Segregation 16

DISCUSSION 17

CONCLUSIONS 18

ACKNOWLEDGMENTS 18

LITERATURE CITED 18

OTHER REFERENCES 20
- Personal Communications 20

LIST OF TABLES

1. Summary of primer pairs screened and polymorphic loci detected by four different techniques. The first column is the number of primers amplified with 6-8 individuals to determine if polymorphic loci were produced. The second column is the number of polymorphic loci detected by each technique and included in the linkage analysis. The third column is the percent of polymorphic loci that have been assigned to a linkage group. 22
2. Number of polymorphic fragments generated by each primer combination in PINEs. all fragments scored segregate in expected Mendelian ratios. "NA" indicates primer combinations that have not yet been analyzed. 23
3. Summary of linkage groups in the pink salmon genome map. 24

LIST OF FIGURES

1. Growth hormone pseudogene amplification products separated in a 2% agarose gel. The first two lanes indicated by the bar on the bottom left of the gel are known males, the third lane is female 95-103. The arrows along the bottom of the gel indicate males that were detected in family 95-103. 25
2. A representative RAPD gel. Bands were separated in a 2% agarose gel. DNA was amplified with Operon Technology Inc. primer J-10. Polymorphic bands are indicated by arrows on the right and scored from the top as *10J516* and *10J500*. 26
3. A representative AFLP gel. Bands were separated in 7% polyacrylamide. DNA was amplified using the primer pair AGG/CAA from the Perkin/Elmer AFLP selective amplification start up module. Polymorphic bands are indicated by the arrows on the right side. the bands were scored starting from the top of the gel as *AGG/CAA185, AGG/CAA180, AGG/CAA175, AGG/CAA140, AGG/CAA129* and *AGG/CAA110*. 27
4. A representative PINE gel. DNA was amplified with the 5' FokI and 5' Tc1 primer. Bands were separated electrophoretically in a 4.5% polyacrylamide gel. Polymorphic bands are indicated by the arrows on the right side. Starting from the top of the gel the bands are scored as *5F5T308*, *5F5T281*, *5F5T278*, *5F5T232*, and *5F5T217*. 28
5. Genetic linkage map of pink salmon. Numbers to the left of the linkage groups indicate recombination rates (centimorgans, cM). To the right of the linkage groups are the locus names. 29

EXECUTIVE SUMMARY

The construction of a genetic linkage map for the pink salmon (*Oncorhynchus gorbuscha*) genome is in its second year. This linkage map will allow the characterization of the genetic impacts of the March 1989 *Exxon Valdez* oil spill on pink salmon populations and help to document the recovery of effected populations in Prince William Sound. A genetic linkage map will be essential for detecting and understanding causes of reduced egg and embryo survival in oiled areas. In addition, a detailed genetic map 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 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. This suggests that genetic damage may have occurred as a result of exposure to oil during early developmental life-stages.

The map is being constructed using haploid progeny to avoid the difficulties associated with dominant markers that may obscure recessive alternatives in diploids. Gametes and tissues of pink salmon were collected from the Armin F. Koernig hatchery in August of 1995. Families of gynogenetic haploid embryos were produced in cooperation with the Alaska Department of Fish and Game by mixing irradiated sperm with eggs from individual females.

One family (number 95-103) was chosen to be the reference family upon which our initial mapping efforts are based. Embryos were confirmed to be haploid using a sex-specific pseudogene locus to identify diploid males and several microsatellites to identify diploid females. After removal of diploid individuals, family 95-103 consists of 94 haploid embryos and their maternal parent.

We are using various methods to detect the large number of polymorphic loci necessary to saturate a linkage map. Initially we focused on randomly amplified polymorphic DNA but this resulted in a limited number of polymorphic loci per reaction. We are currently using two techniques that detect polymorphic loci much more efficiently, Amplification Fragment Length Polymorphisms and Paired Interspersed Nuclear Elements.

We have assigned 137 of the 181 markers analyzed (76%) to one of 41 linkage groups. These 41 linkage groups have two to seven markers each at an average interval of 9.7 cM. The estimated size of the total pink salmon linkage map based on these data is 3,920 cM. This includes 1,328 cM mapped, 504 cM to account for the distance from the end markers to their adjacent telomeres, and 2124 cM in unfilled gaps in the map. The haploid pink salmon genome is approximately 2.72 billion bp; thus, we estimate a physical recombination rate of approximately 694 kbp/cM.

Our results are consistent with the maps constructed in zebrafish (*Danio rerio*) and the Japanese medaka (*Oryzias latipes*). The medaka map contains 227 markers, of which 71% have been assigned to a linkage group. This is very similar to the 76% of the markers that have been assigned to a linkage group in our project. Although our current data set of 181 markers is incomplete, our estimate of total map size is close to what we expected on the basis of the polyploid ancestry of salmonids.

We have made substantial progress toward the completion of a genetic linkage map in pink salmon. New techniques such as Amplified Fragment Length Polymorphisms and Paired Interspersed Nuclear Elements will allow us to consolidate the map earlier than we originally proposed and do so using a single family. As a result of our increased efficiency, we will complete this part of the project ahead of schedule and require less support than we had anticipated from the *Exxon Valdez* Oil Spill Trustee Council.

INTRODUCTION

The construction of a genetic linkage map for the pink salmon (*Oncorhynchus gorbuscha*) genome is in its second year. This map will provide a more thorough understanding of the genetic characteristics of pink salmon. This map will allow the characterization of the genetic impacts of the March 1989 *Exxon Valdez* oil spill on pink salmon populations and help to document the recovery of effected populations in Prince William Sound.

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 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.

A detailed genetic map for pink salmon will be invaluable for documenting the effects that the spill may have had on pink salmon. First, it will be possible by following the inheritance of any DNA lesions identified in Restoration Study \191A to determine if they are micro- or macro-lesions. Second, these lesions can 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. A detailed genetic map will also aid other pink salmon recovery efforts, including estimation of straying rates, description of stock structure, and testing for a genetic basis of marine survival.

