

Exxon Valdez Oil Spill
Restoration Project Annual Report

Construction of a Linkage Map for the Pink Salmon Genome

Restoration Project 97190
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
Kate R. Lindner
Kathy L. Knudsen
Douglas J. Reedy
Kristine L. Pilgrim

Division of Biological Sciences
University of Montana
Missoula, Montana 59812

for:

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

April 1998

Construction of a Linkage Map for the Pink Salmon Genome

Restoration Project 97190 Annual Report

Study History: This project began in FY 96, but the authorization to proceed was delayed until 5 March 1996. Five oral presentations reporting the results of this project have been given at professional meetings or university seminar series. Dr. Fred Allendorf gave an overview of the project at the national meeting of the American Fisheries Society in Dearborn, Michigan, August 1996 and at the Arctic Division of the American Association for the Advancement of Science annual meeting in Valdez, Alaska, September 1997. In addition, Dr. Allendorf presented an overview of this study at the Swedish Agricultural University, Umeå, Sweden in January 1998 and Aarhus University, Aarhus Denmark, in February 1998. Kate Lindner presented at the semi-annual Coast-wide Salmonid Genetics meeting in Seattle, March 1997. Dr. Paul Spruell attended the U.S. Department of Agriculture panel on mapping in aquaculture species in San Diego, January 1998. This panel may result in additional funding for salmonid mapping efforts. A journal article (Spruell, P., K.L. Pilgrim, B.A. Greene, C. Habicht, K.L. Knudsen, K.R. Lindner, G.K. Sage, J.E. Seeb, and F.W. Allendorf. Inheritance of nuclear DNA markers in haploid pink salmon embryos) has been submitted to the Journal of Heredity and is included in this report as Appendix 1. We will submit at least two additional manuscripts within the next year. The first of these manuscripts will describe gene-centromere mapping and the genomic distribution of PCR-amplified loci in pink salmon. It will be submitted to Genetics within the next six months. A complete description of the linkage map will be submitted to the Journal of Heredity within a year.

Abstract: We are constructing a detailed genomic linkage map for pink salmon (*Oncorhynchus gorbuscha*), that will assist recovery efforts for pink salmon in Prince William Sound. In August 1995, families of haploid and gynogenetic diploid pink salmon were produced using gametes collected at the Armin F. Koernig Hatchery in Prince William Sound. We selected one reference family and have analyzed the segregation pattern of 590 DNA fragments in the maternal parent and 94 of her haploid progeny. Of these fragments, 540 have been assigned to one of 56 linkage groups. We have mapped 266 loci in relation to their centromeres using gynogenetic diploid progeny. These data should allow us to reduce the number of linkage groups to 26, the number of chromosomes present in the female pink salmon genome, and will enable us to incorporate data from other salmonid maps onto our map. We remain ahead of schedule to complete the mapping portion of this project and plan to initiate the second portion of the study in August 1998, with the production of pink salmon families for release from the Alaska SeaLife Center.

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

Project Data: *Description of data* - Data are the inheritance of DNA fragments in the haploid and gynogenetic diploid progeny of a single pink salmon female (95-103). The haploid data set

consists of 590 polymorphic DNA fragments in female 95-103 and 94 of her haploid progeny. The diploid data consist of genotypes of 71 gynogenetic diploid progeny from female 95-103 at 266 loci. *Format* - Excel spreadsheets. *Custodian* - 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. *Availability* - Data will be made available to individuals within the reasonable bounds of sharing unpublished data.

Citation:

Allendorf, F.W., P. Spruell, K.R. Lindner, K.L. Knudsen, D.J. Reedy, and K.L. Pilgrim. 1998. Construction of a linkage map for the pink salmon genome, *Exxon Valdez Oil Spill Restoration Project Annual Report* (Restoration Project 97190), Alaska Department of Fish and Game, Habitat and Restoration Division, Anchorage, Alaska.

TABLE OF CONTENTS

LIST OF TABLES	5
LIST OF FIGURES	6
LIST OF APPENDIXES.....	7
EXECUTIVE SUMMARY	8
INTRODUCTION	10
OBJECTIVES	12
METHODS	13
RESULTS	15
DISCUSSION	17
CONCLUSIONS.....	20
ACKNOWLEDGMENTS	20
LITERATURE CITED	21
OTHER REFERENCES	24
Personal Communications	24

LIST OF TABLES

1. Summary of primer pairs screened and polymorphic loci detected by four different techniques. The second column is the number of polymorphic loci detected by each technique. The percentage of those markers assigned to a linkage group is given in the third column. The fourth column is the percent of the loci that are inherited in a co-dominant manner	25
2. Summary of linkage groups in the pink salmon genome map based on inheritance in 94 haploid progeny of a single female.....	26
3. Progeny genotypes at 22 loci in gynogenetic diploid pink salmon. 11 and 22 are homozygotes for alternate maternal alleles; 12 are heterozygotes.	27
4. Results of gene-centromere mapping of PINE loci in gynogenetic diploid family of pink salmon. The number of segregating loci resulting from amplification using each primer combination is given above the diagonal and in parentheses along the diagonal. The average proportion of heterozygotes (y) is given on and below the diagonal. Primer combinations that did not result in scoreable products are indicated by a 0 above the diagonal and a dash on or below the diagonal.	28
5. Comparison of linkage maps from four teleost fishes. Number of each type of marker included on the map are given. LOD and rf are the log odds and recombination fraction used for construction of the map using the MapMaker linkage analysis program (Lander et al. 1987).	29

LIST OF FIGURES

1. Genetic linkage map of pink salmon based on the inheritance of 540 polymorphic loci amplified using the polymerase chain reaction. Numbers to the left of the linkage groups indicate recombination rates (centimorgans, cM). Locus names are to the right of the linkage groups... 30
2. Distribution of distances between DNA markers and their centromeres. The frequency of second division segregation (γ) is given on the X-axis: (a) all markers, (b) microsatellites, (c) AFLPs, and (d) PINEs..... 33

LIST OF APPENDIXES

1. Spruell, P., K. L. Pilgrim, B.A. Greene, C. Habicht, K. L. Knudsen, K. R. Lindner, G. K. Sage, J.E. Seeb, and F. W. Allendorf. Inheritance of nuclear DNA markers in gynogenetic haploid pink salmon. Submitted to the Journal of Heredity.....34

EXECUTIVE SUMMARY

The construction of a genetic linkage map for the pink salmon (*Oncorhynchus gorbuscha*) genome is in its third 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 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 we are constructing will provide the platform to address the genetic impact of the oil spill. The initial framework of the map is being constructed using haploid progeny to avoid the difficulties associated with dominant markers that obscure recessive alternatives in diploids. Gynogenetic diploids from the same family are also being examined to locate the centromere of each chromosome 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. 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 (number 95-103) was chosen to be the reference family upon which our initial mapping efforts were based. Embryos were confirmed to be of gynogenetic origin using a sex-specific pseudogene locus and several microsatellites. After removal of diploid individuals, family 95-103 consists of the maternal parent, 94 of her haploid embryos, and 71 gynogenetic diploid progeny.

Linkage analysis of 590 markers segregating in the gynogenetic haploids produced a genetic map comprised of 56 linkage groups, covering a distance of 4526 centimorgans (cM). Assuming a minimum distance of 25 cM for linkage detection and accounting for all the gaps and unlinked markers the minimum distance of the pink salmon genome is 6744 cM. The haploid pink salmon genome consists of approximately 2.72 billion base pairs or 2.72×10^6 kilobase pairs (kbp) thus, we estimate a physical recombination rate of approximately 406 kbp/cM.

We have generated a large number of markers throughout the genome using haploid embryos and multilocus techniques. Our next goal is to “consolidate” the map and reduce the number of linkage groups to 26, the number of chromosomes in pink salmon. In order to achieve this goal, we have collected gene-centromere distances of 266 loci using gynogenetic diploids. These data will allow us to identify the centromere of each chromosome and assign the linkage groups identified using the haploids to a specific chromosome.

We are also adding additional microsatellites and genes of known function to the linkage map. The microsatellites will serve as land marks, or “anchor loci” that facilitate comparisons between maps. These loci will allow us to compare genetic linkage of odd and even year pink salmon, estimate recombination rates of males and females, and incorporate data from other salmonid linkage maps. The known genes will also serve as anchor loci and will be of great interest during the next phase of this project in which we examine the effects of the marine environment on the pink salmon genome.

We remain ahead of schedule to complete the mapping portion of this project and are prepared to begin the second phase. We will produce families of pink salmon in August 1998 for release from the Alaska SeaLife Center the following spring.

INTRODUCTION

The construction of a genetic linkage map for the pink salmon (*Oncorhynchus gorbuscha*) genome is in its third year. This map will provide a more thorough understanding of the genetic impacts of the March 1989 *Exxon Valdez* oil spill on pink salmon populations. In addition, the map will increase our understanding of the genetic characteristics of pink salmon 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 (Bue et al. 1998). 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 of pink salmon will be invaluable for documenting the effects the spill may have had on pink salmon. Such a 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 must generate reproducible markers and it is preferred that they allow the detection of multiple loci from a single polymerase chain reaction (PCR). These multilocus techniques generate the large number of loci necessary to serve as the framework of the map. The fragments detected by each technique should be dispersed throughout the genome to avoid exclusion of specific chromosomal regions. Finally, the markers must be present in the maternal parent and segregate in a ratio expected in simple Mendelian inheritance models.

Many of the multilocus 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.

Analysis of the segregation of dominant markers in haploids is an efficient method to generate the large number of markers necessary to produce a linkage map. However, this system does have some limitations. We have generated a large number of reproducible markers but have not yet been able to consolidate the map, or reduce the number of linkage groups to 26, the number

of chromosomes present in the female pink salmon genome. In addition, dominant markers, such as amplified fragment length polymorphisms (AFLPs) and paired interspersed nuclear element PCR (PINEs) do not facilitate comparison between maps. These limitations can be addressed using additional crosses and anchor loci, markers generated in single locus PCR such as microsatellites and known genes.

Consolidation of the map is our next goal. Identification of the centromeres of each chromosome would assist in this consolidation. The analysis of gynogenetic diploids is an effective method to identify the centromeres of chromosomes (Thorgaard 1983) and will aid in the consolidation of the map in two ways. First, gynogenetic diploids allow the analysis of recombination rates of individual markers based on the frequency of heterozygotes. These recombination rates are a function of the distance the marker is located from the centromere, the gene-centromere distance. Once markers tightly linked to centromeres have been identified, additional markers linked to the centromeric markers can be assigned to a specific chromosome (Johnson et al. 1996). Second, the analysis of co-segregation between markers used in gene-centromere mapping will identify centromeric groups that are linked thereby consolidating the map.

Comparison between the pink salmon linkage map and other teleost linkage maps are difficult due to the lack of shared anchor loci. Linkage relationships in fishes persist after 300 million years of evolution (Graf 1989), suggesting that many anchor loci should occupy a similar chromosomal position in closely related taxa. As linkage maps of fish become more common it is important to be able to incorporate information from other linkage maps onto the pink salmon linkage map. Therefore, it is necessary to include more anchor loci on the pink salmon linkage map.

We currently are adding numerous microsatellites and know genes to the map that will serve as anchor loci. PCR primers for many salmonid microsatellite loci are published and many of these should amplify homologous loci in pink salmon. Given the high level of polymorphism typical for these markers, screening microsatellite loci is probably the most efficient method to add anchor loci to the map.

Genes of known function can also serve as anchor loci and are particularly interesting in the study of interactions between genes and the environment. For example, the major histocompatibility complex, MHC, has been extensively studied in vertebrates for a variety of reasons including its role in disease resistance or susceptibility. Various class I and class II MHC alleles have been characterized in Pacific and Atlantic salmon, including pink salmon (Katagiri et al. 1996, Miller et al. 1996, Miller et al. 1997). Katagiri et al. (1996) sequenced the entire MHC class I cDNA in pink salmon and described the presence of two main allele types in the alpha 1 domain. The presence or absence of a six base pair insertion differentiates these allele classes. Female 95-103 does not show a length polymorphism at this locus. However, other techniques to detect sequence polymorphisms should allow us to add this and other genes to the linkage map.

Another benefit to generating a linkage map is the ability to identify the sex-determining region in pink salmon. Sex-linked genetic markers are important for the study of population genetics. Several such markers have been described and developed for a variety of species (i.e., mouse, human, chicken and fruit fly) yet few are known in salmonids (Devlin et al. 1991, Du et al. 1993, Allendorf et al. 1994, Forbes et al. 1994, Moran et al. 1996). Sex-linked markers are useful for the gender identification of immature fish, for investigating the genealogy and phylogeny of species, for comparative gene mapping, and for detecting geographic population structures. Sex-linked markers can also help compare rates of recombination between the X and Y chromosomes in males, the X chromosomes in females, and aid in the determination of the size of the non-recombining region on the Y-chromosome.

In this report, we update our progress on the construction of the pink salmon linkage map that is being generated by the analysis of segregation in haploid gynogens. We also report gene centromere distances for 266 loci included on the haploid linkage map. Concordance of results for loci analyzed in both haploid and diploid progeny will confirm the accuracy of the data being used to construct the map. Finally, we compare the genomic distribution of loci amplified by AFLPs and PINES. These data will provide a framework to address questions of pink salmon survival and fitness.

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.

The genomic map research was originally 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 be invaluable for interpreting the results of studies such as Restoration Study \191A in several ways. First, it will be possible by following the inheritance of any DNA lesions 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. However, Restoration Study \191A is no longer ongoing, and thus our work concentrates on objectives 1 - 2 and 5 - 6.

