GENETIC MAPPING IN ATLANTIC SALMON (SALMO SALAR) AND BROWN TROUT (SALMO TRUTTA)

COLIN McGOWAN

GENETIC MAPPING IN ATLANTIC SALMON (*SALMO SALAR*) AND BROWN TROUT (*SALMO TRUTTA*)

by

Colin McGowan

A thesis submitted to the School of Graduate Studies in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Biochemistry

Faculty of Science

Memorial University of Newfoundland

October 1996

St. John's Newfoundland

ABSTRACT

Segregation analysis of polymorphic genetic markers was performed using the hybrid progeny of a brown trout (*Salmo trutta*) x Atlantic salmon (*Salmo salar*) cross. Three types of genetic markers, 1) RAPDs (randomly amplified polymorphic DNA), 2) microsatellites and 3) ESTs (expressed sequence tags) were examined.

1) Of the 271 random primers tested, 151 (56%) generated distinct, reproducible band patterns. A total 893 loci were scored, of which 126 could be attributed to the Atlantic salmon component of the genome and 173 to that of the brown trout. The majority (594) of loci were held in common by both parents. Three and 13 heterozygous loci were identified in the Atlantic salmon and brown trout parents, respectively. All of the markers segregated in a normal (1:1) manner. No linkage was observed in Atlantic salmon. Thirteen significant ($P<0.05$) pair-wise associations were identified in brown trout, five of which were highly significant ($P<0.001$).

The RAPD procedure was also used for bulked segregant analysis of Atlantic salmon male and female DNA, to search for a sex-specific marker. A total of 1152 loci were screened using 200 different primers; however none were tightly linked to a male sex determining factor.

2) Two hybrid families were screened for polymorphisms at 14 microsatellite loci which have been described in previous studies. Seven and eight heterozygous markers were identified in the Atlantic salmon and brown trout parents, respectively. Segregation ratios of variant alleles in the hybrid offspring, were normal (1:1) in all but one case. In brown trout, seven significant ($P<0.05$) linkages were detected among microsatellite and RAPD loci, three of which were highly significant ($P<0.001$). In Atlantic salmon, five significant ($P<0.05$) linkages were identified, three of which were highly significant ($P<0.001$). Of three tightly linked microsatellite pairs identified in Atlantic salmon, only two are conserved in brown trout.

3) The partial sequences of 41 ESTs, isolated from two cDNA libraries, were submitted, by e-mail, for comparison with entries in the EMBL and GenBank data bases. Twenty of the sequences were similar to genes of known protein function, 12 of which were used as probes to search for RFLPs in the genomic DNA of Atlantic salmon and brown trout parents. Five different probes detected RFLPs, but multiple banding patterns were difficult to interpret. Variant alleles at a locus detected with the fibrinogen EST segregated normally (1:1) in hybrid offspring. Linkage was not observed between the fibrinogen marker and any of the microsatellite or RAPD loci.

Three and six, highly significant ($P<0.001$), new linkage groups are reported for Atlantic salmon and brown trout, respectively. Linkage analysis using the hybrid offspring of these two species has been a success, however, faster gains may result from more extensive outbreeding prior to hybridization.

ACKNOWLEDGEMENTS

I would like to thank Dr. Willie Davidson, who invited me to Memorial University and has been my supervisor ever since. I'm proud of what I've accomplished since first arriving here, and owe much of that feeling to him. Dr. Gene Herzberg and Dr. Ban Younghusband have sat on my supervisory committee and have provided thoughtful comments on both the project and thesis. Silvia Bartlett has been there for me since the beginning and was especially helpful during the microsatellite analysis. Dorothy Crutcher edited most of the text and will always be a friend indeed. Roy Ficken produced the photographic prints. Vic Drover characterised the ApoAI microsatellite locus, Thorben Bieger helped with the construction and sequencing of the Atlantic salmon cDNA library and Sheilagh Rahal assisted with the characterization of expressed sequence tags. The thesis has also benefited greatly from conversations and collaborations with Kerry Naish and David Skibinsky. My lab mates have been John Goodier, Carolyn Hurst, Terry Young, Mike Woods, Carolyn Walsh and Hughie Broders. Mark Vardy got me stoked. My fellow RAPDers, Barb Saunders and Michelle Normore, have also been a big help. Finally, I would like to thank my parents for all their support over so many years.

TABLE OF CONTENTS

Chapter 3

Application of the RAPD Technique to the Search for a Sex-Linked Genetic Marker in Atlantic Salmon

Chapter 4

Linkage Mapping in an Atlantic Salmon x Brown Trout F1 Hybrid Family Using Microsatellites

Chapter 5

An Evaluation of Expressed Sequence Tags as Non-Ambiguous Markers for Genetic Mapping in Atlantic Salmon and Brown Trout

Chapter 6

Summary

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

EMBL	European Molecular Biology Laboratory
EST	expressed sequence tag
FISH	flourescent *in situ* hybridization
GH	growth hormone
PCR	polymerase chain reaction
QTL	quantitative trait loci
RAPD	randomly amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
SSCP	single strand conformational polymorphism
STS	sequence tagged site
UBC	University of British Columbia
YAC	yeast artificial chromosome

Chapter 1

General Introduction

1.1 Non-Human Genome Mapping

The human genome project represents the upper extreme of ambition and resolution in genetic mapping. Its goal, to map and sequence the estimated 50,000 to 100,000 genes within the genome of *Homo sapiens*, is expected to provide critical information concerning a broad spectrum of questions in human biology (Guyer and Collins 1995; O'Brien et al. 1993). The results of this effort have already provided fresh insights into the identification, cause, treatment and prevention of many genetic disorders in humans (Yates 1996; Guyer and Collins 1995; Gottesman and Collins 1994; Hoffman 1994). It has also brought about several technological advances which have improved the efficiency and speed with which loci can be mapped and has escalated the pace of genome mapping in non-human organisms.