The construction of a linkage map requires analyzing the genetic transmission of several hundred DNA polymorphisms. These markers must meet several criteria to be useful for mapping purposes. The techniques used to generate these markers must be reproducible and ideally would allow the detection of multiple loci from a single polymerase chain reaction (PCR). The fragments detected by each technique should be dispersed at random throughout the genome to avoid exclusion of specific chromosomal regions. Finally, the markers must be present in the maternal parent and segregate in a 1:1 ratio as expected in simple Mendelian inheritance models.

Many of the techniques commonly used for linkage map construction result in variation that is measured by the presence or absence of a given PCR product. Fragments that segregate in this "dominant" (present) versus "recessive" (absent) manner are problematic in that it is impossible to distinguish individuals that are homozygous dominant (two alleles that produce the DNA fragment) from those that are heterozygous (one allele produces the fragment, the other does not). We are avoiding the difficulties of dominance with these markers by using haploid progeny in which recessive alleles are not obscured by their dominant alternatives (Lie et al. 1994). Although these embryos are not viable, development progresses until just prior to hatching (Stanley 1983), providing an embryo from which we obtain a sufficient quantity of DNA for analysis.

To ensure the accuracy of our map, each individual was confirmed to be haploid before it was included in the reference family. Techniques that simultaneously amplify multiple products are advantageous for construction of a linkage map. However, these techniques have limited power to confirm that an individual is haploid. Three methods were used to detect diploid individuals in the putative haploid families. Flow cytometry was conducted on a sample of embryos to measure the proportion of embryos for which haploid induction was successful. A male-specific locus and 5 polymorphic microsatellite loci were then used to identify individual embryos that were diploid. Diploid individuals were excluded from further analyses.

Our initial efforts focused on two closely related techniques, randomly amplified polymorphic DNA (RAPDs) (Williams et al. 1990) and arbitrarily primed PCR (AP-PCR) (Welsh and McClelland 1990) that have been used to identify polymorphic DNA fragments in many other mapping studies (Postlethwait et al. 1994; Wada et al. 1995). These techniques use anonymous sequences as PCR primers to simultaneously amplify multiple DNA fragments. Although these methodologies have been very useful in other taxa, we detected limited polymorphism in the haploid family that we are using to construct the pink salmon linkage map. Therefore, we began to explore other methods to generate the necessary polymorphisms.

Amplification fragment length polymorphisms (AFLPs) have been used extensively in the construction of genomic maps in plants (Becker et al. 1995, Maheswaran et al. 1997). AFLP analysis consists of three steps. The first step is the "restriction/ligation" step. Two restriction enzymes are used to cut the genomic DNA into many fragments. Double stranded adapters that are specific to the restriction sites are then ligated onto the fragments. The second step is the "pre-selective amplification". During this step the restriction fragments are amplified using two primers that are specific to the synthetic adapters. Each of these primers includes an additional one base extension into the genomic DNA fragment flanked by the adapters. This step amplifies only DNA fragments with those specific bases on each end, reducing the number of DNA fragments available for subsequent amplification. The final step, "selective amplification," uses an aliquot of the pre-selective products as DNA template. Amplification is conducted with primers that are specific to the synthetic adapters with three additional "selective" bases extending into the genomic DNA fragment. The increasing specificity of the primers used to amplify the fragments results in clean, reproducible banding patterns.

The AFLP technique is especially useful for our purposes for two reasons. First, many bands are produced per reaction and, therefore, more scoreable polymorphic loci are produced per unit effort. Second, the selective amplification step uses a subsample of the PCR products of the preamplification. Up to 133 selective amplifications can be completed from a single pre-amplification that originally used only 0.5 μg of genomic DNA. Much more genomic DNA is needed to produce fewer bands using other methods such as RAPDs. This is an important consideration when dealing with the limited amount of tissue available from embryos.

In addition to RAPDs and AFLPs, we are using a technique that uses various repetitive DNA elements that are scattered throughout the genome of salmonid fishes. Kido et al. (1991) described 3 SINEs (short interspersed elements) found in specific members of the family salmonidae. 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. SINEs and transposons occur in high copy number and are believed to be ubiquitously dispersed throughout the genome, making them ideal candidates for genomic mapping efforts.

We have used DNA sequences homologous to salmonid-specific SINEs and the transposon Tc1 as primers to generate multiple DNA fragments from a single PCR reaction a technique we call PINEs (Paired Interspersed Nuclear Elements). The theoretical basis for this procedure is similar to the use of the human SINE AluI to identify human chromosomes in somatic cell hybridization experiments. 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. Greene and Seeb (submitted) used this technique to confirm the parentage of pink salmon fry, demonstrating the potential for including these fragments in our mapping study.

Our study plan has been adapted to utilize newly developed techniques that are more efficient than the RAPDs we had proposed to use in the construct of the map. Both PINEs and AFLPs have proven to be more polymorphic than RAPDs and highly reproducible. Based on the development of these more efficient methodologies, we are confident that we will be able to complete the pink salmon linkage map more rapidly than we initially proposed and at a reduced expense.

Information gained from this study will provide resource managers with insight into the magnitude and persistence of damages sustained by wild pink salmon due to the spill. Efforts to restore damaged pink salmon populations depend upon the ability of fishery managers to identify causes of reduced survival and to monitor their persistence. The potential of long term oil exposures to cause genetic damage needs to be understood so that spawning escapement goals can be adjusted if necessary. In addition, verification of the genetic hypothesis would provide the first evidence that the germline of fish can be affected by exposure to chronic or acute sources of oil pollution .

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.

In this report we review the progress made toward the completion of objectives 1 and 2. This study is ongoing. Objectives 3-6 will be addressed in future years.