METHODS

Samples and Gynogenesis

Gynogenetic Haploid Production

In August 1995, gametes and tissues of 31 pink salmon were collected from the Armin F. Koernig hatchery, Prince William Sound, Alaska. This hatchery stock originated from adult fish collected at several spawning sites in Prince William Sound. Gynogenetic haploid embryos were produced by sperm inactivation as described by Thorgaard et al. (1983). Sperm from four males was pooled prior to UV irradiation, then mixed with eggs from individual females. Females were numbered and their progeny were designated by year class and the number assigned to that female (e.g. family 95-103). Embryos from these families were incubated until just prior to hatching when they were collected and preserved in ethanol. DNA extraction and confirmation of gynogenetic origin was completed on the haploids as described in Spruell et al. (submitted, Appendix 1).

Gynogenetic Diploid Production

Gynogenetic diploid progeny were produced using gametes from the same parents used to produce the haploid progeny. Eggs were fertilized with sperm that had been UV irradiated. Diploidy was restored by a heat shock that causes the retention of the second polar body (Thorgaard et al. 1983). These diploid embryos are viable and were raised until they reached sufficient size (approximately 40 mm, total length) for allozyme analysis. Individuals were sacrificed and tissue samples collected and frozen at -80C for allozyme analysis. In addition, the caudal peduncle and fin were collected for DNA extraction. DNA extraction was completed as described in Spruell et al. (submitted, Appendix 1).

Individuals were confirmed to be gynogenetic diploid progeny of female 95-103 using three types of genetic markers. Allozyme analysis indicated that four individuals contained alleles that were not present in female 95-103 (J. Seeb, pers. comm.). These four individuals plus two additional individuals were confirmed to be males using a Y-linked pseudogene as described in Spruell et al. (submitted, Appendix 1). We screened eight microsatellite loci and detected no additional individuals that contained alleles that could not be attributed to female 95-103. The

six individuals that were not gynogenetic diploids may have resulted from incomplete inactivation of sperm (Thorgaard et al. 1983) or may have been individuals from another treatment. These six individuals were removed from subsequent analyses.

Selection of a Reference Family

A single family (95-103) was chosen as the reference family to generate the linkage map. DNA was extracted from 150 putative haploid embryos in this family. Based on microsatellite and Y-linked pseudogene data we detected 29 diploid individuals and eliminated them from subsequent analysis. Ninety-four of the remaining 131 contained sufficient DNA for extensive analysis. Gene-centromere mapping studies were conducted on a total of 71 individuals from the gynogenetic diploid treatment.

Single Pair Matings

Diploid embryos from a single pair mating (96-A14; female 96-18 x male 96-30) were collected to screen for sex-linked markers using bulked segregant analysis (Giovanoni et al. 1991, Michelmore et al. 1991). These individuals were sexed with a growth hormone pseudogene as described in Spruell et al. (submitted, Appendix 1), and DNA was quantified using a scanning spectrofluorometer. Offspring from this family were used to produce sex-specific mixtures of DNA for bulked segregant analysis. Sex-specific mixtures were created by combining 0.5 μ g of DNA from each of 10 males and 10 females. The two gender specific DNA mixtures were screened and visually examined for sex-specific amplification products.

Markers

Genetic markers are included on the haploid linkage 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. Five hundred and ninety fragments met these criteria and were included in the analysis of joint segregation ratios. RAPDs, AFLPs, and PINEs were amplified as described in Spruell et al. (submitted, Appendix 1).

Microsatellite loci were amplified as reported by the original authors with minor modifications. Primers and annealing temperatures are as follows: *FGT1*, 51°C (Sakamoto et al. 1994); *μ SAT60*, 55°C (Estoup et al. 1993); *OTS1*, 55°C (Hedgecock, personal communication); *OTS100*, *OTS101*, 55°C (Nelson et al. submitted) *SSA85* and *SSA197*, 57°C (O'Reilly et al. 1996); *OC12*, 56°C (Condrey et al. in press); *OGO1C*, *OGO7*, 60°C; *OGO5*, 64°C; *OGO8*, 56°C (Olsen pers com.). PCR products were electrophoresed in a 7% denaturing polyacrylamide gel and visualized with a Hitachi FMBIO-100 fluorescent imager.

Primers developed by Katagiri et al. (1996; 5' TGA CTC ACT CCC TGA AGT A 3' and 5' CTC CAC TTT GGT TAA AAC G 3') were used to amplify the MHC class 1 gene in females

95-103, 95-104, 95-114, and 95-120. PCR conditions followed those published by Katagiri et al. (1996).

Genotypes of the 71 individuals confirmed to be gynogenetic diploid offspring of female 95-103 were scored at 157 AFLP loci, 101 PINE loci, and eight microsatellite loci. The techniques for each marker type are identical to those used for amplification of markers in the haploid embryos. (Spruell et al. submitted, Appendix 1).

Map Construction

We used MapMaker software (Lander et al. 1987) to assign markers segregating in haploid progeny to linkage groups. Grouping of markers was completed using a minimum LOD score of 3.0 and a maximum recombination fraction (θ) of 0.25 ($P < 0.001$). The linkage phase of the markers we are using is unknown. Therefore, the segregation pattern of each locus is entered into the program in both possible phases. This resulted in a data set that was too large to use the MapMaker "group" command on the entire set. The markers present in each group were assigned to chromosomes and all additional markers were analyzed using the "assign" command. The Kosambi mapping function was used to calculate the genetic distance.

RESULTS

Haploid Linkage Map

Linkage analysis using 590 markers (Table 1) produced a genetic map comprised of 56 linkage groups and covering a distance of 4526 centimorgans (cM) (Table 2, Fig. 1). Seventeen groups include five markers or less. The largest group contains 41 markers. Fifty markers remain unlinked. Assuming a minimum distance of 25 cM for linkage detection and accounting for all the gaps and unlinked markers, the minimum size of the pink salmon genome is 6744 cM. The haploid pink salmon genome consists of approximately 2.72 billion base pairs or 2.72×10^6 kilobase pairs (kbp; Johnson et al. 1987) thus, we estimate a physical recombination rate of approximately 406 kbp/cM

Single Locus PCR

Twelve of 36 microsatellite primers successfully amplified in female 95-103 segregate in the haploid family 95-103. The loci *OGO7* and *FGT1* are duplicated but could be included on the map because only one locus of each of the duplicated pairs segregates in the haploid progeny.

In order to increase the number of microsatellites we can add to the map, we have initiated collaborations with the laboratory of Drs. Roy Danzmann and Moira Ferguson at the University of Ontario in Guelph, and the laboratory of Dr. Gary Thorgaard at Washington State University. Kate Lindner recently spent a week working with Dr. Takashi Sakamoto, a post-doctoral fellow at the University of Ontario, to screen 108 microsatellite primers in the female pink salmon upon

which our map is based. This collaboration resulted in 28 microsatellites that are polymorphic in female 95-103 and are currently being placed on the map.

We are in the process of placing known genes such as MHC on the pink salmon linkage map. Amplification of the alpha 1 domain of MHC I in four pink salmon females produced both type A1 and A2 alleles. Female 95-103 was homozygous for type A1 alleles. However, female 95-114 possesses both A1 and A2 alleles and they segregate in her haploid progeny.

Sex-Linkage Group

Bulked segregant analysis uses two DNA mixtures that differ only in the region of interest to identify genetic markers linked to a trait. PCR products from two gender-specific pools of DNA were analyzed for sex-specific amplification products. Since the males and females included in the pooled DNA were siblings, these two pools were genetically similar except for the sex-determining region in males that does not undergo recombination with the X chromosome. Only those loci that are sex-linked and segregating in this cross could be detected. However, more than 3000 fragments were screened using 74 AFLP primer pairs (with an average of 34 fragments per primer pair), and 33 PINE primer pairs (with an average of 22 fragments per primer pair) and no sex-linked markers were identified.

Gene-Centromere Mapping

Segregation of Markers

Gene-centromere distances (Fig. 2) were estimated by the proportion of gynogenetic diploid progeny that were heterozygous (y). Since gynogenetic progeny are the result of second polar body retention, heterozygotes can only be produced if there is a crossover between the centromere and that marker (Thorgaard et al. 1983). These recombination rates are a function of the distance the marker is located from the centromere, the gene-centromere distance. Once markers that are tightly linked to centromeres have been identified, additional markers linked to the centromeric markers can be assigned to a specific chromosome (Johnson et al. 1996). The analysis of markers in the gynogenetic diploids will also provide an independent measure of linkage. This will allow us to fill the gaps between linkage groups and thereby reduce the total number of linkage groups, as well as confirm linkages identified with the haploids.

Table 3 shows the proportion of homozygotes and heterozygotes at 22 codominant loci. Seven of these are microsatellite loci, and 11 are either AFLP or PINE loci. Four of these loci were identified as fragments produced by two different primer pairs that exhibited perfect alternative segregation in the haploids. Some loci mapped very close to their centromere (e.g., *5F-745*, $y=0$), and others are far from their centromere or distal (e.g., *OGO8*, $y=1.00$). There is no evidence for differential survival of the two homozygous types at any of the 22 loci. In addition, the similar proportion of the two homozygous classes provides independent support that the four pairs of fragments that showed perfect alternate segregation in the haploids map to the same locus.

Heterozygotes cannot be differentiated from the dominant homozygotes at dominant markers. To estimate the proportion of heterozygotes at these markers, we assumed equal numbers of each homozygote class. The frequency of second division segregation (y) can then be estimated by

$$y = \frac{(N_i - 2N_{aa})}{N_i}$$

where N_i is the total number of progeny screened and N_{aa} is the observed number of recessive homozygotes.

Distribution of Markers

The distribution of y values for all loci (Fig. 2a) confirms that the markers we are scoring are distributed along the length of the chromosome. However, examination of each marker type indicates that there are differences between their distributions. AFLPs and PINEs show statistically different distributions for y (contingency chi-square, $P < 0.001$). AFLPs are underrepresented on the tips of chromosomes (Fig. 2c). The distribution of PINEs is biased toward centromeric and telomeric regions (Fig. 2d). We currently have too few microsatellites mapped to provide the statistical power necessary to test their distribution (Fig. 2b).

We tested each PINE primer to investigate the possibility that a single element or a subset of elements were responsible for the non-random distribution of PINE fragments. Although there is a tendency for fragments produced using the HpaI 3' primer to be more telomeric, there are no other obvious relationships between the primers used and the distribution of y values for PINEs (Table 4).

DISCUSSION

Haploid Linkage Map

In comparison to previously described fish linkage maps, the size of the pink salmon map is large, 6744 cM (Table 5). Due to the polyploid ancestry of salmonids we expect our map to be larger than linkage maps of non-salmonid species, such as zebrafish (*Danio rerio*) and medaka (*Oryzias latipes*). There is also a reduced rate of recombination in males relative to females (May and Johnson 1990). Therefore, we also expect the female pink salmon map to be larger than the male rainbow trout (*Oncorhynchus mykiss*) map (Young et al. 1998).

Gene-Centromere Mapping

Segregation of Codominant Markers

We observed an equal number of homozygotes in each class at 22 codominant markers. All of these markers met our criteria for inclusion on the haploid linkage map (e.g. 1:1 segregation) and therefore it is unlikely any are linked to a recessive lethal. However, unlike haploids, gynogenetic diploid development proceeds through hatching. Therefore, the expression of a linked lethal allele at or after hatching would effect the survival of gynogenetic diploids but not the haploids. Such an allele could result in an overestimate of the gene-centromere distance for dominant markers by decreasing the number of homozygous recessive gynogenetic diploids. Our data provide evidence that linkage between an allele of a mapped locus and a recessive lethal is not a common occurrence. These data are consistent with Allendorf et al. (1986) who used a similar approach to test for linkage between recessive lethal alleles and allozyme loci.

Distribution of Markers

We are using the frequency of recombination to estimate the distance between a marker and the centromere. However, the correspondence between genetic distance and physical distance is problematic for distal markers. All markers that are more than 50 cM away from the centromere will have a y equal to 1.0 causing an apparent accumulation of markers in telomeric regions, even if the markers are uniformly distributed throughout the genome.

Our data are concordant with previous attempts to examine the distribution of these markers in salmonids. In the few cases where there are apparent discrepancies, experimental design may account for the observed differences. Young et al. (1998) inferred that AFLPs were centromeric in rainbow trout based on the presence of a cluster of tightly linked AFLPs at the center of most of their linkage groups. We observed a reduction of AFLPs at the extreme tips of chromosomes but have not observed the tight clustering found in the rainbow trout map. Young et al. (1998) suggested that the reduced rate of recombination in males relative to females may cause the apparent centromeric clustering of AFLPs. If so, the genetic distance calculated using recombination in males would be an underestimate of the physical distance separating those markers. Although there is a reduction in the number of AFLPs with y values approaching one, it is not clear that reduced male recombination alone would result in the differences observed in the two maps.

Another possibility is that the difference in AFLP distribution between the two maps reflects a bias in the base composition of certain genomic regions. The majority of AFLP polymorphisms result from point mutations at one of six selective bases added to the synthetic primers used to amplify the AFLP products. Although both linkage maps were produced using the same restriction enzymes, Young et al. (1998) included different preselective bases on both of the primers used to generate AFLPs. If centromeric regions are biased toward these bases the accumulation of AFLPs near the centromeres may result.

The distribution of the PINEs is somewhat more complex. We observe a lack of PINE fragments in the middle region of chromosomal arms. PINEs are concentrated at both the centromere, as described by Greene and Seeb (1997), and at the telomeres. Greene and Seeb (1997) report the centromeric distribution of fragments amplified using primers homologous to *Sma*I and *Tc*I. Our data does not indicate a tendency for either of these elements to be centromeric (Table 4). The experimental design of Greene and Seeb (1997) was such that markers heterozygous but located far from the centromere would not be detected. We observe a similar pattern as described by Greene and Seeb (1997) if the markers with a y greater than 0.7 are eliminated from our data.