There are several reasons for wanting to map genomes other than our own. Organisms such as *Escherichia coli* (Nilson et al. 1996), *Caenorhabditis elegans* (Hodgkin et al.

1995; Kawaga et al. 1995; Waterston and Sulston 1995; Ogura et al. 1994; Ebert et al. 1993) and *Drosophila melanogaster* (Banfi et al. 1996; Stathakis et al. 1995), provide simple systems which can be easily manipulated for the study of developmental processes, mutation, or gene regulation. The mouse (*Mus musculus*), a more appropriate model for human genetics, plays a significant role in understanding mammalian physiology, genetic disorders and gene structure (Dietrich et al. 1996 and 1994; Melo et al. 1996; Keightley et al. 1996; Rubin and Barsh 1996).

Potential economic benefits have driven researchers to map the genomes of many agriculturally important species. Genetic maps for crops such as maize (*Zea mays*) (Coe and Neuffer 1993; Gardiner et al. 1993), rice (*Oryza sativa*) (Causse et al. 1994; Tanksley et al. 1993) and tomato (*Lycopersicon esculentum*) (Broun and Tanksley 1996; Tanksley 1993), as well as livestock such as cow (*Bos taurus*) (Barendse et al. 1994; Womak et al. 1993), pig (*Sus scrofa domestica*) (Rohrer et al. 1994; Echard et al. 1993) and chicken (*Gallus gallus*) (Crooijmans et al. 1996; Cheng et al. 1995; Bitgood and Somes 1993) are expected to improve the effectiveness of breeding programs through marker assisted selection of desired traits and the mapping of quantitative trait loci (QTLs)

(Whittaker et al. 1995; Lande and Thompson 1990). This involves identifying associations between variable genetic markers and complex phenotypic traits such as growth rate and production (Georges et al. 1995; Kjaer et al. 1995; Andersson et al. 1994), behaviour (Hunt et al. 1995), or life-history characteristics (Mitchell-Olds 1996; Bradshaw et al. 1995). As of 1993, genetic maps of varied resolution have been published for over 100 non-human, eukaryotic and prokaryotic organisms including mammals, birds, reptiles, fish, invertebrates, plants and bacteria (O'Brien 1993).

The development of so many maps for such a wide variety of organisms has given rise to the field of comparative genome mapping. Just as zoologists have developed theories about the evolutionary relevance of particular morphological traits by making comparisons between species, biochemists have taken the same approach to understanding the adaptive significance of genome organization (Koop and Nadeau 1996; O'Brien et al. 1993; Morizot 1990). The growing volume of mapping data has provided numerous examples of conserved linkage (conservation of linkage as well as gene order) and conserved synteny (conservation of linkage regardless of gene order) among diverse taxonomic groups of both plants (Shields 1996; Grivet et al. 1994; Prince et al. 1993; Weeden et al. 1992) and animals (Dubcovsky et al. 1996; Koop and Nadeau 1996; Johansson et al. 1995; Morizot 1990). Extensive conservation

of chromosomal arrangement contrasts with earlier proposals which suggested a reshuffling of linkage groups over substantial evolutionary periods (Morizot 1990). Comparative mapping can also be used to predict candidate genes for genetic disorders in humans by analyzing conserved linkages in smaller, simpler vertebrate genomes such as the pufferfish, *Fugu rubripes* (Koop and Nadeau 1996; Trower et al. 1996).

1.2 Genome Mapping in Fish

Fish represent the most primitive class of vertebrates, and have served as simple models for the investigation of genetic disorders, vertebrate development and genome structure. Platyfish and swordtails of the genus *Xiphophorus* have played an important role in cancer research since the discovery of malignant melanomas which can spontaneously develop in certain hybrids (Schartl 1995; Wellbrock et al. 1995; Fornzler et al. 1991). A genetic map, comprised of 80 loci in 20 linkage groups, has been constructed for *Xiphophorus* using an interspecies backcross strategy (Nairn et al. 1996; Morizot et al. 1993 and 1991; Narine et al. 1992). The majority of these loci are genes which code for electrophoretically detectable proteins; however, several genes associated with melanoma formation and skin pigmentation have also been mapped.

The transparent embryos of zebrafish (*Brachydanio rerio*) have made it a preferred model for the study of vertebrate development (Driever et al. 1994). The ability to produce gynogenetic haploids which proceed normally though early development has simplified the detection of recessive, early-acting mutations which can be artificially induced (Driever et al. 1994; Rossant and Hopkins 1992). Haploid genetics has also advanced the construction of highly saturated genetic maps for the zebrafish. The current linkage map consists of 652 polymerase chain reaction (PCR) based markers covering all of the 25 zebrafish chromosomes (2N=50) and includes centromere map distances. The map incorporates several visible mutations and is expected to facilitate the positional cloning of mutant genes which are involved in embryonic development (Johnson et al. 1996; Kauffman et al. 1995; Postlewait et al. 1994).