METHODS

Gamete Collection and DNA Isolation

Gametes and tissues of pink salmon were collected from the Armin F. Koernig hatchery in August of 1995. Families of gynogenetic haploid embryos were produced in cooperation with the Alaska Department of Fish and Game by mixing irradiated sperm with eggs from individual females. Embryos from these families were incubated until just prior to hatching at which time the eggs were preserved in alcohol and sent to the University of Montana for DNA extraction and genetic analysis.

DNA was isolated from one family of 150 haploid embryos as well as muscle and liver tissue from their parents. DNA was isolated from pink salmon embryos using standard Purgene(TM) protocol (Gentra Systems Inc.). The concentration of DNA was determined using a scanning spectrofluorometer.

A single family (number 95-103) was chosen as the reference family to be used to generate the linkage map. An acceptable number of embryos were produced in this family. In addition, initial screening with microsatellites was consistent with the embryos being predominately haploid and confirmed Mendelian segregation. Flow cytometric analysis confirmed that approximately 90% of the embryos were haploids (Chris Habicht, pers. comm.). However, it is not possible to do flow cytometric confirmation on those individuals to be used in the linkage map.

Analysis of the sex-specific pseudogene followed published methods (Devlin et al 1991, Forbes et al. 1994). Male pink salmon are clearly distinguishable from females based on the presence of a lower molecular weight band that is absent in females (Figure 1). PCR primer sequences were: left, 5'-tttctctacgtctacattct-3'; right 5'-gtctggctagggtactcca-3' (courtesy R. H. Devlin 1996). Ten-µl PCR mixtures contained 10 mM Tris-HCL pH 8.3, 50 mM KCl, 4.0 mM $MgCl_2$, 0.2 mM of each dNTP, 4.5 pmoles of each primer, 0.5 U Taq DNA polymerase and 50-100 ng purified genomic DNA. The PCR profile was 93°C for 1 minute, and thirty cycles of 92°C for 1 minute, 54°C for 1 minute and 72°C for 2 minutes. PCR products were electrophoresed on a 2.0% agarose gel containing ethidium bromide in TAE buffer (Ausubel et al. 1989).

Embryos in family 95-103 were also screened with five microsatellites to identify diploid females. Ten-µl PCR mixtures contained 10 mM Tris-HCL pH 8.3, 50 mM KCl, 4.0 mM $MgCl_2$, 0.2 mM of each dNTP, 4.5 pmoles of each primer, 0.5 U Taq DNA polymerase and 20 ng purified genomic DNA. PCR products were electrophoresed on a 7% denaturing polyacrylamide gel and visualized using a Hitachi FMBIO-100 fluorescent imager.

Identification of Useful Markers

Genetic markers are included on the map based on two criteria. First, polymorphic fragments in the haploids had to be present in female 95-103. Second, the segregation of each fragment in the progeny had to be 1:1 as expected under simple Mendelian genetic models. 181 fragments met this expectation and were included in the analysis of joint segregation ratios.

RAPDs

Simultaneous amplification of RAPD fragments was completed on DNA from the entire family of 94 haploid progeny and their mother using 96 well microtiter plates. A total reaction volume of 10 µL consisted of 0.2mM of each dNTP, 4mM $MgCl_2$, 6.7mM Operon Technology Inc. RAPD primer, 0.25U Amplitaq DNA polymerase Stoffel fragment (Perkin/Elmer), 1X Stoffel buffer, 10ng of DNA and HPLC H_20 to a volume of 10µl. Thermal cycling was performed in a MJ Research, PTC-200 DNA engine. Two cycles of higher stringency PCR were performed with the following thermal profile: denaturation at 96°C for 5 seconds, annealing at 40°C for 20 seconds, and extension at 72°C for 30 seconds. This was followed by 43 cycles with an annealing temp of 36°C for 20 seconds and a final extension at 72°C for 2 minutes. PCR products were mixed with a bromophenol blue dye and loaded on a 2% agarose gel and subjected to electrophoresis. Gels were soaked in an ethidium bromide solution and products were visualized using a Hitachi FMBIO-100 fluorescent imager.

AFLPs

The AFLP restriction/ligation and pre-selective amplification steps were completed following the Perkin-Elmer/Applied Biosystems AFLP plant mapping protocol with a few modifications outlined below. The following thermal profile was used for the pre-selective

amplification: 72°C for 2 minutes, 30 cycles of denaturing at 94 °C for 1 second, annealing at 60 °C for 30 seconds, extension at 72 °C for 2 minutes. The selective amplification was completed using 1.5 µL of the pre-selective amplification products as DNA template, 0.5 µl EcoRI selective primers, 0.5 MseI selective primers, 2mM $MgCl_2$, 0.17mM of each dNTP, 2X Amplitaq PCR buffer, 0.5U Amplitaq DNA polymerase, and HPLC H_2O as needed to result in a 10 µl total volume. The following thermal profile was used for the selective amplification: initial denaturation at 96°C for 2 minutes is 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 resulting in a final annealing temp of 59°C. An additional 30 cycles with the annealing temperature at 59°C for 30 seconds was completed. Products were mixed with formamide loading buffer, denatured by heating to 80 °C for 15 min., and electrophoresed on a 7% polyacrylamide gel. The gel was scanned using an Hitachi FMBIO image scanner.

PINEs

PCR amplification of DNA fragments flanked by interspersed elements were conducted in a total volume of 10 µl. Each reaction contained approximately 25 ng of genomic DNA, 1 µl 1X Perkin-Elmer PCR buffer, 4.5 mM $MgCl_2$,0.2 mM of each dNTP, 5.0 pmoles of primer and 0.5 U Taq. Reactions were completed in an MJ Research thermocycler under the following profile: 94C for min. followed by cycles of 94C for min., 60C for min. 74C for min.