Young et al. (1998) used a subset of SINE-homologous sequences as probes in Southern blots to detect variation in rainbow trout and report a uniform distribution of the resulting fragments. The comparison of our data with that of Young et al (1998) is not direct because of the differences between the methods used to detect polymorphisms. However, the apparent telomeric accumulation of PINEs in the pink salmon data may be the result of many markers being at a physical distance greater than can be detected using gene-centromere mapping. Young et al. (1998) inferred the position of the centromere from the accumulation of AFLPs but did not conduct gene-centromere mapping. Therefore, their data are derived solely from the co-segregation of linked markers and are not bounded by y values that underestimate physical distance.

Y-linked markers

We are continuing our efforts to isolate markers linked to the Y chromosome. We have examined a substantial number of loci using bulked segregant analysis but have not isolated male-specific markers using that technique. The size of the male-specific region in salmonids is unknown. However, the limited differentiation between the X and Y chromosomes suggests that the region may be quite small. If this is true, an extremely large number of random polymorphic markers would have to be screened to detect a marker linked to the sex-determining region. More direct methods to isolate markers linked to the sex-determining region are being explored.

Anchor Loci

Techniques such as AFLPs and PINEs are valuable in that they rapidly produce many markers that can serve as the framework for a linkage map. However, these techniques produce markers with unknown genomic characteristics. Anchor loci that enable comparisons among closely related taxa should be included to maximize the benefit of linkage maps. We are focusing on adding additional microsatellite loci and known genes to the map to serve as anchor loci.

Although known genes are of particular interest for mapping, polymorphism at these markers is frequently at the level of the DNA sequence. For example, our initial attempts to map MHC were hampered by the lack of length polymorphism in female 95-103. We are investigating techniques such as heteroduplex analysis, single strand conformational polymorphism, and

denaturing gradient gel electrophoresis, to address this issue. These techniques allow the detection of sequence polymorphisms without actually sequencing the product. In addition, once the map has been consolidated, it will be much easier to add markers to the map based on the analysis of other females that are polymorphic for markers of particular interest.

Microsatellites that amplify in closely related taxa can also serve as anchor loci. Several research groups are working to develop and map salmonid microsatellites. We plan to collaborate with those groups to maximize the number of microsatellites we are able to place on the pink salmon linkage map. The visit to the Guelph lab and subsequent addition of numerous microsatellite loci is an example of the value of such collaboration. These cooperative efforts also facilitate the incorporation of the data generated by other research groups onto the pink salmon map and encourage comparisons between maps. These comparisons should provide an increased understanding of the salmonid genome and residual tetrasomy.

CONCLUSIONS

We have made substantial progress toward the completion of a genetic linkage map in pink salmon. The combination of data from the haploid linkage map and the gene-centromere studies should allow us to consolidate the map and incorporate the data available from other salmonid maps more readily. We are moving toward the consolidation of the map and are adding anchor loci onto the framework produced using AFLPs and PINEs. In August 1998, we anticipate producing families of pink salmon for release from the Alaska Sea Life Center. The release of these families will mark the beginning of our application of the map to the assessment of the genomic response to environmental selective forces.

ACKNOWLEDGMENTS

We thank James E. Seeb and Chris Habicht from the (ADFG) for providing samples and producing the gynogenetic progeny for this project. Thank you James E. Seeb for all of your helpful comments. In addition we would like to thank Jeff Olsen (University of Washington) and Dennis Hedgecock (University of California, Davis) for sharing unpublished microsatellite primer sequences. We also thank Takashi Sakamoto, Roy Danzmann, and Moira Ferguson for sharing unpublished microsatellite primers and their hospitality.

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.
- 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.
- 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:in press.
- Condrey, M. J. and P. Bentzen. 1998. Characterization of coastal cutthroat trout (*Oncorhynchus clarki clarki*) microsatellites, and their conservation in other salmonids. *Molecular Ecology* 7:in press.
- 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.
- Du, S. J., R. H. Devlin and C. H. 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-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.
- Giovannoni, J. J., R. A. Wing, M. W. Ganai, and S. D. Tanksley. 1991. Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucl. Acids Res.* 16:6553-6558.
- Graf, J. D. 1989. Genetic mapping in *Xenopus laevis*: Eight linkage groups established. *Genetics* 123: 389-398.

- 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*). *Mol. Mar. Biol. Biotech.* 6: 328-338.
- 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.
- 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.
- Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly et al. 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.
- May, B., K. R. Johnson, and J. E. Wright. 1989. Sex linkage in salmonids: Evidence from a hybridized genome of brook trout and Arctic charr. *Biochemical Genetics* 27: 291-301.
- 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. *Proc. Natl. Acad. Sci. USA* 88: 9828-9832.
- Miller, K. M., Withler, R. E. 1996. Sequence analysis of a polymorphic *Mhc* class II gene in Pacific salmon. *Immunogenetics* 43: 337-351.
- Miller, K. M., Withler, R. E. 1997. *Mhc* diversity in Pacific salmon: Population structure and trans-species allelism. *Hereditas* 127: 83-95.
- Moran, P., J. L. Martinez, E. Garcia-Vazquez, and A. M. Pendas. 1996. Sex chromosome linkage of 5S rDNA in rainbow trout (*Onchorhynchus mykiss*). *Cytogenet. Cell Genet.* 75: 145-150.
- Nelson, R. J. and T. D. Beacham. (submitted). Isolation and Characterisation of microsatellite loci useful for study of Pacific Salmon.

- 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.
- Spruell, P., B. A. Greene, C. Habicht, K.L. Knudsen, K. R. Lindner, K. L. Pilgrim, G. K. Sage, J. E. Seeb, and F. W. Allendorf. (submitted) Inheritance of nuclear DNA markers in haploid pink salmon embryos. *Journal of Heredity*.
- Stanley, J. G. 1983. Gene expression in haploid embryos of Atlantic salmon. *J. Hered.* 74:19-22.
- Thorgaard, G. H., F. W. Allendorf, and K. L. Knudsen. 1983. Gene-centromere mapping in rainbow trout: High interference over long map distances. *Genetics* 103: 771-783.
- 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.

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.

Jeffrey B. Olsen, Marine Molecular Biotechnology Laboratory, University of Washington, 3707 Brooklyn Ave. NE, Seattle, WA 98105-6715.

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

Table 1. Summary of primer pairs screened and polymorphic loci detected by four different techniques. The second column is the number of polymorphic loci detected by each technique. The percentage of those markers assigned to a linkage group is given in the third column. The fourth column is the percent of the loci that are inherited in a co-dominant manner.

Marker Type	Number of Polymorphic Loci	Percent Assigned to a Linkage Group	Percent of Markers Codominant
AFLP	387	91	5
PINE	157	96	13
RAPD	34	74	0
Microsatellites	12	100	92 ^α
Total	590	92	9

^α *SSA197* could not be scored as a codominant marker due the presence of a null allele in female 95-103 (Spruell et al. submitted, Appendix 1).

Table 2. Summary of linkage groups in the pink salmon genome map based on inheritance in 94 haploid progeny of a single female (95-103).

Number of Markers in Group	Number of Linkage Groups	Average Size (cM)
1 - 5	17	31
6 - 10	18	48
11 - 15	12	110
16 - 20	5	156
21 - 25	2	183
26 - 30	1	243
50 - 60	1	418

Table 3. Progeny genotypes at 22 loci in gynogenetic diploid pink salmon. *11* and *22* are homozygotes for an alternate maternal allele, *12* are heterozygotes. The χ^2 values test for equal number of each homozygous class, 1 df. Different primer pairs refer to combinations of primer pairs that amplify alternately segregating alleles.

Loci	Progeny Genotypes			Proportion heterozygotes (<i>y</i>)	χ^2
	<i>11</i>	<i>12</i>	<i>22</i>		
Microsatellites					
<i>OTS-1</i>	34	13	22	0.19	2.57
<i>OGO-1</i>	24	17	27	0.25	0.18
<i>STRμ60</i>	16	46	9	0.65	1.96
<i>SSA-85</i>	22	18	31	0.25	1.53
<i>OTS-101</i>	26	21	23	0.30	0.18
<i>OC-12</i>	9	47	13	0.68	0.73
<i>OGO-8</i>	0	71	0	1.00	--
AFLPs					
<i>AGA/CTC-204</i>	28	13	26	0.19	0.07
<i>ACC/CAC-328</i>	16	40	10	0.61	1.38
<i>AAT/CTG-293</i>	15	30	20	0.46	0.71
<i>AAC/CGT-157</i>	21	23	18	0.37	0.23
PINEs					
<i>33.6/5T-384</i>	2	66	2	0.94	--
<i>5F-745</i>	33	0	37	0.00	0.23
<i>5H/5T-76</i>	27	19	22	0.28	0.51
<i>5H/5T-125</i>	0	68	0	1.00	--
<i>5H/5T-224</i>	22	22	24	0.32	0.09
<i>5F/5T-217</i>	0	70	0	1.00	--
<i>3H/5T-182</i>	27	15	29	0.21	0.07
Different Primer Pairs					
<i>5H/5S-191:</i>					
<i>ACG/CTG-322</i>	3	65	1	0.94	1.00
<i>5F/5T-308:</i>					
<i>3H/5H-318</i>	0	69	1	0.99	1.00
<i>5H/5T-77:</i>					
<i>3H/5T-182</i>	28	13	27	0.19	0.02
<i>5H/5T-76:</i>					
<i>3H/5T-183</i>	22	19	27	0.28	0.51

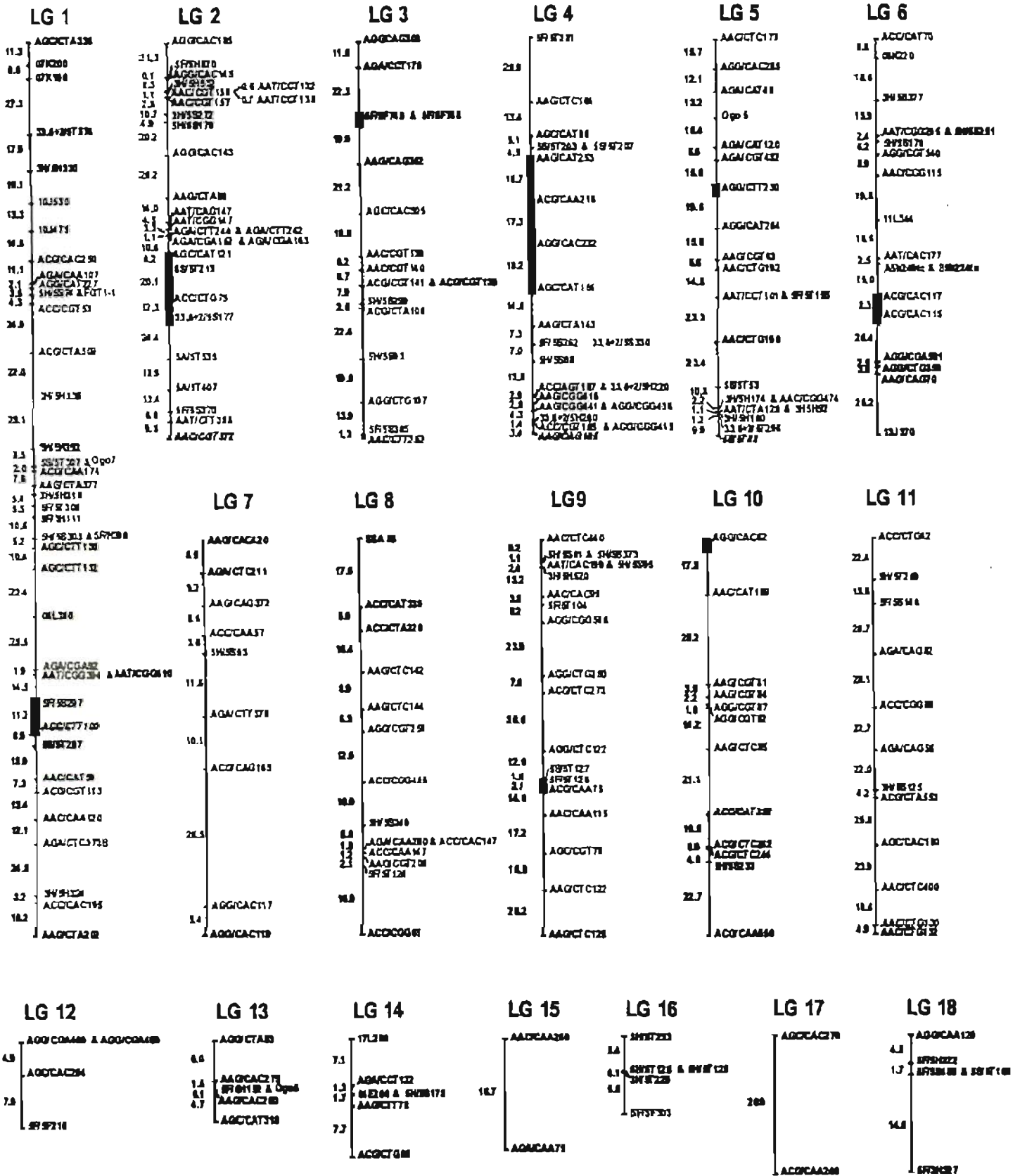
Table 4. Results of gene-centromere mapping of PINE loci in a gynogenetic diploid family of pink salmon. The number of polymorphic loci resulting from amplification using each primer combination is given above the diagonal and in parentheses along the diagonal. The average proportion of heterozygotes (γ) is given on and below the diagonal. Primer combinations that were not used in the map because they did not result in scoreable products in either haploids or diploid progeny are indicated by a 0 above the diagonal and -- on or below the diagonal. One primer pair, indicated by NA, has not yet been analyzed.