The pufferfish (*Fugu rubripes*) genome has very little noncoding DNA and is an estimated one-sixth to one-eighth the size of humans (Elgar et al. 1996; Koop and Nadeau 1996; Brenner et al. 1993). It has become an efficient model for determining the structural organization of large genes which may be difficult to sequence in other vertebrates. Problems associated with the husbandry of *Fugu* have precluded standard linkage analysis and have restricted mapping endeavours to

physical methods such as contig construction, pulse field gel electrophoresis or large scale sequencing (Elgar et al. 1996). Its compact size however, has made it useful for predicting the structure and function of homologous human genes (Maheshwar et al. 1996; Baxendale et al. 1995) and investigating the evolution of multigene families (Venkatesh et al. 1996). It has also generated several examples of linkage conservation between fish and humans (Elgar et al. 1996; Trower et al. 1996). Trower et al. (1996) have used the pufferfish to identify candidate genes for familial Alzheimer's disease in humans. Mapping of fish genomes has also provided information on the conservation of genome organization over long evolutionary periods. Morizot (1990) identified four syntenic groups which have been conserved between different orders of fish species and three which have been maintained over the 900 million years of evolutionary time which separates mammals from fish (450 million years in each lineage) (Koop and Nadeau 1996). Eleven linkages are known to be conserved between humans and the pufferfish (Elgar et al. 1996; Trower et al. 1996). These examples suggest possible constraints on the disruption of certain genomic regions or may be the result of non-random translocation frequencies (Goodier and Davidson 1993; Morizot 1990). Additional mapping of homologous genes is required before any model can be thoroughly tested.

The increasing economic importance of aquaculture has created the desire to map the genomes of fish species which are cultured for food. Progress on species such as tilapia (*Oreochromis niloticus*) (Naish et al. 1995), channel catfish (*Ictalurus punctatus*) (Morizot 1994) and Atlantic salmon (*Salmo salar*) (Lie et al. 1994a and b) has lagged behind that of agricultural species such as pig or chicken; however, recent growth in the consumption of farmed seafood has started to drive technological advances in this area (Ahmed and Anderson 1994). As with agricultural species, the goal has been to obtain maps which can be used to improve the efficiency of breeding programs through marker assisted selection and the mapping of QTLs (O'Reilly and Wright 1995; Ferguson 1994; Lie et al. 1994a; Goodier and Davidson 1993; Thorgaard 1992). So far, the most extensive mapping of economically important species has occurred in the salmonids (the salmons, trouts and chars).

1.3 Gene Mapping in Salmonids

Despite their economic importance, the majority of mapping projects in salmonid fish have been aimed at comprehending the evolutionary implications of genome organization. Observations regarding DNA content, chromosome arm number and duplicate gene loci with a disomic mode of

inheritance, indicate that salmonids have evolved from a tetraploid ancestor and have undergone a process of diploidization (Johnson et al. 1987; Allendorf and Thorgaard 1984; May et al. 1979; Ohno et al. 1970). Chromosomal rearrangements, such as centric fusions and pericentric inversions, have resulted in divergent karyotypes. Diploid chromosome number and arm number vary from ~58 and 74 for Atlantic salmon (*Salmo salar*) to ~60 and 104 for the Pacific species of the genus *Oncorhynchus* and ~80 and 100 for brown trout (*Salmo trutta*), chum salmon (*Oncorhynchus keta*) and *Salvelinus* species such as brook trout (*Salvelinus fontinalis*), lake trout (*Salvelinus namaycush*) and Arctic char (*Salvelinus alpinus*) (Hartley 1987).

Another interesting feature of salmonid genetics has been the observation of nonrandom assortment of segregating loci in male parents, resulting in an excess of recombinant genotypes among progeny (May et al. 1980; Davisson et al. 1973; Morrison 1970). This has been referred to as pseudolinkage and is believed to result from multivalent pairing of homeologous metacentric chromosome arms, coupled with alternate centromeric segregation (Wright et al. 1983). The occurrence of pseudolinkage supports theories of linkage conservation over long evolutionary periods by facilitated translocations

between residual homeologous chromosomes (Morizot 1990). Further gene mapping in this order of fish is expected to improve our understanding of the evolutionary mechanisms behind karyotypic divergence and linkage conservation.

To increase the number of heterozygous loci and simplify the analysis of segregation, many studies have taken advantage of the ability in certain species to form highly heterozygous, fertile hybrids, which can be backcrossed with homozygous parental species (Johnson et al. 1987; May et al. 1980). Others have demonstrated linkage though gene centromere mapping of gynogenetic diploids, produced by activating eggs with sperm that have been irradiated with ultraviolet (UV) light, followed by heat shock treatment to induce retention of the second polar body (Johnson et al. 1987; Allendorf et al. 1986; Seeb and Seeb 1986).

May and Johnson (1993) have combined the data from several independent investigations of linkage in a variety of salmonid species to create a composite linkage map for these fish. It is composed of 55 isozyme loci arranged into 18 classical linkage and 6 pseudolinkage groups. The majority of linkages which have been discovered in salmonids appear to be highly conserved among all species that have been tested (May and Johnson 1993). More polymorphic genetic markers must be developed before highly saturated genetic maps can be constructed for the salmonids. Their tetraploid genome and

phenomena such as pseudolinkage are expected to complicate this effort (Ferguson 1994).