Map Construction

We used MapMaker software (Lander et al. 1987) to assign markers to linkage groups. The program determines which fragments are linked by comparing the segregation patterns of each locus in all individuals. We do not know the linkage phase of the markers we are using without examining the segregation patterns. Therefore, the segregation pattern of each locus is entered into the program in both possible phases. The more similar the segregation pattern between loci the closer they are on the chromosome. The proportion of recombinant types must be 0.365 or less for us to consider a pair of loci to be linked. Those that exceed this value are not significantly ($P<0.01$) different than expected due to independent assortment.

RESULTS

Confirmation of Haploidy

All progeny in family 95-103 were screened to confirm that there was sufficient DNA for subsequent analyses and that they were haploid. Spectrophotometry determined that DNA extractions from 12 individuals did not yield a sufficient quantity of DNA for subsequent analyses.

To prevent any misinterpretations that might result from including diploid individuals, we screened all members of the family 95-103 with two types of single locus DNA markers. The first, a Y chromosome-linked intron from a growth hormone pseudogene is found only in males (Devlin et al. 1991; Forbes et al. 1994). The presence of this marker in an individual is an unambiguous screen to detect diploid males. Haploids are produced by eliminating the paternal chromosome complement therefore, no Y-chromosomes should be found in a pure haploid family. Amplification of the sex-linked growth hormone pseudogene (Figure 1) identified 28 individuals that were males. These individuals were removed from the study.

Failure of haploid induction could also produce diploid females. Analysis with the sex-specific pseudogene will not identify these individuals. Five microsatellites were used to screen the remaining individuals. Fourteen additional individuals were confirmed to be diploid based on presence of two alleles at one or more microsatellite loci. These individuals were also eliminated from any additional analyses.

Nomenclature

The construction of a linkage map relies on various techniques to generate hundreds of markers. All of the loci included on the map must be named in a standard manner. This will allow for a clear understanding of the techniques used to produce the loci and will facilitate communication between labs. In general, each locus is named by alpha-numeric characters that designate the specific PCR primers used followed by the length of the amplified product in base pairs.

The nomenclature used to designate microsatellite loci uses the locus name described by the original authors. The names usually begin with a three letter abbreviation designating the genus and species from which the locus was isolated. This abbreviation consists of the first letter of the genus name capitalized, followed by the first two letters of the species name in lower case. The species designator is then followed by a number to indicate the specific locus. For example, *Ots1* (Dennis Hedgecock, pers. comm.) is microsatellite number one cloned from chinook salmon (*Oncorhynchus tshawytscha*). A few of the earliest microsatellites isolated do not conform to this standardized nomenclature (e.g. *μSAT60*, Estoup et al. 1993) (summarized by Olsen et. al 1996). One locus, *FGT1* (Sakamoto et al. 1994) was duplicated in pink salmon. The nomenclature for those loci follows the format used in allozyme analysis (*FGT1-1,2*; Shaklee et al. 1990)

The nomenclature for the loci produced using RAPD primers followed that used for the zebrafish (*Danio rerio*) genome map (Postlethwait et al. 1994). For example, *O5L250* is a 250 base pair fragment generated using Operon primer 5L.

The nomenclature for AFLPs follow the same format as Young and Thorgaard (pers. comm). The name starts with the three base selective primer extensions used to produce the loci and ends with the length of the fragment measured in base pairs; e.g. *AAA/CAT250.*

PINE loci are named using the same format as Spruell et al. (in prep.) The names start with a number designating the end of the element from which the primer was derived (3' or 5') followed by a one letter designation of the element from which the primer was derived. If more than one primer was used during the amplification the second primer is named in the same manner as the first and the primers are placed in alphabetical order. The name is followed by the length of the fragment measured in base pairs. For example, *5F3H250* is a 250 base pair fragment amplified using a primer homologous to the 5' end of the FokI SINE and the 3' end of the HpaI SINE. Consistent use of this nomenclature will minimize the confusion that might arise from arbitrarily assigned names.

Individual Segregation

One hundred and eighty-one fragments conformed to a 1:1 ratio ($P<0.001$). Table 1 provides a summary of the techniques used to amplify these fragments and the number of fragments detected by each technique that have been assigned to a linkage group.

Microsatellites

We screened 31 microsatellite loci in pink salmon. Unlike the other markers we are using, microsatellites are usually inherited in a codominant fashion, that is both alleles produce a visible amplification product. Female 95-103 is heterozygous at 5 microsatellite loci which have been added to the map (Table 1). Four of the five polymorphic microsatellites we have analyzed are inherited in a codominant manner, alternate alleles were scored based on varying sizes of PCR products.

The fifth locus, *Ssa197* (O'Reilly et al. 1996) was suspected to contain a null allele (an allele that results in no amplification product) based on previous population analyses in which a significant deficiency in heterozygotes had been detected (J. E. Seeb, pers. comm., K. L. Pilgrim, unpublished data). The presence of the null was confirmed by analyzing the segregation patterns for this locus in the haploid progeny of four females, including female 95-103, that produced a single PCR product. Half of the progeny from each female produced the same allele as the mother, the other half produced no PCR product, confirming the presence of a null allele at this locus. This locus was subsequently scored based on the presence of or absence of a PCR product and added to the map.

RAPDs

We devoted our initial efforts toward RAPDs. However, only 35 polymorphic markers were identified out of 140 primers that were screened. Each RAPD primer produced 2-8 bands 400-1500 bp in length (Figure 2). The low number of polymorphic bands results from both the low number of bands amplified and the failure of many primers to detect any polymorphism (Table 1).

AFLPs

The AFLP technique is much more productive than RAPDs. Each AFLP primer combination results in the amplification of at least 30 bands ranging from 60-450 bp (Figure 3). We have screened 14 of the 64 available primer combinations and have yet to use a combination that does not produce at least one useable polymorphism. Most primer combinations produce many more polymorphisms with an average of 7.5 polymorphic bands being produced by amplification with each primer pair (Table 1). Assuming the 50 remaining AFLP primer combinations produce a comparable number of polymorphic fragments, we should be able to accumulate approximately 375 additional loci using these 64 primer combinations. It is also possible to try additional primer combinations to generate even more AFLP markers.