Source of primer sequence	Source of primer sequence					
	Tcl	FokI 5'	HpaI 5'	HpaI 3'	SmaI 5'	33.6+2
Tcl	--	10	9	5	16	6
FokI 5'	0.65	0.46 (4)	NA	4	5	2
HpaI 5'	0.61	NA	--	9	0	3
HpaI 3'	0.58	0.89	0.95	--	8	0
SmaI 5'	0.51	0.59	0.63	1.00	--	4
33.6+2	0.67	0.86	0.49	--	0.63	--

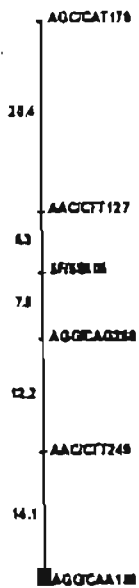
Table 5. Comparison of linkage maps from four teleost fishes. Number of each type of marker included on the map are given. LOD and *rf*, the log odds and recombination fraction respectively, are criteria used for construction of the map using MapMaker linkage analysis program.

	Medaka Wada et al. 1995	Zebrafish Johnson et al. 1996	Rainbow Trout Young et al. 1998	Pink Salmon
Number of Markers	170	652	476	590
Number of Linkage Groups	28	25	42	56
Number of Chromosomes	24	25	30	26
Estimated size (cM)	2480	2790	2627	6744
kbp/cM	323	590	913	406
AFLPs	0	0	332	387
PINEs	0	0	40	157
RAPDs	161	620	5	34
Microsatellites	0	27	2	12
Sequence tagged sites	0	5	0	0
Minisatellites	0	0	96	0
Recombination fraction (<i>rf</i>)	0.40	0.26	0.35	0.25
Log Odds (LOD)	3	3	3	3

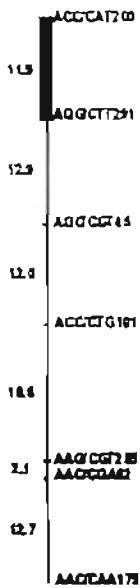
Figure 1. Genetic linkage map of pink salmon based on the inheritance of 540 polymorphic loci. Numbers to the left indicate recombination rates (cM). Locus names are to the right. Centromeres are indicated by black rectangles.



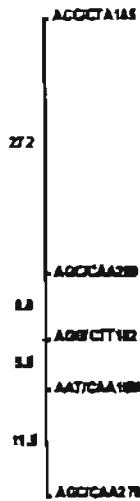
LG19



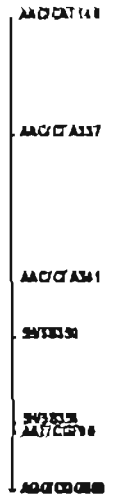
LG 20



LG 21



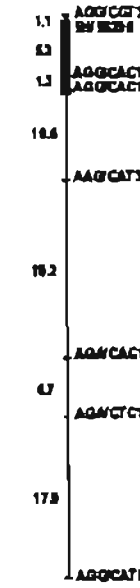
LG 22



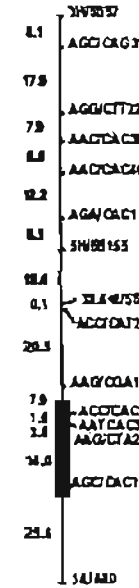
LG 23



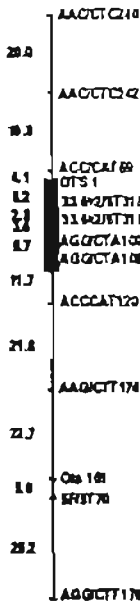
LG 24



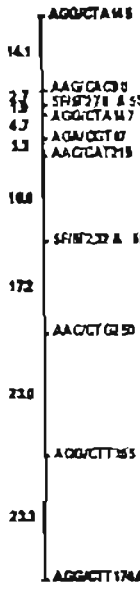
LG 25



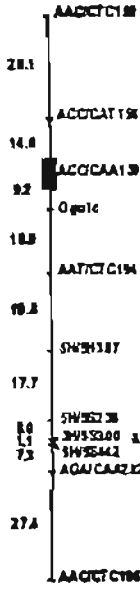
LG 26



LG 27



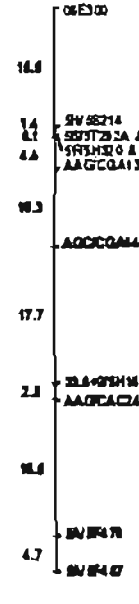
LG 28



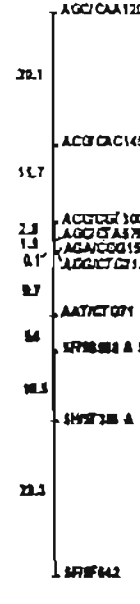
LG 29



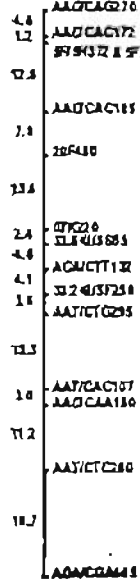
LG 30



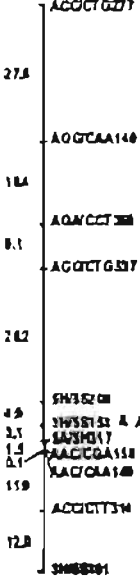
LG 31



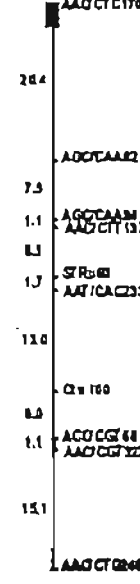
LG 32



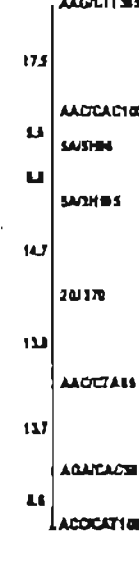
LG 33



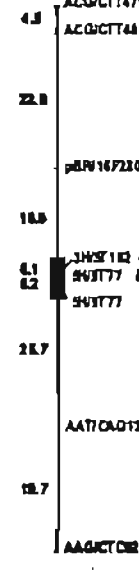
LG 34



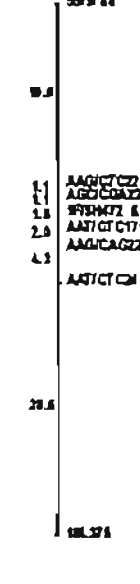
LG 35



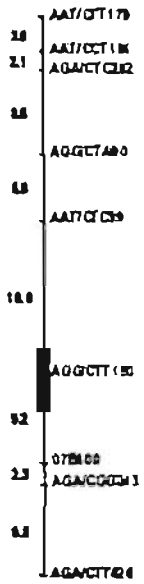
LG 36



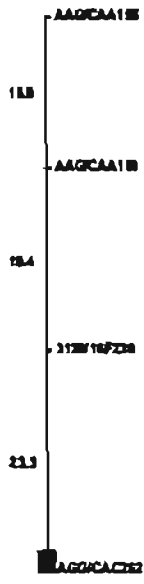
LG 37



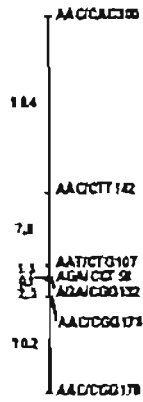
LG 38



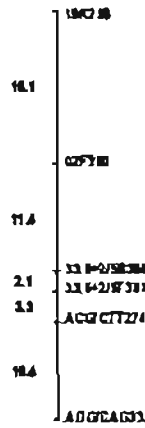
LG 39



LG 40



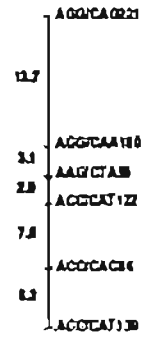
LG 41



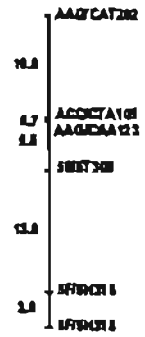
LG 42



LG 43



LG 44



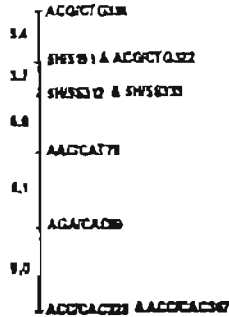
LG 45



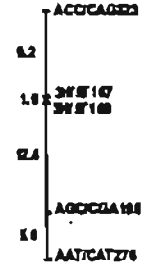
LG 46



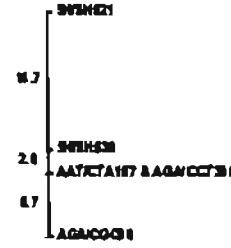
LG 47



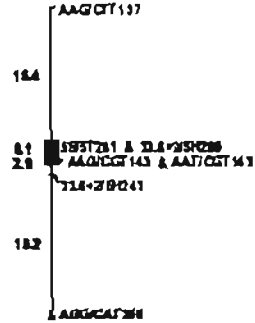
LG 48



LG 49



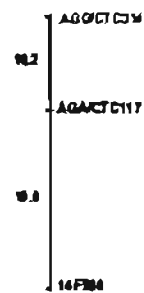
LG 50



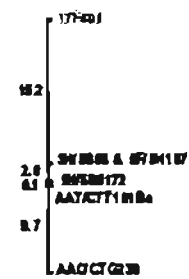
LG 51



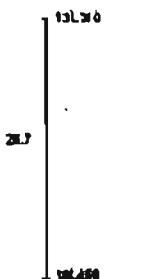
LG 52



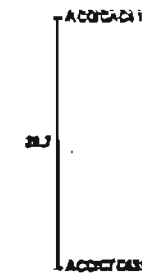
LG 53



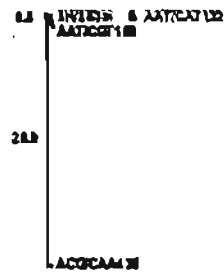
LG 54



LG 55



LG 56



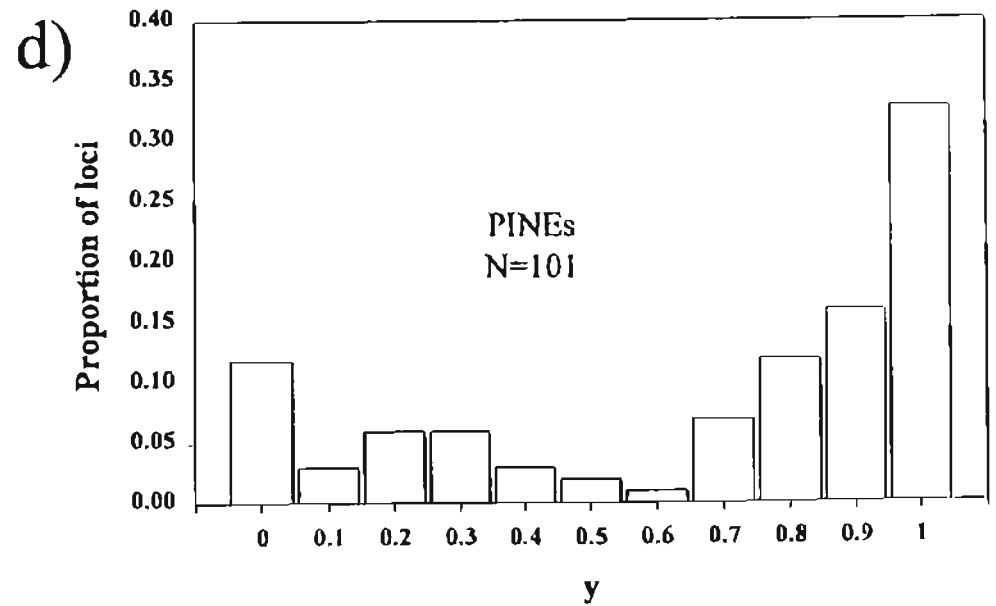
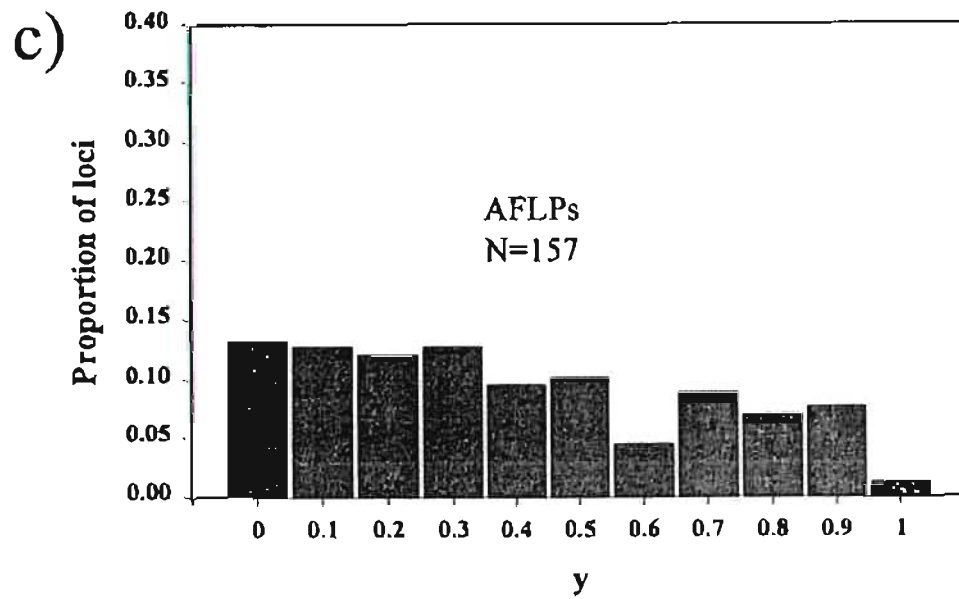
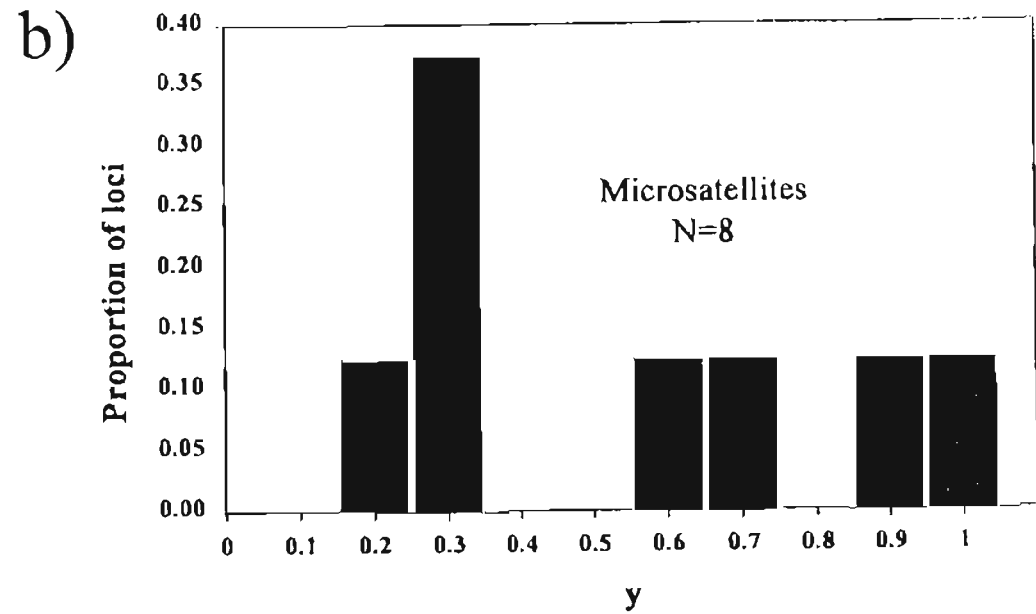
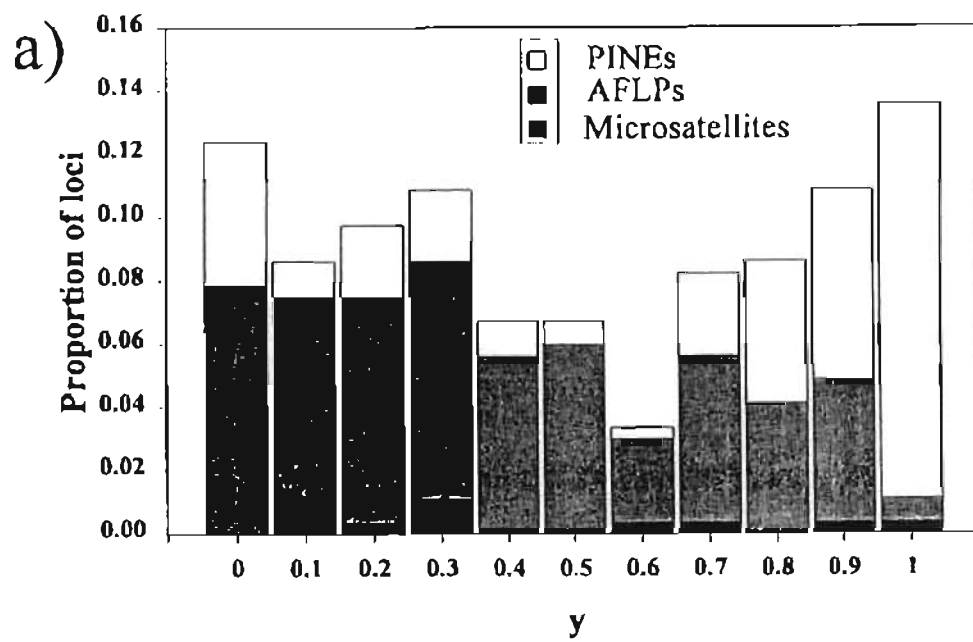