1.4 Gene Mapping in Atlantic Salmon and Brown Trout

Atlantic salmon (*Salmo salar*) and brown trout (*Salmo trutta*) are the two predominant species of the genus *Salmo*. Their importance to both commercial and recreational fisheries has resulted in extensive genetic analysis of wild populations for purposes of stock identification and conservation (McConnell et al. 1995a and b; Wilson et al. 1995; Stephen and McAndrew 1990; Ferguson 1989; Stahl 1987). Genomic analysis in both these species has started to become more relevant. The growing economic significance of Atlantic salmon as an internationally cultured finfish, has created the need for a genetic linkage map which can be used to enhance the selection of health and production related traits (Lie et al. 1994a). From an evolutionary standpoint, the limited sequence divergence observed in the mitochondrial DNA of Atlantic salmon and brown trout (McVeigh et al. 1991; Gyllensten and Wilson 1987) contrasts their disparate karyotypes (Hartley 1987; Section 1.3). Comparative mapping within this genus should provide fresh insight into the structural organisation and evolution of salmonid genomes.

The progress of gene mapping in *Salmo* has been limited.

Only four linkage groups involving isozyme loci have been reported for brown trout (May and Johnson 1993; Johnson et al. 1987; Taggart and Ferguson 1984). Prodohl et al. (1994) observed joint segregation of a minisatellite locus and a sex determining factor in brown trout, and a tight linkage between two minisatellite loci which has been conserved in Atlantic salmon (Taggart et al. 1995). Linkage of any two isozyme loci has not been demonstrated in Atlantic salmon. Some physical mapping of reiterated gene families such as ribosomal RNA or histones, has taken place in both brown trout and Atlantic salmon (Pendas et al. 1995, 1994 and 1993). Further development of fluorescent *in situ* hybridization (FISH) in *Salmo* and its application with single locus probes, is sure to enhance the rate of gene mapping in these species (Phillips and Reed 1996).

There have been two problems with gene mapping in Atlantic salmon and brown trout. The first has been an inability to produce fertile, highly heterozygous hybrids, which can be backcrossed with a homozygous parental species (Johnson and Wright 1986). This strategy has benefitted mapping projects in other salmonids by increasing the number of polymorphic loci which can be screened and simplifying the observation of segregating loci (Johnson et al. 1987). Alternative strategies, such as the production of haploid

gynogens (Lie et al. 1994) or segregation analysis of two haploid genomes simultaneously in hybrid progeny (Kubisiak et al. 1995; Grattapaglia and Sederoff 1994), will have to be tested. A second problem has been a low number of polymorphic enzyme loci which has prevented genetic analysis over broad areas of either genome (Taggart et al. 1995; Ferguson 1994; Davidson et al. 1986). Several projects were undertaken to investigate different techniques and develop new genetic markers which can detect high levels of variability in the genomic DNA of salmonid fish. The following chapters describe some of these techniques.

1.5 Objectives and Approach

The objective of this project was to advance the genomic analysis of Atlantic salmon and brown trout. The approach has been to analyze the segregation of polymorphic genetic markers in the hybrid progeny of a brown trout x Atlantic salmon cross. By using hybrids, the genetic contribution of each parental species can be easily distinguished and treated as separate haploid genomes.

Three types of genetic markers have been tested. Randomly amplified polymorphic DNA (RAPDs) (chapters 2 and 3) and microsatellites (chapter 4) are both PCR based techniques which can furnish the large number of polymorphic loci

necessary for extensive coverage of an entire genome. Expressed sequence tags (ESTs) (chapter 5) are highly informative, non-ambiguous genetic markers which are essential for the comparative analysis of the salmonid genome with the genomes of other fish and other vertebrates.

Chapter 2

Linkage Mapping in a Brown Trout x Atlantic Salmon F1 Hybrid Family Using Randomly Amplified Polymorphic DNA

2.1 Introduction

2.1.1 Randomly amplified polymorphic DNA

Randomly amplified polymorphic DNA (RAPD) analysis involves the amplification of genomic DNA using single primers of arbitrary nucleotide sequence (Williams et al. 1990; Welsh and McClelland 1990). The primers are typically ten nucleotides in length, and are annealed to the template DNA at low temperatures in a low stringency PCR reaction. The technique is expected to sample evenly throughout the entire genome (Lynch and Milligan 1994), with variant loci identified by either the absence of a PCR product or a change in its size (Ewayne-Shoemaker et al. 1994). A locus will fail to amplify if a mutation in the template DNA results in the loss of a priming site. Insertions or deletions within a locus can cause a change in the size of a PCR product, or a loss of that product if the template becomes too large to amplify, or too small to be detected by electrophoresis.

The RAPD analysis is both a practical and economical

method for detecting a large number of DNA polymorphisms. It requires no *a priori* sequence information so that large numbers of unique primers can be synthesized, in bulk, at little expense to the researcher. PCR products can be separated according to size on standard agarose gels and visualised by ethidium bromide staining, thus eliminating much of the time and expense associated with radioactive labelling or probing. Finally, a single RAPD primer can generate several PCR products, each of which can be treated as an individual locus. Consequently, a genome can be screened with hundreds of primers at thousands of loci in a relatively short period of time.