PINEs

PINEs also produce many more polymorphic loci than RAPDs. Like AFLPs a minimum of 30 bands are amplified in each PCR (Figure 4). We have scored all 94 haploid offspring with 5 PINE primer combinations. An average of 13.4 polymorphic loci are produced with each of these combinations (Table 2). In two cases it appears that PINE fragments are segregating as codominant alleles that vary in length at a single locus. This is not unexpected, an insertion or deletion occurs between the SINE primer sites will produce a this result.

Joint Segregation

We have assigned 137 of the 181 markers analyzed (76%) to one of 41 linkage groups (Figure 5). These 41 linkage groups have two to seven markers each at an average interval of 9.7 cM (Table 3). The estimated size of the total pink salmon linkage map based on these data is 3,920 cM. This includes 1,328 cM mapped in Figure 5, 504 cM to account for the distance from the end markers to their adjacent telomeres, and 2124 cM in unfilled gaps in the map. The haploid pink salmon genome is approximately 2.72 billion bp (Johnson et al. 1987); thus, we estimate a physical recombination rate of approximately 694 kbp/cM.

There does not appear to be a bias in which techniques have generated the fragments that are included on the map. However, there does appear to be a tendency toward tight linkage of fragments amplified by PINEs. Seven of 11 pairs of loci that map within 2 map units are both PINE-derived fragments (Figure 5). Of those seven pairs, two are generated using the same PINE primers and map to the same location. These are probably length polymorphism alleles being codominantly inherited. Excluding those two pairs, five of nine (0.556) pairs of tightly linked loci are PINE-derived fragments. This is more than would be expected (0.173) based solely on the frequency of PINEs on the map. Four other pairs of fragments from various techniques also show absolute linkage.

DISCUSSION

Our results are consistent with the maps constructed in zebrafish (Postlethwait et al. 1994; Johnson et al. 1996) and medaka (*Oryzias latipes*; Wada et al. 1995). The medaka map contains 227 markers, of which 71% have been assigned to a linkage group. This is very similar to the 76% of the markers that have been assigned to a linkage group in our project. Although our current data set of 181 markers is incomplete, our estimate of total map size (3920 cM) is close to what we expected on the basis of the polyploid ancestry of salmonids. The zebrafish map is estimated to be 2900 cM (590 kbp/cM) and the medaka map is estimated to be 2480 cM (323 kbp/cM).

Our study plan has been altered slightly from that outlined in our initial proposal. The generation of polymorphic loci using RAPDs proved to be more difficult than we anticipated. Therefore, other techniques that are more efficient than RAPDs are currently being used. AFLPs are particularly valuable. Our current estimates indicate that we may detect sufficient polymorphisms to saturate the map using this technique alone. This technique uses two PCR amplifications which allows many more reactions to be done using a limited amount of genomic DNA. This improvement has eliminated the need to use multiple families in the construction of the map. We should be able to generate enough polymorphic markers using only individuals from family 95-103 to complete the map.

After we consolidate the number of linkage groups to the number of chromosomes (27; 25 autosomes, the X-chromosome, and the Y-chromosome). We will also begin placing loci used in population genetic analysis of pink salmon (allozymes and microsatellites) onto the map using centromere-linkage analysis of half-tetrads (Johnson et al. 1996). In addition, other loci of special importance will be placed onto the map (e.g., growth hormone loci and the major histocompatibility loci (MHC; Katagiri 1996). These loci will be used as landmarks to compare recombination rates in males and females and the linkage maps of odd- and even-year fish (O'Brien et al. 1993)

In addition to its management applications, the completion of this map will result in a significant contribution to our understanding of genomic organization and the distribution of repetitive DNAs. Our data are currently not sufficient to adequately address the distribution of fragments derived by particular techniques. However, there does appear to be a trend toward PINE-derived fragments mapping in close proximity to one another. The tight clustering of certain fragments might be explained by the biases in the genomic distribution of repetitive DNA. These sequences are generally thought to be ubiquitously distributed throughout the genome. However, mapping studies in other taxa have detected chromosomal regions that preferentially accumulate repetitive DNAs. Comparisons of these results to those of the pink salmon linkage map will be particularly interesting in light of the tetraploid history of salmonids (Allendorf and Thorgaard 1984).

CONCLUSIONS

We have made substantial progress toward the completion of a genetic linkage map in pink salmon. New techniques such as AFLPs and PINEs will allow us to consolidate the map more rapidly than we originally proposed and do so using a single family. As a result of our increased efficiency, we will complete this project ahead of schedule and require less support than we had anticipated from the EVOS trustee council.

ACKNOWLEDGMENTS

We thank James E. Seeb and Chris Habicht from the (ADFG) for providing samples and producing the haploid embryos for this project. In addition, we thank Jim Seeb for encouraging us to use the AFLP technique and his general harassment. Ben A. Greene (ADFG) provided helpful suggestions that led to our use of PINEs for mapping haploid pink salmon.

LITERATURE CITED

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.

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.

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.

Du, S. J., R. H. Devlin and C. H. Hew. 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-Ψ. DNA and Cell Biol. 12:739-751.

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

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. Natl. Acad. Sci. USA 91:1628-1631.

Devlin, R. H., B. K. McNeil, T. D. D. Groves and E. M. Donaldson, 1991. Isolation of a Y-chromosomal DNA probe capable of determining genetic sex in chinook salmon (*Oncorhynchus tshawytscha*) Can. J. Fish. Aquat. Sci. 48:1606-1612.

Greene, B. A. and J. E. Seeb. (in review) 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*).