Figure 2. Distribution of distances between loci and their centromeres.

APPENDIX

Inheritance of nuclear DNA markers in gynogenetic haploid Pink Salmon

INHERITANCE OF NUCLEAR DNA MARKERS
IN GYNOGENETIC HAPLOID PINK SALMON

P. Spruell¹, K. L. Pilgrim¹, B. A. Greene^{2,3}, C. Habicht², K. L. Knudsen¹, K. R. Lindner¹,
G. K. Sage^{1,3}, J. E. Seeb², and F. W. Allendorf¹

¹Division of Biological Sciences, University of Montana, Missoula, MT 59812

²Alaska Department of Fish and Game, 333 Raspberry Road, Anchorage, AK 99518

³current address: United States Geological Survey, 1011 East Tudor Road, Anchorage,
AK 99503

Corresponding Author: Paul Spruell, (406) 243-5503, FAX (406) 243-4184,
spruell@selway.umt.edu

Acknowledgments: Supported by the *Exxon Valdez* Oil Spill Trustee Council (96160) and the National Science Foundation (DEB-9300135). We thank Robert H. Devlin for helpful suggestions on the use of the growth hormone pseudogene primers and generously sharing his unpublished primer sequences, Joan Strange for her assistance in sequence analysis, David Arsenault for technical assistance, and the Prince William Sound Aquaculture Corporation and the Valdez Fisheries Development Association for providing adult fish and gametes.

Running Title: Inheritance in haploid pink salmon

ABSTRACT

We describe the inheritance of 460 PCR-based loci in the polyploid-derived pink salmon (*Oncorhynchus gorbuscha*) genome using gynogenetic haploid embryos. We detected a length polymorphism in a growth hormone gene (*GH-2*) intron that is caused by a 81-base pair insertion homologous to the 3' end of the salmonid short interspersed repetitive element (SINE) SmaI. Such insertion polymorphisms within species bring into question the use of SINEs as phylogenetic markers. We confirmed that a microsatellite locus encodes a PCR-null allele that is responsible for an apparent deficit of heterozygotes in a population sample from Prince William Sound. Another set of microsatellite primers amplified alleles of the same molecular weight from both loci of a duplicated pair. In our analysis of several PCR-based multilocus techniques, we failed to detect evidence of co-migrating fragments produced by duplicated loci. Segregation analysis of PCR-based markers using gynogenetic haploid embryos ensures that the interpretation of molecular variation is not complicated by heterozygosity, diploidy, or gene duplication. We urge investigators to test the inheritance of polymorphisms in salmonids prior to using them to measure genetic variation.

INTRODUCTION

Fishes of the family Salmonidae comprise a monophyletic group descended from a single tetraploid ancestor (Allendorf and Thorgaard 1984; Behnke 1992). Salmonids have extensive gene duplication at protein loci resulting from this polyploid event (Allendorf and Thorgaard 1984). Studies of DNA sequences have confirmed the presence of many duplicate genes. For example, Agellon et al. (1988) reported duplicated growth hormone genes in rainbow trout (*Oncorhynchus mykiss*), and several other hormones have been found to be encoded by duplicated genes in *Oncorhynchus* species (Hiraoka et al. 1993). In addition, Dautigny et al. (1991) described the sequence divergence between two rainbow trout lysozyme genes.

The polyploid derived genome of salmonids has resulted in complex patterns of segregation and inheritance that have been revealed by the investigation of isozyme loci. Only disomic inheritance has been reported in females. Most loci in males are also inherited disomically. However, some loci show variable patterns of segregation in males, ranging from disomic ratios in some populations to tetrasomic ratios in other populations (Allendorf and Danzmann 1997). The residual tetrasomic inheritance observed in males apparently results from a two-stage pattern of pairing during male meiosis in which homologous chromosomes pair first followed by homeologous pairing. Disjunction of paired chromosomes occurs so that homologs segregate at the first meiotic division in males. Recombination events between homeologs produce segregation ratios approaching tetrasomic expectations for loci that are distant from their centromere, and therefore, more likely to be exchanged between homeologs.

The extensive gene duplication in salmonids makes genetic interpretation of molecular variation more difficult than in diploid species. Isoloci (two loci that result from a duplication event and share alleles with identical electrophoretic mobility) are especially problematic and constitute approximately 25% of isozyme markers in rainbow (Allendorf and Thorgaard 1984). Individuals have four gene copies at isoloci, and it is

difficult to determine how many copies (doses) of a particular allele are present in an individual. In addition, genotypes cannot be determined unambiguously, and there is no way to assign observed variation to a particular locus of the pair without extensive experimental matings (Waples 1988).

There are inherent difficulties in using the polymerase chain reaction (PCR) to study genetic variation. Preferential amplification of alleles at a single locus because of priming site polymorphisms and amplification of multiple paralogous loci are both potentially serious problems. Hare et al. (1996) have encountered and discussed these problems in an analysis of anonymous nuclear DNA markers in American oysters (*Crassostrea virginica*). These problems are likely to be even more serious in organisms such as salmonids that, as a result of their polyploid ancestry, have more duplicated loci. PCR primers designed without detailed knowledge of differences between paralogous loci may or may not amplify sequences from both loci. Moreover, even if only one locus is amplified, it will be difficult to ensure that homologous loci are being studied when comparing samples from two populations or two species. The complexities of tetrasomic inheritance and sex-specific recombination in salmonids further confounds these problems.

The complications in interpreting molecular variation in salmonids make it important to test the genetic basis of observed variation with inheritance experiments. Fortunately, external fertilization and well developed culture systems makes salmonids amenable to direct analysis of inheritance. Gametes can be stored and mixed together as desired to produce many full-sib groups from the gametes of a single male or female. In addition, methods of genome manipulation are available to produce large numbers of gynogenetic diploid and gynogenetic haploid progeny that provide more powerful methods of genetic analysis (Thorgaard and Allen 1987).

Examination of gynogenetic haploids provides an efficient system to test for Mendelian segregation and linkage without the complications associated with diploidy

and heterozygosity. For example, the use of haploid embryos avoids the difficulties associated with dominant PCR markers (those in which alleles are expressed as the presence or absence of an amplification product) since recessive alleles are not obscured by their dominant alternatives (Lie et al. 1994). Haploid embryos are not viable; however, they do develop until just prior to hatching (Stanley 1983), providing an embryo from which a sufficient quantity of DNA can be isolated to complete most analyses.

In this paper, we describe the inheritance of a variety of PCR-based markers in haploid pink salmon (*Oncorhynchus gorbuscha*). These include an intron length polymorphisms in a gene encoding growth hormone, eight microsatellite loci, and over 400 other loci detected by the presence or absence of specific fragments produced by several techniques that amplify multiple fragments from a single set of PCR primers. We also test for the presence of duplicated loci encoding fragments amplified by multilocus PCR-based techniques. The primary objective of this study is to detect and describe hundreds of genetic markers in the pink salmon genome so that we can eventually construct a linkage map that will allow us to better understand the transmission genetics of this polyploid-derived species.

MATERIALS AND METHODS

Samples and Haploid Gynogenesis

In August of 1995, gametes and tissues of 31 pink salmon were collected from the Armin F. Koernig hatchery, Prince William Sound, Alaska. This hatchery stock originated from adult fish collected at several spawning sites in 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 was pooled prior to UV irradiation, then mixed with the eggs from individual females. Females were numbered and their progeny were designated by year class and the number assigned to

that female (e.g. family 95-103). Embryos from these families were incubated until just prior to hatching when they were collected and preserved in ethanol.

Muscle or liver tissue was collected from each parent and embryos were dissected away from the egg chorion and yolk sack. DNA was isolated from these tissues using the Puregene(TM) DNA isolation kit (Gentra Systems Inc.). The concentration of DNA was determined using a scanning spectrofluorometer. DNA extractions from haploid embryos yielded an average of 45.30 μ g of DNA.

Prior to segregation analysis, we screened all putative gynogenetic haploid individuals to eliminate diploids that could be produced by the failure of sperm inactivation. We first used a Y chromosome-specific growth hormone pseudogene (Du et al. 1993, Forbes et al. 1994; primer sequences: 5'-TTTCTCTACGTCTACATTCT-3'; 5'-GTCTGGCTAGGGTACTCCA-3'; courtesy R. H. Devlin) to identify diploid males. Since haploids were produced by excluding the paternal chromosome complement, any individual containing a Y chromosome must be diploid. Males were identified based on the presence of a 143-bp fragment that is absent in females (Figure 1); males were eliminated from subsequent analyses. Failure of haploid induction could also produce diploid females. To identify diploid females, embryos were screened with six non-duplicated microsatellite loci described later in this paper. Individuals that had more than one allele at any of these loci were excluded from inheritance analysis.

Growth Hormone (GH) Intron

We amplified intron C of *GH-2* using previously described PCR primers and conditions (Forbes et al. 1994). PCR products were electrophoresed on a 2.0% agarose gel containing ethidium bromide in TAE buffer (Ausubel et al. 1989) and visualized with a Hitachi FMBIO-100 fluorescent imager. PCR amplification products from haploid individuals were purified from agarose gels using the GENECLEAN kit (BIO 101 Inc.) and sequenced by direct automated sequencing (Applied Biosystems Inc.).

Microsatellites

Analysis using seven previously described microsatellite primer sets followed the conditions reported by the original authors with minor modifications. Primers and annealing temperatures are as follows: Fgt-1 and Fgt-4, 51°C (Sakamoto et al. 1994); One μ 3, 52°C (Scribner et al. 1996); μ Sat60, 55°C (Estoup et al. 1993); Ots1, 55°C (Hedgecock, personal communication); Ssa85 and Ssa197, 57°C (O'Reilly et al. 1996). PCR products were electrophoresed on a 7% denaturing polyacrylamide gel and visualized with a Hitachi FMBIO-100 fluorescent imager.

Randomly Amplified Polymorphic DNA (RAPDs)

RAPD fragments were amplified in a total reaction volume of 10 μ L consisting of 10ng of genomic DNA, 6.7mM Operon Technology Inc. RAPD primer, 4.0 mM MgCl₂, 0.2mM of each dNTP, 1X Stoffel buffer, and 0.25U Amplitaq DNA polymerase Stoffel fragment (Perkin-Elmer). 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 electrophoresed on a 2% agarose gel containing ethidium bromide and visualized with a Hitachi FMBIO-100 fluorescent imager.