There are certain drawbacks associated with the RAPD procedure. The majority of RAPD markers are dominant, such that variant null alleles cannot be scored in heterozygous individuals. When mapping diploid organisms with RAPDs, it is important to choose a breeding strategy in which the observation of segregating alleles inherited from one parent, is not interfered with by dominant alleles at the same loci, inherited from a second parent (Williams et al. 1991; Sall and Nilsson 1994). This can be accomplished through the analysis of backcross progeny (Lin and Ritland 1996; Peltier et al 1994), haploid gynogens (Postlethwait et al. 1994; Johnson et al. 1996), recombinant inbred lines (Reiter et al. 1992) or

hybrids (Grattapaglia and Sederoff 1994; Kubisiak et al. 1995; Plomion et al. 1995). Several authors have indicated how minor changes, in the parameters of the amplification reaction or the quality of the DNA template, can result in different banding patterns (Bardakci and Skibinski 1994; Devos and Gale 1992; Williams et al. 1993; Weeden et al. 1992; Hadrys et al. 1992). Inconsistency in protocols can make it difficult to transfer information between different projects. Accordingly, marker quality and the robustness of the amplification reaction must always be taken into consideration. Errors can occur when two or more different loci are represented by PCR products of equal size, and cannot be distinguished from one another. This problem is magnified when comparing the results of different species or divergent strains. Rieseberg (1996) tested the sequence homology of comigrating fragments from three closely related species of sunflower (*Helianthus annuus*, *Helianthus petiolaris* and *Helianthus anomalus*) and found that only 79% of loci were suitable for comparative genetic studies. To improve their utility, polymorphic RAPD loci can be converted into sequence tagged sites (STSs) which are amplified using longer primers of greater specificity (Johnson et al. 1996).

2.1.2 Genetic linkage mapping using RAPDs

Regardless of its weaknesses, the RAPD technique has been

a powerful tool for genomic analysis and has been used extensively in the production of genetic maps for several plant species (Peltier et al. 1994; Kubisiak et al. 1995; Lin and Ritland 1996; Grattapaglia and Sederoff 1994; Plomion et al. 1995; Kiss et al. 1993; Whisson et al. 1995). Some of these maps have included morphological traits, such as flower shape or colour (Rajapakse et al. 1995; Chaparro et al. 1994), as well as quantitative trait loci (QTLs) such as disease resistance or cold acclimation (Hombergen and Bachmann 1995; Grattapaglia et al. 1995; Mouzeyar et al. 1995; Cai et al. 1994; Martin et al. 1991). RAPDs have also been used for the marker assisted selection of QTLs (Miklas et al. 1996).

In many cases, genetic linkage mapping with RAPDs has resulted in comprehensive coverage of entire plant genomes. Kubisiak et al. (1995) mapped an estimated 61.7% of the slash-pine (*Pinus elliottii*) genome and 80% of the longleaf pine (*Pinus palustris*) genome through analysis of their hybrid offspring. Similarly, Grattalapaglia and Sederoff (1994) mapped approximately 95% of both the *Eucalyptus grandis* and *Eucalyptus urophylla* genomes using 558 RAPD markers which were generated from only 151 primers.

Despite its frequent application in plant genetics, there are few examples of gene mapping with RAPDs in animals. Promboon et al. (1995) mapped 168 RAPD markers into 29 linkage

groups for the silkworm (*Bombyx mori*) genome. Sixty-eight RAPD loci have been assigned to the genetic map of chicken (Burt et al. 1995). Postlethwait et al. (1994) used 401 RAPD markers, generated from only 134 random primers, to construct a recombination map for the zebrafish (*Danio Brachydanio rerio*) genome. Observation of segregating loci was simplified by using hybrid females to produce haploid gynogen progeny in which the recessive null alleles were not obscured by their dominant alternative. To refine this map, Johnson et al. (1996) converted several RAPD markers into STSs, and used them as anchor loci for aligning several maps constructed from different families. The current zebrafish map consists of 619 RAPD markers, 14 STSs constructed from RAPDs, five STSs located within genes and 14 simple sequence repeats (SSRs). All 25 chromosomes have been mapped including the positions of each centromere.

2.1.3 Objectives

This chapter describes the application of the RAPD technique to linkage analysis in the genus *Salmo*. The approach has been to screen a brown trout x Atlantic salmon F1 hybrid family, where the genetic contribution of each parent can be easily distinguished, with a series of unique primers of arbitrary nucleotide sequence. The goal has been to identify polymorphic loci which can be used to construct

genetic linkage maps for both species. In order to maximise the number of loci screened, efforts were concentrated on a single family.

2.2 Materials and Methods

2.2.1 Brood stock and hybrid families

Adult anadromous brown trout and adult anadromous Atlantic salmon were collected from the North East Placentia River, Newfoundland (47°16'N, 53°50'W). Interspecific hybrid families were produced and reared until first feeding, at which point 50 to 100 fry from each family were killed and individually frozen at - 70°C (See McGowan and Davidson 1992 or McGowan 1992 for more details). Blood samples were taken from the caudal vein of all parents. Red blood cells and plasma were separated by centrifugation then stored frozen at - 70°C.