Johnson, O. W., F. M. Utter, and P. S. Rabinovitch. 1987. 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.

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. Proc. Natl. Acad. Sci. 88:2326-2330.

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.

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.

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.

Olsen, J. B., J. K. Wenburg, and P. Bentzen, 1996. Semiautomated multilocus genotyping of Pacific salmon (*Oncorhynchus spp.*) using microsatellites. Mol. Mar. Biol. Biotech. 5:259-272.

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.

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

Postlethwait, J.H. and 14 other authors. 1994. A genetic linkage map for the zebrafish. Science 264:699-703.

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

Shaklee, J. B., F. W. Allendorf, D. C. Morizot, and G. S. Whitt. 1990. Gene nomenclature for protein-coding loci in fish. Trans. Am. Fish. Soc. 119:2-15.

Spruell, P., B. A. Greene, C. Habicht, K.L. Knudsen, K. R. Lindner, K. L. Pilgrim, J. E. Seeb, and F. W. Allendorf. (in prep.) Inheritance of nuclear DNA markers in haploid pink salmon embryos.

Spruell, P. and G. H. Thorgaard, 1996. SINE sequences detect DNA fingerprints in salmonid fishes. Heredity 317-324.

Stanley, J. G. 1983. Gene expression in haploid embryos of Atlantic salmon. J. Hered. 74:19-22.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hones, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: an new technique for DNA fingerprinting. Nucl. Acids Res. 23:4407-4414.

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.

Welsh, J. and M. McClelland, 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucl. Acids Res. 18:7213-7218.

Williams, J. G., A. R. Kubelik, K. J. Livak, J. A. Rafalski, and S. V. Tingey, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acids Res. 18:6531-6535.

OTHER REFERENCES

Personal Communications

Chris Habicht, Alaska Department of Fish and Game, Genetics Program, 333 Raspberry Road, Anchorage AK 99518.

Dennis Hedgecock, University of California, Davis, Bodega Marine Laboratory, Bodega Bay CA 94923.

James E. Seeb, Alaska Department of Fish and Game, Genetics Program, 333 Raspberry Road, Anchorage AK 99518.

Gary H. Thorgaard, Department of Zoology, Washington State University, Pullman, WA 99164-4236.

William P. Young, Department of Biological Sciences, Northern Arizona University Flagstaff, AZ 86011-5640.

Table 1. Summary of primer pairs screened and polymorphic loci detected by four different techniques. The first column is the number of primers amplified with 6-8 individuals to determine if polymorphic loci were produced. The second column is the number of polymorphic loci detected by each technique and included in the linkage analysis. The third column is the percent of polymorphic loci that have been assigned to a linkage group.

	# of Primer Pairs screened	Polymorphic loci detected	% of Loci assigned to a linkage group
RAPDs	140	35	71%
u-Sats	30	5	100%
PINE	5	67	85%
AFLP	14	75	65%
Total	189	182	

Table 2. Number of polymorphic fragments generated by each primer combination in PINEs. all fragments scored segregate in expected Mendelian ratios. "NA" indicates primer combinations that have not yet been analyzed.

	Primer				
Primer	HpaI5'	HpaI3'	FokI5'	SmaI5'	Tc15
HpaI5'	NA				
HpaI3'	NA	NA			
FokI5'	9	NA	NA		
SmaI5'	32	NA	10	NA	
Tc15'	5	NA	11	NA	NA

Table 3. Summary of linkage groups in the pink salmon genome map.

Number of markers in linkage group	Number of linkage groups	Average size (cM)
2	6	18.6
3	3	24.6
4	6	38.4
5	5	57.0
6	2	85.8
7	1	138.9

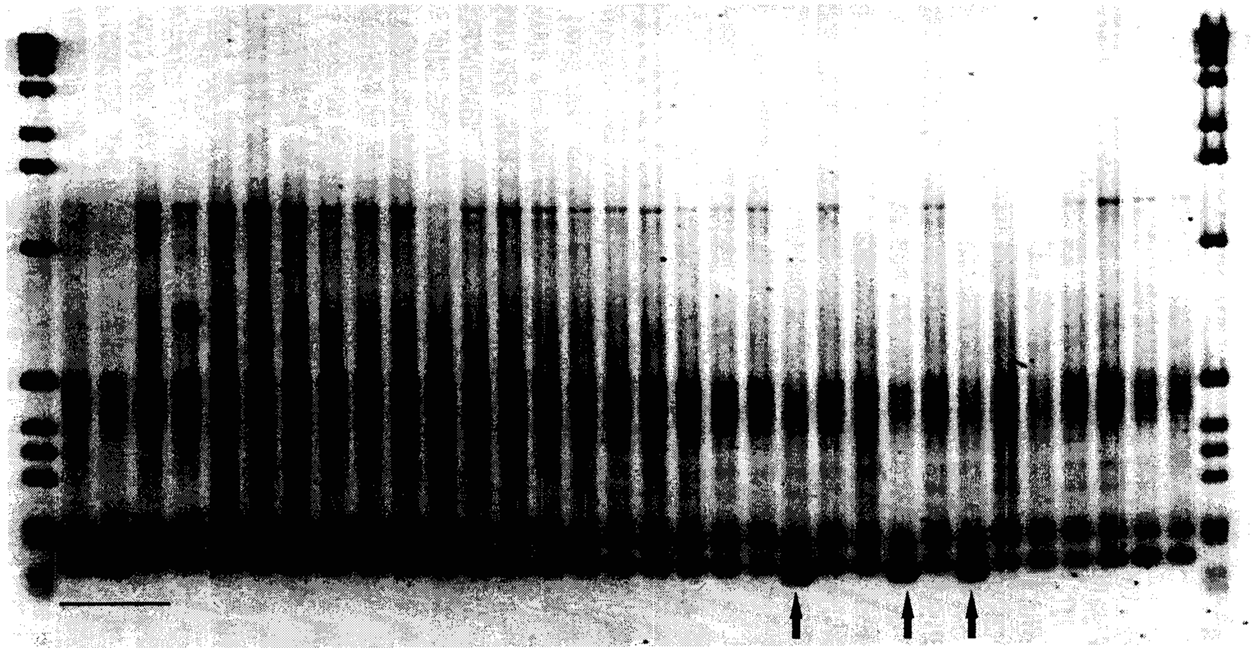

Figure 1
Growth hormone pseudogene amplification products separated in a 2% agarose gel. The first two lanes indicated by the bar on the bottom left of the gel are known males, the third lane is female 95-103. The arrows along the bottom of the gel indicate males that were detected in family 95-103.