Amplified Fragment Length Polymorphisms (AFLPs)

AFLP restriction/ligation and pre-selective amplification steps were completed following the Perkin-Elmer/Applied Biosystems AFLP plant mapping protocol with modifications outlined below. Thermal cycling was performed in a MJ Research, PTC-200 DNA engine. The 10 μ L PCR mixture for the selective amplification consisted of 1.5

μL of the pre-selective amplification products as DNA template, 0.5 μl EcoRI selective primers, 0.5 μl MseI selective primers, 2 mM MgCl₂, 0.1 mM of each dNTP, 2X Amplitaq PCR buffer, and 0.5 U Amplitaq DNA polymerase. The following thermal profile was used for the selective amplification: 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 six cycles resulting in a final annealing temp of 59°C. An additional 30 cycles with an annealing temperature of 59°C for 30 seconds were also completed. Products were electrophoresed on a 7% denaturing polyacrylamide gel and visualized using an Hitachi FMBIO-100 fluorescent imager.

Paired Interspersed Nuclear Element PCR (PINEs)

PCR amplification of anonymous DNA fragments flanked by SINEs (short interspersed elements) and the Tc1 transposon was conducted in a total volume of 10 μl. Primers were designed on the basis of published sequences (Table 1). Each reaction contained approximately 20 ng of genomic DNA, 1 μl 1X Perkin-Elmer PCR buffer, 2.5 mM MgCl₂, 0.2 mM of each dNTP, 0.38 pM labeled primer and 0.5 U Perkin-Elmer AmpliTaQ DNA polymerase Stoffel fragment. Reactions were completed in an MJ Research thermocycler using an annealing temperature of 60°C. Products were electrophoresed on a 4.5% denaturing polyacrylamide gel and visualized using an Hitachi FMBIO-100 fluorescent imager.

Nomenclature

A standard method for naming microsatellite loci in salmonids has been informally adopted (see Olsen et al. 1996). Primers are typically named after the species from which they are derived: Ssa (Atlantic salmon, *Salmo salar*) and Ots (chinook salmon, *Oncorhynchus tshawytscha*). Some microsatellites were named prior to the

advent of this standardized nomenclature. The μ Sat60 primers were isolated from brown trout (*Salmo trutta*), and the Fgt1 primers were isolated from rainbow trout but named fish GT-repeat. Locus names are the primer pair name in upper-case and italics (e.g., *OTS1*) to make them analogous to the nomenclature for allozyme loci (Shaklee et al. 1990). The nomenclature for duplicated loci follows the format used for duplicated allozyme loci (Shaklee et al. 1990). For example, *FGT1-1,2* designates isoloci produced by the Fgt-1 primer set.

We followed nomenclature used for zebrafish (*Brachydanio rerio*) in naming loci for RAPD markers (Johnson et al. 1996). The name consists of the name of the 10 nucleotide long primer followed by the approximate size of the amplification product. Thus, the locus *20A.760* is amplified by primer A20 and results in a 760-bp amplification product.

The nomenclature for AFLPs is consistent with Young et al. (in press) in their description of a rainbow trout linkage map. The names start with the three base selective primer extensions used to produce the loci and end with the length of the fragment measured in base pairs (e.g. *AAA/CAT250*).

PINE loci are named using a number designating the end of the element from which the primer was derived (3' or 5') followed by a one letter designator for the element from which the primer was derived. If more than one primer was used during the amplification the primers are placed in alphabetical order. The primer designations are followed by the length of the fragment. For example, the locus *5F3H250* amplifies a 250-bp fragment using a primer sequence from the 5' end of the FokI SINE and the 3' end of the HpaI SINE.

Alleles are designated as **p* for the presence of a product and **a* for the absence of a product for multifragment PCR-based techniques (RAPDs, AFLPs, and PINES). Alleles that differ in length are designated by a number representing their size. For

example, *5F3H250*270* designates an allele encoding a 270 bp-fragment at a PINE locus at which the common allele encodes a 250-bp fragment.

RESULTS

Growth Hormone

The *GH-2* intron C primers gave products of two different lengths (446-bp and 527-bp; Figures 2 and 3). We designated the two alleles that produce these fragments as *GH-2*C446* and *GH-2*C527* according to the nomenclature guidelines for protein-coding loci in fish (Shaklee et al. 1990). Sequencing revealed that this length difference is caused by an 81-bp insert that is nearly identical to the 3' end of the consensus sequence of the *SmaI* SINE in pink salmon (Kido et al. 1991; Figure 2).

Seven of the 31 adults were heterozygotes at this locus and the remaining 24 were **C446* homozygotes. Two females, 95-105 and 95-115 were heterozygous at this locus; their seventy-two haploid progeny displayed the expected 1:1 Mendelian segregation for these alleles.

Microsatellites

The seven microsatellite primer sets examined were polymorphic in the 31 adult fish. Four microsatellite loci (*FGT4*, *ONE μ 3*, *OTS1*, and *μ SAT60*) were in Hardy-Weinberg proportions in the adult fish (Table 2), and exhibited expected Mendelian segregation (e.g., Table 3). *SSA85* exhibited expected Mendelian segregation, but was not used in the analysis of adult fish because genotypes were difficult to score. This locus had a minimum of 12 alleles in the adult fish and each allele produced multiple "stutter" bands (Litt and Luty 1989; Hayashi 1994). The overlapping patterns of these additional products makes unambiguous identification of alleles impossible when alleles of similar size are present. The alleles in the heterozygous females were sufficiently different in size such that there was no overlap in the patterns generated in haploids from a single female.

One of the microsatellite primer sets (*Fgt1*) produced phenotypes that indicated more than two alleles in diploid individuals and more than one allele in haploids.

FGT1-1,2 has previously been described as duplicated isoloci in sockeye salmon (*Oncorhynchus nerka*; Allendorf et al. submitted) and rainbow trout (Young et al. in press). Inheritance results confirm that *FGT1-1,2* are also isoloci in pink salmon (Table 4). This is seen most clearly in the progeny from female 95-106 who possessed both the *155 and *157 alleles. All 37 of her progeny received both alleles; thus, she must have been homozygous at both *FGT1* loci (-1 and -2) as indicated in Table 4. This locus was not scored in the sample of adult fish because of the problems in scoring doses at isoloci (Allendorf et al. submitted).

PCR amplification of *SSA197* produced 15 alleles and a highly significant excess of apparent homozygotes in the adult fish (Table 2). In addition, no PCR products were detected in two of the 31 fish. The cause of these results became clear in the segregation experiments. All four single-banded females for which we examined haploid progeny were actually heterozygotes for a PCR-null allele (*SSA197*a*) that produced no amplification product (Table 5). Approximately one-half of the progeny from each of these females had the same fragment as the mother, the other half produced no PCR product (Fig. 4). PCR products were detected at all other loci in the two adults and all haploid progeny that contained the null, eliminating the possibility that these results were caused by poor quality DNA samples.

We reanalyzed the genotypes at this locus in the 31 adult fish including the null allele. We assumed that all apparently homozygous fish were heterozygotes for a null allele and the two fish lacking product were null homozygotes (*SSA197*a/a*). The estimated frequency of *SSA197*a* under these assumptions is 0.258 and the observed genotypic proportions do not differ from Hardy-Weinberg expectations (Table 2).

Multilocus Primer Sets

We screened 140 RAPD primers, or primer pairs, in the haploid progeny from female 95-103. Each primer set produced approximately 5-8 fragments from 400 to 1500

bp. We detected 36 repeatable presence / absence polymorphisms amplified by 25 RAPD primers. All of these markers demonstrated Mendelian segregation in 94 haploid progeny from female 95-103.

We screened 77 AFLP primer combinations in the haploid progeny from female 95-103. Each AFLP primer combination produced at least 30 bands ranging from 50 to 600 bp. We selected 43 primer combinations that amplified 284 clear polymorphisms that segregated in 94 progeny from female 95-103. Almost all of the AFLP polymorphisms were presence / absence differences. However, four of the polymorphisms appeared to be caused by a length polymorphism within a fragment. For all of these polymorphisms, individuals had one of two different sized fragments produced by the same primer combination.

We used DNA sequences of salmonid-specific SINEs and the transposon Tc1 as primers to generate multiple DNA fragments from a single PCR. This procedure is similar to the use of the human SINE AluI to identify human chromosomes in somatic cell hybridization experiments (Nelson et al. 1989). Primers identical 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 pairs of primers may be used to generate multilocus patterns (Greene and Seeb in press).

A minimum of 30 fragments is amplified by each combination of primers (Fig. 5). We have scored 94 haploid offspring from female 95-103 with 16 PINE primer combinations that produce a total of 131 polymorphic loci. In six cases it appears that PINE fragments are segregating as codominant alleles that vary in length (Fig. 5).

DISCUSSION

The examination of haploid embryos is a powerful tool for segregation analysis (Slettan et al. 1997). It allows the unambiguous detection of the transmission of recessive alleles to progeny. Similarly, it allows the *direct detection of PCR-null alleles at*

microsatellite loci, such as *SSA197*a*. Haploid progeny also facilitates the direct sequencing of allelic variants without the problems of heterozygosity (either known or cryptic).

The AFLP technique is particularly well suited for use with haploid embryos. The amount of DNA required for analysis is an important consideration when dealing with the limited amount of tissue available from haploid embryos. Two properties of AFLPs maximize the information that can be obtained from the limited DNA available. First, many bands are produced per reaction, and, therefore, more polymorphic loci are produced per PCR amplification. 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.

Segregation of Recessive Alleles

The occurrence of isoloci makes the interpretation of recessive markers problematic in salmonids. Observed segregation patterns may result from a pair of isoloci (e.g., *LOCUS-1,2*) that are both heterozygous for alleles associated with the presence or absence of a particular fragment (*LOCUS-1*p/a; LOCUS-2*p/a*). In this case, we expect a 3:1 ratio of presence to absence of the fragment in haploid progeny (25% *p/p*: 50% *p/a*: 25% *a/a*). In addition, this same ratio will result if a female is heterozygous (**p/a*) at two non-homologous loci that happen to produce fragments of the same size.

It is difficult to distinguish between a 1:1 and 3:1 ratio for an individual fragment except with very large sample sizes. However, the presence of such pairs of loci segregating 3:1 for the presence or absence of a fragment should affect the observed segregation ratios. In the absence of any such cases, we expect our observed segregation

ratios to fit a binomial distribution with an expectation of 0.5 (1:1 segregation). The presence of markers segregating 3:1 should result in a "shoulder" in the distribution at a value of 0.75. There is perhaps a slight excess of loci segregating with a value of 0.6 or greater at 451 loci segregating from female 95-103 (Figure 6). We conclude that fragments segregating 3:1 represent at most a small fraction of the total fragments that we have examined.

PINEs

SINEs and transposons occur in high copy number and are believed to be ubiquitously dispersed throughout the genomes of many species (Okada 1991). These characteristics make PINEs potentially valuable tools for genomic mapping efforts. Unlike other multilocus techniques, the primers used to generate PINEs are based on repetitive elements known to exist in the salmonid genome. In addition to generating markers, the inclusion of PINEs in our mapping efforts may also increase our understanding of SINEs and transposons in the salmonid genome.

Others have used the presence or absence of families of SINEs or specific SINEs to make phylogenetic inferences (Kido et al. 1991; Murata et al. 1993; Murata et al. 1996). However, the mechanisms of SINE amplification are not entirely known and evidence is accumulating that the genomic distribution of SINEs may be more complex than previously believed (Spruell and Thorgaard 1996; Takasaki et al. 1997; Young et al. in press).

The insertion that we have described in intron C of *GH-2* corresponds to the 3' end of the *Sma*I element (Figure 2). This result is consistent with the observations of Spruell and Thorgaard (1996) and Young et al. (in press) who suggested that the sequences corresponding to some regions of SINEs may be distributed independently of the remainder of the element. Moreover, this *GH-2* insertion is not present in other *Oncorhynchus* species for which this intron has been sequenced (Figure 1; *O. mykiss*,

Agellon et al. 1988; *O. nerka*, Devlin 1993; *O. kisutch*, Forbes et al. 1994; *O. keta*, Shen et al. 1993). Thus, the insertion seen in the *GH-2*C446* allele apparently occurred after pink salmon diverged from other *Oncorhynchus* species.

This pink salmon-specific insertion is unexpected if the amplification of SmaI and SmaI-related sequences occurred in a common ancestor of pink and chum (*Oncorhynchus keta*) salmon as proposed by Kido et al. (1991). Takasaki et al. (1997) also report a lack of concordance between the presence or absence of specific SmaI elements within pink and chum salmon. They propose several possible explanations for this phenomenon. Among those is the possibility of temporal differences in amplification within lineages. If this hypothesis is correct, the possibility of insertion polymorphisms within species must be addressed before these elements are used as phylogenetic markers.

Gene Duplication

Duplicated loci are extremely difficult to use for population genetic analysis. Accurate estimation of allele frequencies at isoloci requires determining the numbers of copies of each allele in individuals (Waples 1988). Isoloci at allozymes are routinely used for population genetic analysis. This is possible because there is a correspondence between band intensity and doses of an allele present in allozymes. (Shaklee and Phelps 1992; Allendorf and Danzmann 1997). In addition, the presence of heteromeric isozymes and tissue-specificity of many loci aid in estimating doses for enzymes (Waples 1988).

It is difficult to determine how many doses of each allele are present in PCR-based techniques because the amount of amplified product may not accurately reflect allelic doses (Wagner et al. 1994). The many alleles present at most microsatellite loci will also make analysis and allele frequency estimation much more difficult. For a tetrasomic locus with n alleles, there are $(n+3)!/(n-1)!$ different genotypes (p. 610, Hartl and Clark 1989). Thus, there are 330 possible genotypes at *OTS1* with eight alleles. May et al. (1997) have recently suggested a method for estimating doses at microsatellite loci that may be helpful when working with duplicated loci.