2.2.2 DNA isolation

DNA isolations were carried out in sterile 1.5 ml Eppendorf microcentrofuge tubes, using the entire body of each hybrid fry, or a 10 μl sample of red blood cells from each parent. Samples were digested overnight at 37°C in 375 μl of DNA extraction buffer (0.5% N-lauroylsarcosine, 0.2 M EDTA) and 25 μl of proteinase K (20 mg/ml). This was followed by a 1 hr incubation at 37°C, with an additional 10 μl of DNAse

free RNAse (10 mg/ml). Three successive organic extractions were performed using 400 μl phenol (Tris saturated, pH 8.0), 400 μl phenol:chloroform:isoamyl alcohol (25:24:1) and 400 μl chloroform:isoamyl alcohol (24:1), respectively. The purified DNA was precipitated with two volumes of cold (-20°C) 95% ethanol and incubated at -20°C for 30 min. After a 20 min centrifugation (10000 x *g*), the supernatant was discarded and the DNA precipitate washed with 100 μl of cold 70% ethanol. This was followed by a 5 min centrifugation (10000 x *g*) and removal of as much ethanol as possible. DNA samples were dried under a vacuum for 15 min, rehydrated at 37°C for 30 min in 200 μl of sterile water, then stored at 4°C. Each sample was quantified using UV (λ = 260 nm) spectrophotometry and a portion diluted to a working concentration of approximately 15 μg/ml. Dilute samples were stored at -20°C.

2.2.3 PCR conditions and electrophoresis

RAPD reactions were carried out in 25 μl volumes of 50 mM KCl, 10 mM Tris-HCl (pH 9), 0.1% Triton X-100, 2 mM MgCl, 0.2 mM each of dATP, dTTP, dCTP and dGTP (Pharmacia Biotech, Baie D'Urfe, PQ), 1 unit of sequencing grade Taq DNA polymerase (Promega Corp., Madison, WI), 1.2 μM of 10-base random primer (Biotechnology Laboratory and Department of Forestry, University of British Columbia, Vancouver, BC; Series 2, 3 and 4) and 25-30 ng of template DNA. A negative control (no

template DNA) was run for each primer.

Reactions were overlaid with a drop of mineral oil and subjected to the following temperature profile: initial denaturation at 95°C for 5 min; 34 cycles of 94°C for 45 sec, 37°C for 45 sec and 72°C for 1 min; final extension at 72°C for 5 min. All reactions were run using a Perkin-Elmer DNA Thermal Cycler (model 480, Perkin-Elmer, Foster City, CA) and once completed, were stored at 4°C.

PCR products were separated according to size using a 1.5% agarose (type I:low EEO, Sigma-Aldrich, Mississauga, ON): 1.5% NuSieve® GTG®, low melting point agarose (FMC BioProducts, Rockland, ME) submersible gel, with a 0.5 X Tris-borate EDTA (TBE) buffer. A mixture of λ DNA, digested with *HindIII*, and ϕ-x-174-RF (replicative form) DNA, digested with *HaeIII*, was included on each gel as a molecular weight marker. A constant voltage of 3.0 V/cm was applied for 4 hours. Gels were stained with ethidium bromide (0.5 μg/ml) for fifteen minutes then rinsed for one hour in running tap water. Banding patterns were visualised with a UV transilluminator and recorded by photography on Polaroid® 665 professional positive/negative black & white film (Polaroid, Cambridge, MA) using a red filter.

2.2.4 Data analysis and mapping

The image of each gel was transferred from photograph to

computer disk using a ScanJet Plus flatbed scanner (Hewlett Pakard, Mississauga, ON). The number and sizes of DNA fragments, generated by each primer, were determined from the digitized image using the PRO-RFLP molecular weight software, version 1.37 (DNA ProScan Inc., Nashville, TN). Only highly reproducible, easily scored polymorphic markers were selected for further evaluation.

Formal locus names for polymorphic RAPD markers follow the scheme used by Johnson et al. (1996) in which the name of the 10 nucleotide long primer is followed by the size of the marker in base pairs and letters indicating the species of origin. For example, the marker UBC215:350:*Str* is amplified by University of British Columbia primer 215, is 350 base pairs in length, and is found in *Salmo trutta* (brown trout). Markers found in Atlantic salmon (*Salmo salar*) are identified by the letters *Ssa*.

In order to maximise the number of loci screened, efforts were concentrated on a single hybrid family (HFJ91) consisting of 36 full-sib hybrid progeny, an Atlantic salmon female parent, and a brown trout male parent. Initial screening for variable loci was conducted on six of the hybrid offspring and the two parents. Segregation frequencies for variant alleles were estimated from the entire family.

Alleles at each locus were tested for departure from a

normal Mendelian segregation ratio (1:1) using a chi-square (χ^2) statistic. When analysing for joint segregation between loci, the genetic contribution from each parent was treated separately, as though it were a haploid genome. Linkage analysis was carried out using a procedure similar to that used in other salmonid linkage studies (May et al. 1979; Taggart and Ferguson 1994; Taggart et al. 1995; Prodohl et al. 1995), using equations originally described by Mather (1951). The four possible genotypes observed in the hybrid offspring were grouped as follows:

a_1 = AB

a_2 = Ab

a_3 = aB

a_4 = ab

A chi-square for independent segregation of alleles at two loci was calculated using the equation:

$$\chi^2 = (a_1-a_2-a_3+a_4)^2/N \quad (df = 1) \qquad (2.1)$$

where N is the total number of progeny tested.