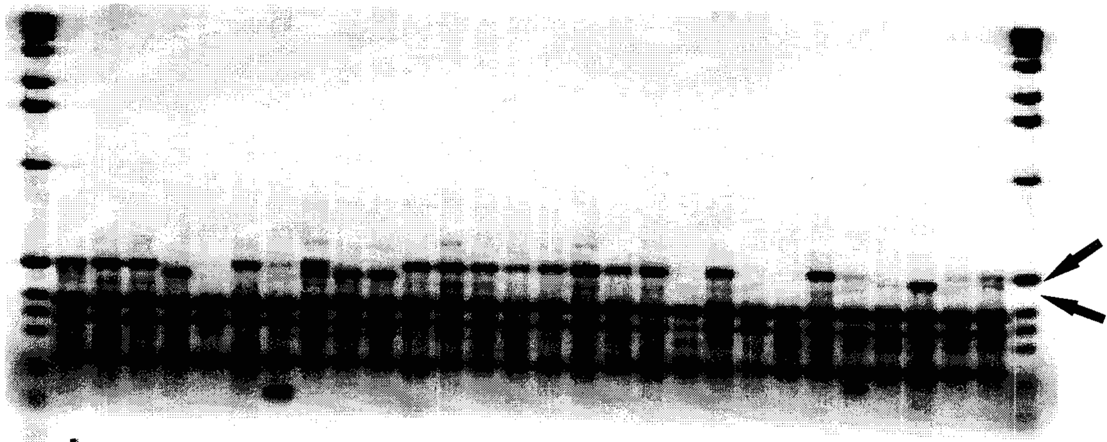

Figure 2
A representative RAPD gel. Bands were separated in a 2% agarose gel. DNA was amplified with Operon Technology Inc. primer J-10. Polymorphic bands are indicated by the arrows on the right and scored from the top as *10J516* and *10J500*.

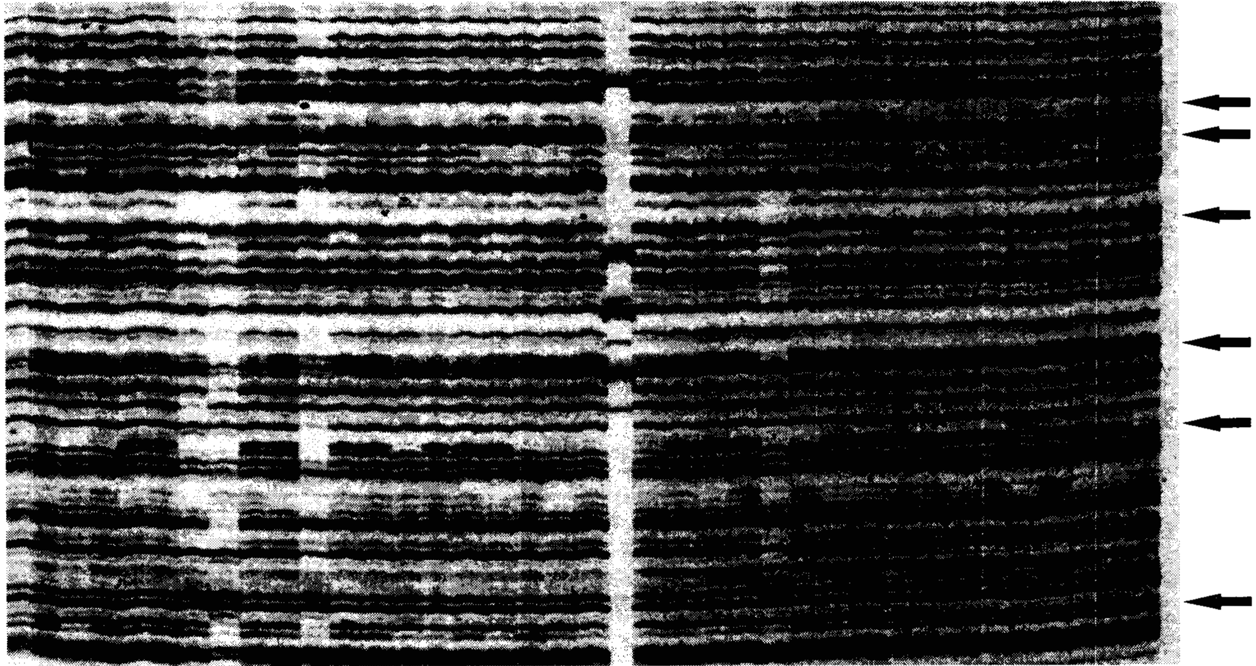

Figure 3
A representative AFLP gel. Bands were separated in 7% polyacrylamide. DNA was amplified using the primer pair AGG/CAA from the Perkin/Elmer AFLP selective amplification start up module. Polymorphic bands are indicated by the arrows on the right side. The bands were scored starting from the top as *AGG/CAA185, AGG/CAA180, AGG/CAA175, AGG/CAA140, AGG/CAA129*, and *AGG/CAA110.*