Perhaps the best way to deal with duplicated microsatellite loci in salmonids is to not use them for population genetic analysis. There are enough microsatellite markers available to obtain a sufficient number of markers without using duplicated microsatellites. Approximately 25% of isozyme markers in rainbow trout are encoded by isoloci (Allendorf and Thorgaard 1984). We would expect the proportion of microsatellites encoded by isoloci to be somewhat less than this since their higher mutation rate will cause more rapid divergence between alleles at two loci that are no longer undergoing residual tetrasomic inheritance (Allendorf and Danzmann 1997). Nevertheless, the processes of diploidization in salmonids is incomplete and we would expect recombination between homeologs to transfer alleles between some microsatellite loci (Allendorf and Danzmann 1997).

Duplicated microsatellite loci in salmonids can be used in many applications (e.g., paternity and kinship analysis). However, it is critical that the inheritance of such loci be tested in the population being investigated because of PCR null alleles and the possibility of residual tetrasomy in some populations and not others (Allendorf and Danzmann 1997).

The many nuclear DNA markers available offer a wealth of opportunities for greatly improving our understanding of the transmission and population genetics of salmonids. Nevertheless, problems in genetic interpretation are in some ways greater in the direct examination of DNA itself using PCR than in the study of genetic variation in proteins. Allozyme electrophoresis only detects functional genes so that pseudogenes are not a complication. Moreover, the tissue specific expression of protein loci has been used to identify specific loci within sets of paralogous loci (Ferris and Whitt 1979). For example, only one of the two paralogous duplicates of the vertebrate lactate dehydrogenase B gene (LDH-B) in salmonids is expressed in liver tissue, and the product of the other paralogous locus predominates in heart tissue. This pattern of expression has been conserved over a long period of evolutionary time, and is shared among all species

of two of the three subfamilies of salmonids: Thymallinae (grayling) and Salmoninae (trout, salmon, and char) (Allendorf and Thorgaard 1984). This consistent pattern of expression makes it easy to identify each of the two paralogous LDH-B loci in salmonids. PCR-based markers are amplified solely on the basis of DNA sequences present, therefore, differentiation of pairs of paralogous loci is impossible without inheritance data.

The analysis of gynogenetic haploids is a powerful tool for understanding the transmission of genetic markers in salmonid fishes. Population frequencies of PCR-amplified gel bands alone will not be adequate to understand the genetic basis and significance of observed variation in salmonids. We urge investigators to use inheritance studies to confirm the genetic basis of observed polymorphisms in salmonids whenever possible.

REFERENCES

- Agellon LB, Davies SL, Lin CM, Chen TT, and Powers DA, 1988. Rainbow trout has two genes for growth hormone. *Mol Reprod Devel* 1:11-17.
- Allendorf FW, and Danzmann RG, 1997. Secondary tetrasomic segregation of MDH-B and preferential pairing of homeologues in rainbow trout. *Genetics* 145:1083-1092.
- Allendorf FW, and Thorgaard G, 1984. Polyploidy and the evolution of salmonid fishes. In: *The evolutionary genetics of fishes*, (Turner BJ, ed.), New York: Plenum Press; 1-53.
- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, and Struhl K, 1989. *Current protocols in molecular biology*. New York: John Wiley & Sons.
- Behnke RJ, 1992. *Native trout of western North America*. Bethesda, MD: American Fisheries Society Monograph 6.
- Dautigny A, Prager EM, Pham-dinh D, Jolles J, Pakdel F, Grinde B, and Jolles P, 1991. Complementary DNA and amino acid sequences of rainbow trout (*Oncorhynchus mykiss*) lysozymes and their implications for the evolution of lysozyme and lactalbumin. *J Mol Evol* 32:187-198.
- Devlin RH, 1993. Sequence of sockeye salmon type 1 and 2 growth hormone genes and the relationship of rainbow trout with Atlantic and Pacific salmon. *Can J Fish Aquat Sci* 50:1738-1748.
- Du SJ, Devlin RH, and Hew CL, 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-psi. *DNA and Cell Biol* 12:739-751.
- Estoup A, Presa P, Krieg F, Vaiman D, and Guyomard R, 1993. (CT)(n) and (GT)(n)

- microsatellites - a new class of genetic markers for *Salmo trutta* l (brown trout). *Heredity* 71:488-496.
- Ferris SD, and Whitt GS, 1979. Evolution of the differential regulation of duplicate genes after polyploidization. *J Mol Evol* 12:267-317.
- Forbes SH, Knudsen KL, North TW, and Allendorf FW, 1994. One of two growth hormone genes in coho salmon is sex-linked. *Proc Natl Acad Sci USA* 91:1628-1631.
- Greene BA, and Seeb JE, 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*). *Mol Mar Biol Biotech* 4:331-341.
- Hare MP, Karl SA, and Avise JA, 1996. Anonymous nuclear DNA markers in the American oyster and their implications for the heterozygote deficiency phenomenon in marine bivalves. *Mol Biol Evol* 13:334-345.
- Hartl DL, and Clark AG, 1989. *Principles of Population Genetics*, 2nd ed. Sunderland, MA: Sinauer Assoc.
- Hayashi K, 1994. Manipulation of DNA by PCR. In: *The polymerase chain reaction*, (Mullis KB, Ferre F, and Gibbs RA, eds.), York, Pennsylvania: The Maple Press Company; 3-13.
- Hiraoka S, Suzuki M, Yanagisawa T, Iwata M, and Urano A, 1993. Divergence of gene expression in neurohypophysial hormone precursors among salmonids. *Gen Comp Endocrin* 92:292-301.
- Johnson SL, Gates MA, Johnson M, Talbot WS, Horne S, Baik K, Rude S, Wong JR, and Postlethwait JH, 1996. Centromere-linkage analysis and consolidation of the zebrafish genetic map. *Genetics* 142:1277-1288.
- Kido Y, Aono M, Yamaki T, Matsumoto K, Murata S, Saneyoshi M, and Okada N,

1991. Shaping and reshaping of salmonid genomes by amplification of tRNA-derived retroposons during evolution. *Proc Natl Acad Sci USA* 88:2326-2330.
- Lie O, Slettan A, Lingaas F, Olsaker I, Hordvik I, and Refstie T, 1994. Haploid gynogenesis: A powerful strategy for linkage analysis in fish. *Anim Biotech* 5:33-45.
- Litt M, and Luty JA, 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am J Hum Gen* 44:397-401.
- May B, Krueger CC, and Kincaid HL, 1997. Genetic variation at microsatellite loci in sturgeon: Primer sequence homology in *Acipenser* and *Scaphirhynchus*. *Can J Fish Aquat Sci* 54:1542-1547.
- Murata S, Takasaki N, Saitoh M, and Okada N, 1993. Determination of the phylogenetic relationships among Pacific salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. *Proc Natl Acad Sci USA* 90:6995-6999.
- Murata S, Takasaki N, Saitoh M, Tachida H, and Okada N, 1996. Details of retropositional genome dynamics that provide a rationale for genetic division: the distinct branching of all the Pacific salmon and trout (*Oncorhynchus*) from the Atlantic salmon and trout (*Salmo*). *Genetics* 142:915-926.
- Nelson DL, Ledbetter SA, Corbo L, Victoria MF, Ramirez-Solis R, Webster TD, Ledbetter DH, and Caskey CT, 1989. *Alu* polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources. *Proc Natl Acad Sci USA* 86:6686-6690.
- Okada N, 1991. SINEs: Short interspersed repeated elements of the eukaryotic genome. *Trends Ecol Evol* 6:358-361.
- Olsen JB, Wenburg JK, and Bentzen P, 1996. Semiautomated multilocus genotyping

- of Pacific salmon (*Oncorhynchus* spp.) using microsatellites. *Mol Mar Biol Biotech* 5:259-272.
- O'Reilly PT, Hamilton LC, McConnell SK, and Wright JM, 1996. Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. *Can J Fish Aquat Sci* 53:2292-2298.
- Sakamoto T, Okamoto N, Ikeda Y, Nakamura Y, and Sato T, 1994. Dinucleotide-repeat polymorphism in DNA of rainbow trout and its application in fisheries science. *J Fish Biol* 44:1093-1096.
- Scribner KT, Gust JR, and Fields RL, 1996. Isolation and characterization of novel salmon microsatellite loci: Cross-species amplification and population genetic applications. *Can J Fish Aquat Sci* 53:833-841.
- Shaklee JB, and Phelps SR, 1992. Chinook salmon NADP-dependent cytosolic isocitrate dehydrogenase: Electrophoretic and genetic dissection of a complex isozyme system and geographic patterns of variation. *Biochem Genet* 30:455-489.
- Shaklee JB, Allendorf FW, Morizot DC, and Whitt GS, 1990. Gene nomenclature for protein-coding loci in fish. *Trans Am Fish Soc* 119:2-15.
- Shen X, Wang Y, Welt M, Liu D, and Leung FC, 1993. Molecular cloning and sequence analysis of the chum salmon growth hormone genomic gene. *GenBank Accession Num* L04688.
- Slettan A, Olsaker I, and Lie O, 1997. Segregation studies and linkage analysis of Atlantic salmon microsatellites using haploid genetics. *Heredity* 78:620-627.
- Spruell P, and Thorgaard GH, 1996. SINE sequences detect DNA fingerprints in salmonid fishes. *Heredity* 76:317-324.
- Stanley JG, 1983. Gene expression in haploid embryos of Atlantic salmon. *J. Hered* 74:19-22.

- Takasaki N, Yamaki T, Hamada M, Park L, and Okada N, 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.
- Thorgaard GH, and Allen SK, 1987. Chromosome manipulation and markers in fishery management. In: *Population Genetics and Fisheries Management*, (Utter F and Ryman N, eds.), Seattle, WA: Univ. Washington Press; 319-331.
- Thorgaard GH, Allendorf FW, and Knudsen KL, 1983. Gene-centromere mapping in rainbow trout: High interference over long map distances. *Genetics* 103:771-783.
- Wagner A, Blackstone N, Cartwright P, Dick M, Misof B, Snow P, Wagner GP, Bartels J, Murtha M, and Pendleton J, 1994. Surveys of gene families using polymerase chain reaction - PCR selection and PCR drift. *Syst Biol* 43:250-261.
- Waples RS, 1988. Estimation of allele frequencies at isoloci. *Genetics* 118:371-384.
- Young WP, Wheeler PA, Coryell VH, Keim P, and Thorgaard GH, in press. A detailed linkage map of rainbow trout produced using doubled haploids. *Genetics*.

Table 1. Primer sequences used for Paired Interspersed Nuclear Element PCR (PINEs) and references.

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
Tcl 5'	GTATGTAAACTTCTGACCCACTGG	Greene and Seeb in press

Table 2. Summary of genetic variation at five microsatellite loci in adult pink salmon from Prince William Sound. F is the fixation index (the proportional excess of heterozygotes).

Locus	No. individuals	No. alleles	Heterozygosity		F
			Observed	Expected	
<i>FGT4</i>	31	2	0.290	0.398	0.271
<i>ONEμ3</i>	31	3	0.548	0.505	-0.085
<i>OTS1</i>	31	8	0.806	0.791	-0.019
<i>μSAT60</i>	31	4	0.290	0.414	0.300
<i>SSA197</i>	29	15	0.586	0.912	0.361 ***
<i>SSA197</i> ^a	31	16	0.935	0.885	-0.056

^a including the null allele

*** $P < 0.001$

Table 3. Inheritance of *OTS1* in gynogenetic haploid progeny.

Female		Progeny Phenotype						
Number	Genotype	218	220	222	224	226	228	230
95-101	222/226	-	-	22	-	17	-	-
95-102	218/224	21	-	-	16	-	-	-
95-103	224/230	-	-	-	53	-	-	38
95-104	224/224	-	-	-	39	-	-	-
95-105	220/226	-	20	-	-	15	-	-
95-106	220/224	-	21	-	17	-	-	-
95-115	226/228	-	-	-	-	17	18	-

Table 4. Inheritance of *FGT1-1,2* isoloci in gynogenetic haplod progeny.

Female		Progeny Phenotype			Female Genotype
Number	Phenotype	155/155	155/157	157/157	
95-101	155/157	-	20	19	<i>155/157 157/157</i>
95-102	155/157	20	16	-	<i>155/155 155/157</i>
95-103	155/157	-	44	47	<i>155/157 157/157</i>
95-104	155/157	-	24	14	<i>155/157 157/157</i>
95-105	155/157	-	18	18	<i>155/157 157/157</i>
95-106	155/157	-	37	-	<i>155/155 157/157</i>
95-115	155/157	17	19	-	<i>155/155 155/157</i>

Table 5. Inheritance of *SSA197* in gynogenetic haploid progeny.