The recombination frequency (r) between linked loci was calculated using the equation:

$$r = (a_1+a_4)/N \quad (2.2)$$

or

$$r = (a_2+a_3)/N \quad (2.3)$$

when alleles at linked loci were in a *trans* or *cis* configuration, respectively.

2.3 Results

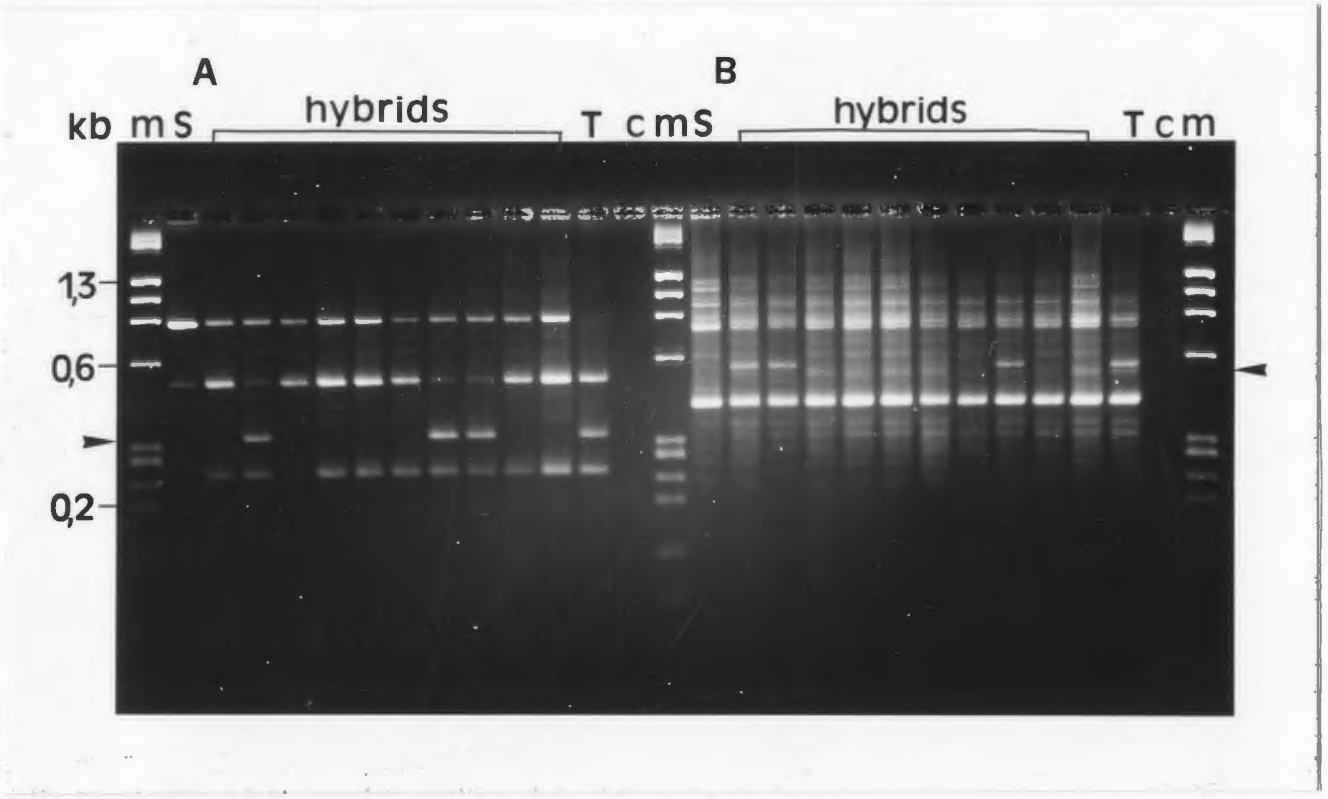
2.3.1 Initial screening of RAPD loci

A total of 271 random primers were tested. Of these, 156 (58%) generated distinct band patterns which were conserved among family members and could be scored. The GC content of primers which worked (average 6.5 out of 10) was significantly higher than for those which failed to amplify (average 5.6 out of 10) ($Z = 6.2$, Mann-Whitney two-tailed test, $P<0.001$, $df = \infty$). This might be due to the greater number of hydrogen bonds, which results in a more stable product during the annealing reaction.

A single primer typically amplified several DNA fragments of varied size and staining intensity (Figure 2.1). Bright bands were highly reproducible and easily scored, whereas faint bands were generally difficult to replicate and were consequently left unrecorded. In the initial screening of hybrid offspring, 893 loci were scored, of which 126 could be

Figure 2.1

Amplification products generated from an Atlantic salmon (S), a brown trout (T) and their hybrid offspring (HFJ91), using random primers UBC226 (A) and UBC273 (B). Arrows indicate dominant polymorphic loci in the brown trout parent which are absent from the Atlantic salmon parent, and can be observed segregating in the hybrid offspring. Control reactions are identified by a **c** and the molecular weight marker (λ DNA, digested with *HindIII*, and ϕ-x-174-RF DNA, digested with *HaeIII*) by an **m**. The sizes of marker bands, in kilo base-pairs (kb), are given on the left.


A
B
kb m S
hybrids
T c m S
hybrids
T c m
1,3
0,6
0,2

attributed to the Atlantic salmon component of the genome and 173 to that of the brown trout (appendix A). The majority (594) of loci were held in common by both parents. A list of those primers which failed to amplify is presented in appendix B.