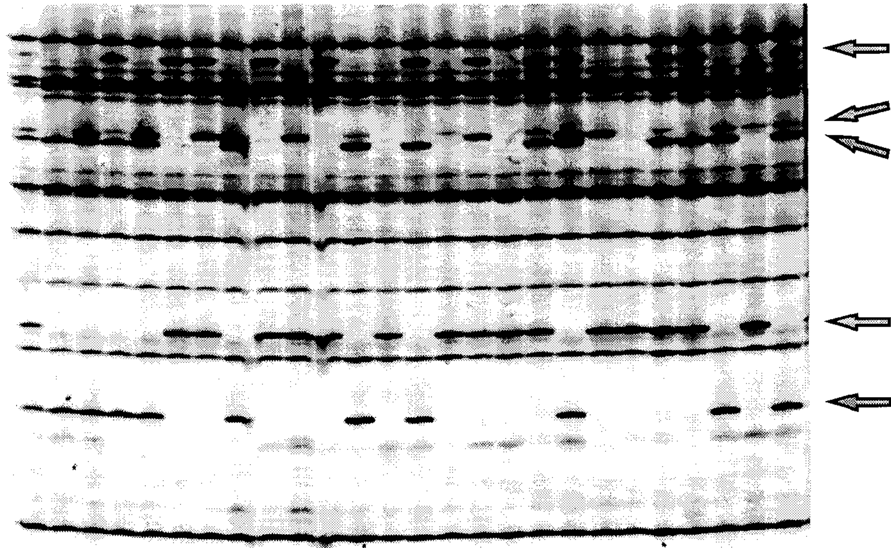

Figure 4

A representative PINE gel. DNA was amplified with the 5'Fok1 and 5'Tc1 primer. Bands were separated electrophoretically in a 4.5% polyacrylamide gel. Polymorphic bands are indicated by the arrows on the right side. Starting from the top of the gel the bands were scored as *5F5T308*, *5F5T281*, *5F5T278*, *5F5T232*, and *5F5T217*.

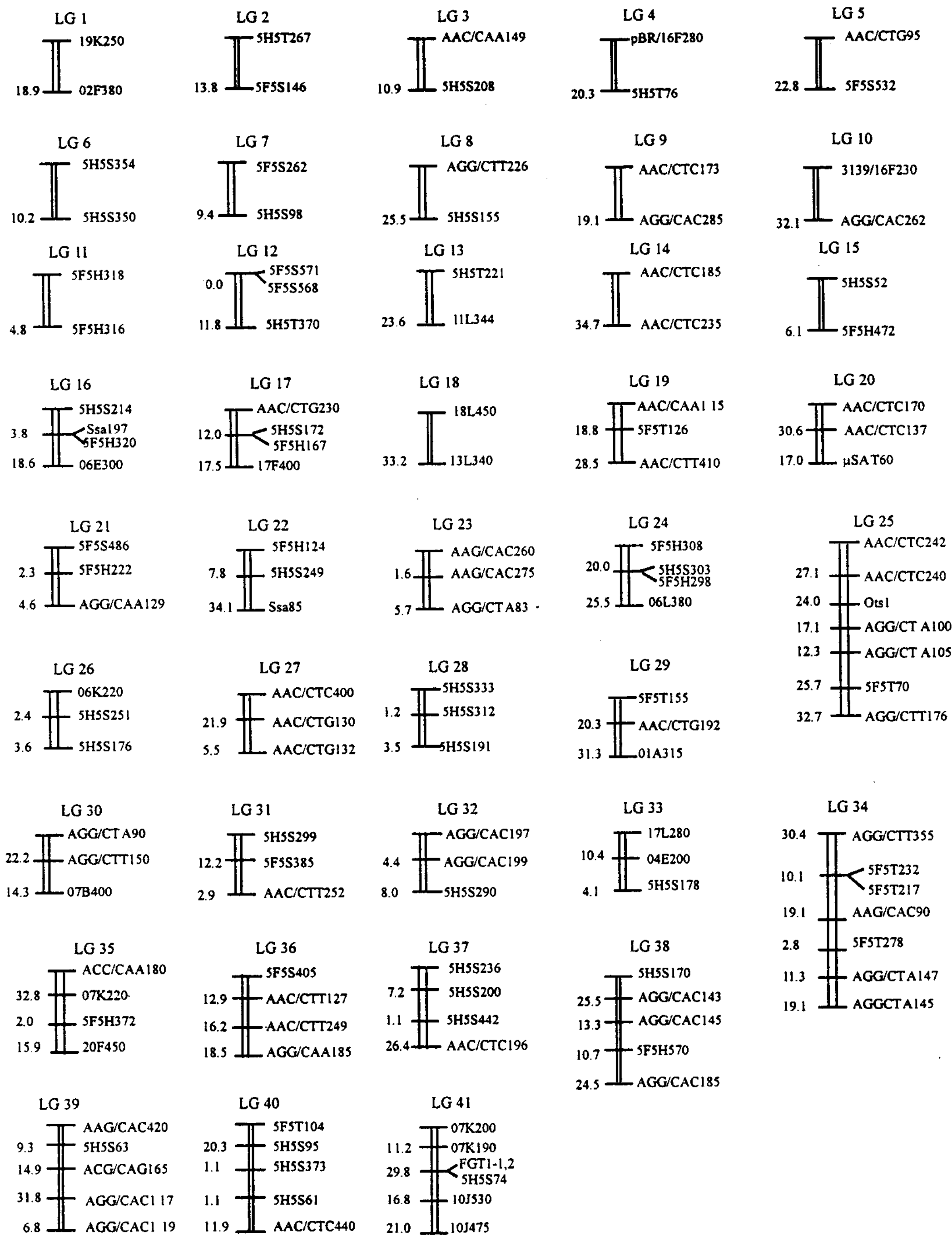

Figure 5

Genetic linkage map of pink salmon. Numbers to the left of the linkage groups indicate recombination rates (centimor gans, cM). To the right of the linkage groups are the locus names.