Female		Progeny Phenotypes									Female Genotype
Number	Phenotype	130	142	146	154	162	164	190	196	null	
95-101	146/162	-	-	21	-	18	-	-	-	-	<i>146/162</i>
95-102	162	-	-	-	-	15	-	-	-	20	<i>162/a</i>
95-103	164	-	-	-	-	-	46	-	-	55	<i>164/a</i>
95-104	196	-	-	-	-	-	-	-	22	17	<i>196/a</i>
95-105	130/154	41	-	-	22	-	-	-	-	-	<i>130/154</i>
95-106	142	-	20	-	-	-	-	-	-	18	<i>142/a</i>
95-115	154/190	-	-	-	24	-	-	12	-	-	<i>154/190</i>

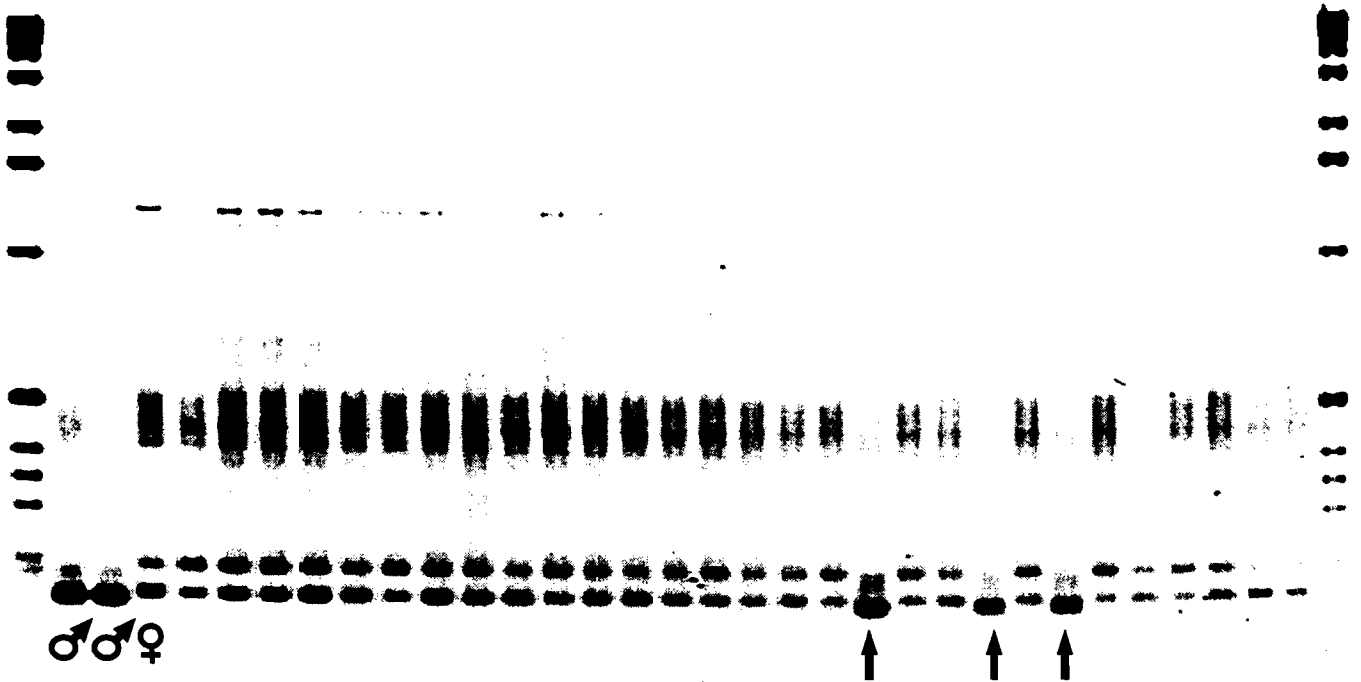


Figure 1. Growth hormone pseudogene amplication products separated in a 2% agarose gel. Individuals of known sex are indicated by the symbols on the bottom left of the gel. The arrows along the bottom of the gel indicate males that were detected in family 95-103.

Sockeye	AACCCATCACTCTCTAATCGGCGGTTTCTCTACGTCTACATTCTCCAGCAATGTGTCATG	359
GH2 * C446 T..... G..... A.....	360
GH2 * C527 T..... AG..... A.....	360
Sockeye	TAAA	363
GH2 * C446	364
GH2 * C527 TAATA- TAATAATA- TAATAATATATGCCATTTAGCAGACGCTTTTATCCAAAGC	417
Sma I	TAATAATAATAATAATA- TAATATATGCCATTTAGCAGACGCTTTTATCCAAAGC	105
Sockeye	TGATATGGCATCTCAAGCTGTACAATTACAA	394
GH2 * C446	395
GH2 * C527	GACTTACAGTCATGTGTGCATACATAAA 	476
Sma I	GACTTACAGTCATGTGTGCATACATTCT	
Sockeye	CTCAACTTCATTTTCTAATAATCTGTGGTTTCTCTACATCTACACACACAG	445
GH2 * C446	446
GH2 * C527	527

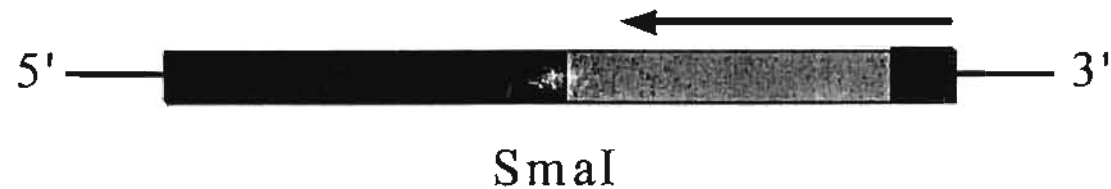


Figure 2. Aligned sequences of GH-2 intron C from sockeye (Devlin 1993) and pink salmon. The 81 bp insert found in GH2*C527 is indicated by the dark bars in the sequences. The region that corresponds to the 81 bp insert and its orientation relative to the GH-2 gene is denoted by the arrow above the element. The solid shaded area corresponds to the tRNA-related region, the hatched region corresponds to the tRNA-unrelated region, and the open region is the AT rich region Okada (1991).

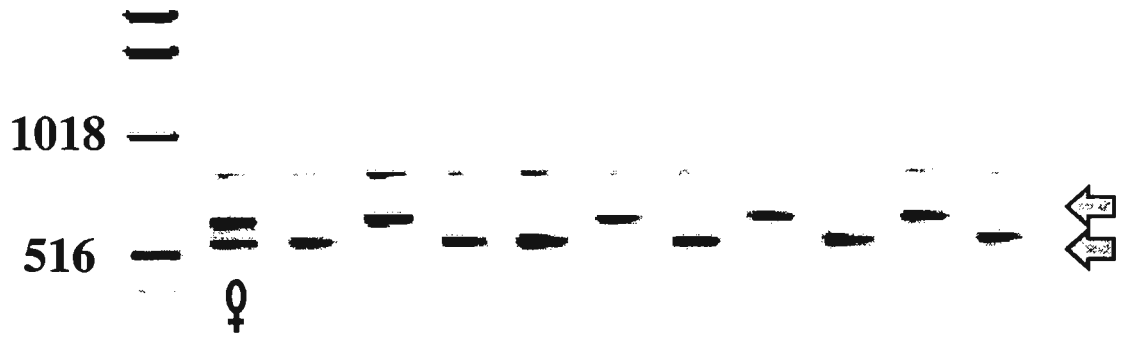


Figure 3. Segregation of GH-2 polymorphism. The individual denoted by the female symbol is female 95-105, the 10 individuals to the right are haploid offspring from this female segregating for the GH-2*C446 and *C527 alleles indicated by the arrows.



Figure 4. Segregation of a “null” allele (no PCR product is amplified) at *SSA197*. The individual denoted by the symbol is female 95-106, the 14 individuals to the right are haploid offspring from this female. The arrow indicates allele *SSA197**142.

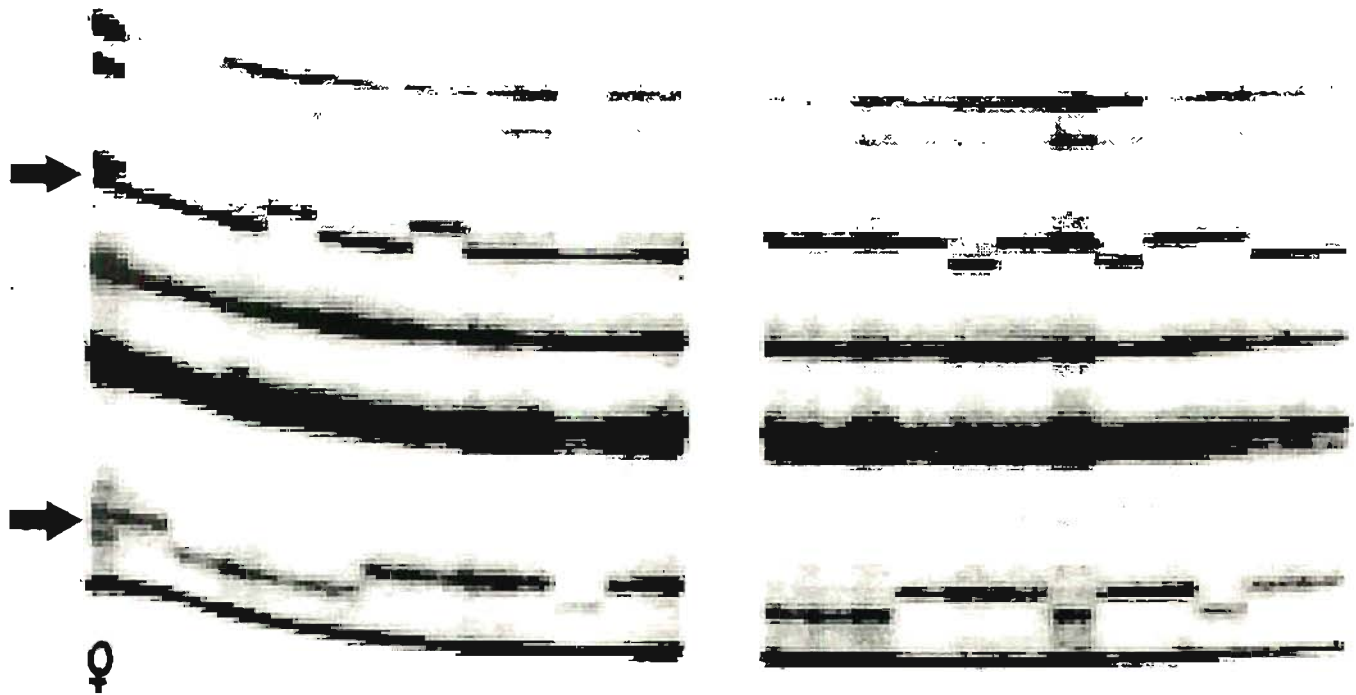


Figure 5. Hpa 3' and TcI PINE primer amplification products separated on a 4.5% polyacrylamide gel. Female 95-103 is indicated. The two arrows along the left side of the gel indicate length polymorphisms: top bands 181/182, bottom bands 166/167 bp.

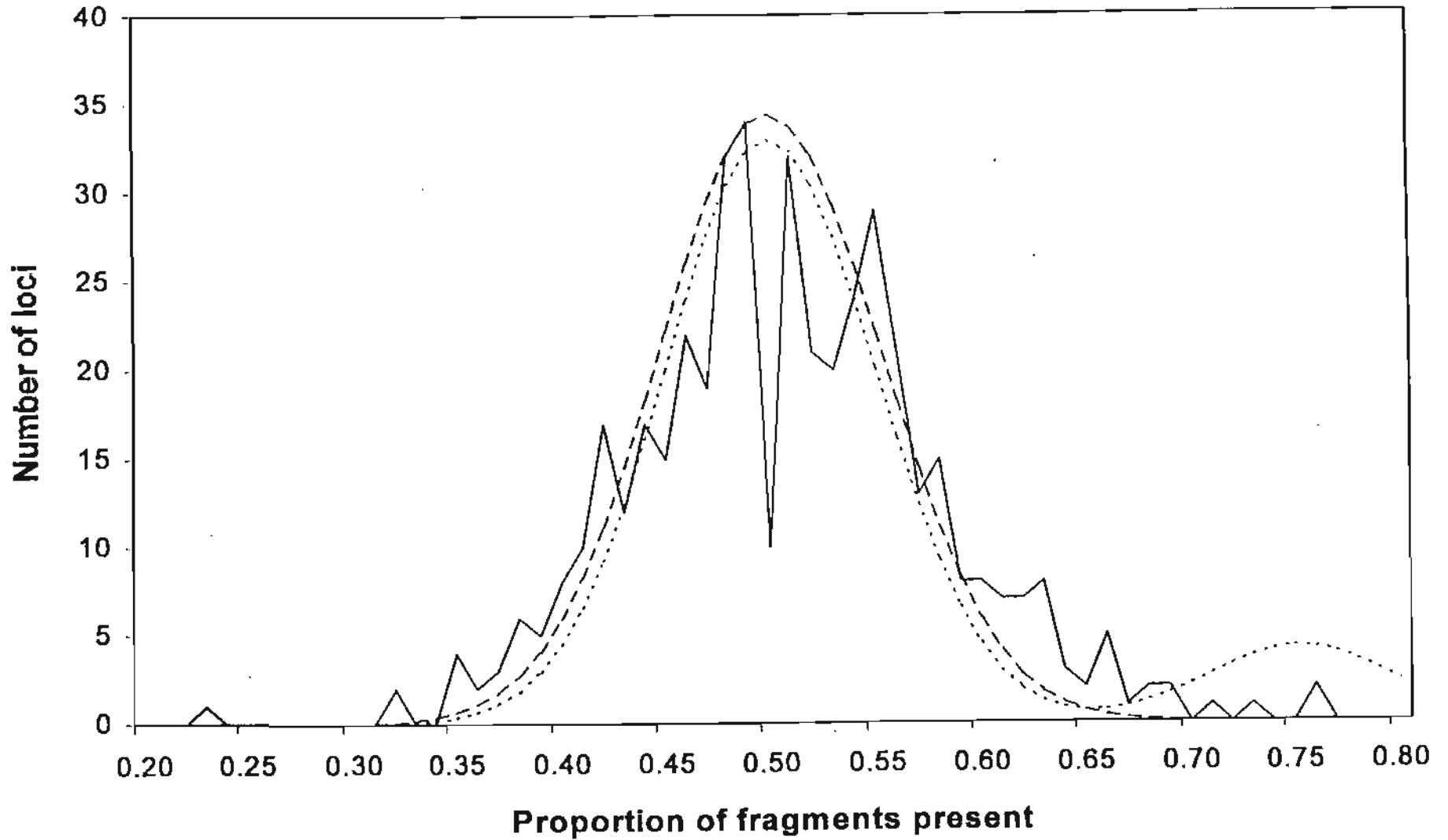


Figure 6. Distribution of segregation ratios for 451 fragments in haploid progeny from female 95-103 (solid line). The dashed line is the expected value of 0.5 (1:1) segregation. The dotted line is the expected binomial distribution for 90% of the fragments segregating 1:1 and 10% of the fragments segregating 3:1.