2.3.2 Identification of variable RAPD loci

Three heterozygous loci were identified in the Atlantic salmon parent (Table 2.1). Each locus was amplified by a different primer. Alleles at one locus (UBC234:778:*Ssa*) were codominant, with hybrid progeny inheriting one of two bands which differed in electrophoretic mobility (Figure 2.2). The other two markers were dominant such that null alleles could only be identified by their absence in some of the hybrid offspring. Alleles at all three loci were inherited in a normal Mendelian fashion with no significant deviations from a 1:1 segregation ratio.

Thirteen heterozygous loci, amplified by 12 different primers, were identified in the brown trout parent (Table 2.2). All of the markers were dominant and all of the alleles segregated in a normal 1:1 manner. One primer (UBC348) amplified two variant markers in brown trout as well as an Atlantic salmon locus (Figure 2.3). The individual genotypes

Table 2.1
Chi-square test for Mendelian (1:1) segregation of alleles in an Atlantic salmon x brown trout hybrid family (HFJ91), at three polymorphic loci inherited from the Atlantic salmon parent.

RAPD marker	Primer sequence (5' to 3')	N	Progeny genotype observed/(expected) A	a	χ^2 (df=1)	Probability
UBC234:778:*Ssa*	TCCACGGACG	35	22(17.5)	13(17.5)	2.32	0.5>P>0.1
UBC348:352:*Ssa*	CACGGCTGCG	35	20(17.5)	15(17.5)	0.72	0.9>P>0.5
UBC368:220:*Ssa*	ACTTGTGCGG	30	15(15.0)	15(15.0)	0.00	P>0.9

Figure 2.2

Amplification products generated from an Atlantic salmon (S), a brown trout (T) and their hybrid offspring (HFJ91), using random primer UBC234. Arrows indicate codominant alleles which differ in molecular weight. Either one of the two bands present in the Atlantic salmon parent, was observed in the hybrid progeny, but never both at the same time. Control reactions are identified by a **c** and the molecular weight marker (λ DNA, digested with *HindIII*, and ϕ-x-174-RF DNA, digested with *HaeIII*) by an **m**. The sizes, in kilo base-pairs, of marker bands are given on the left.

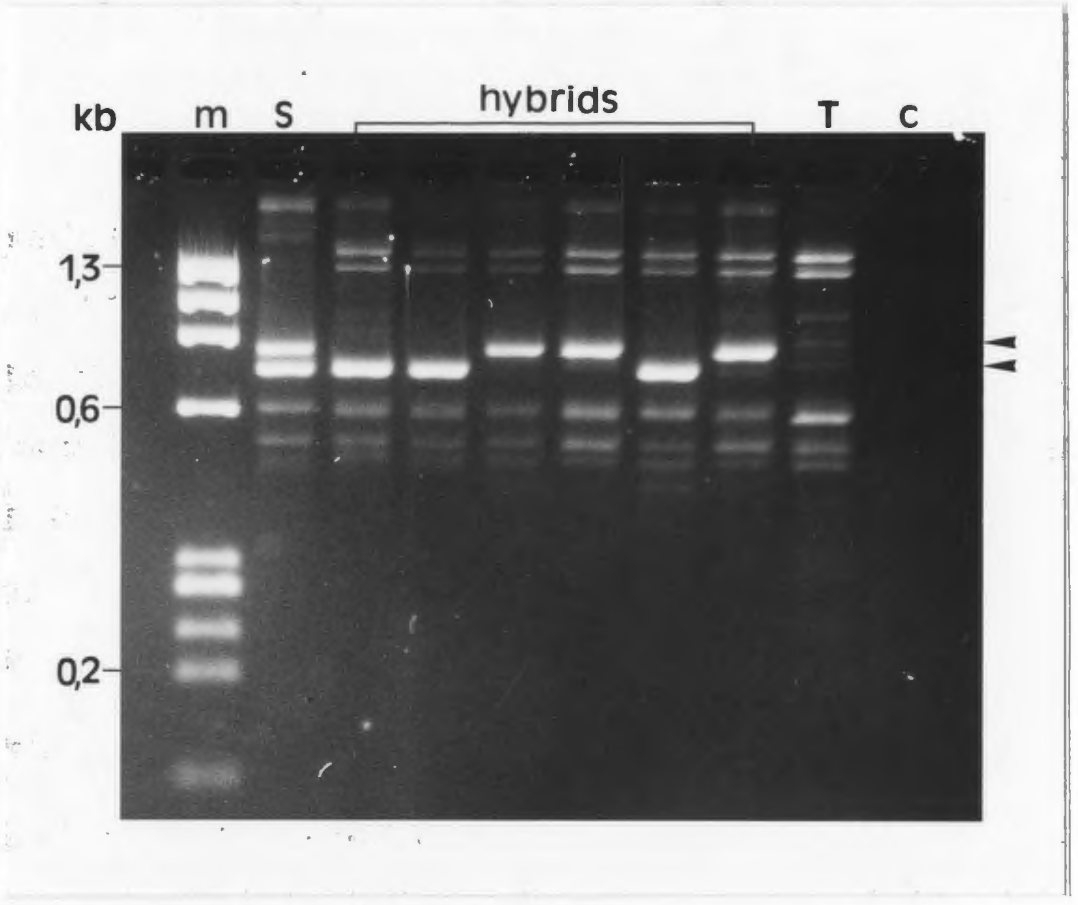
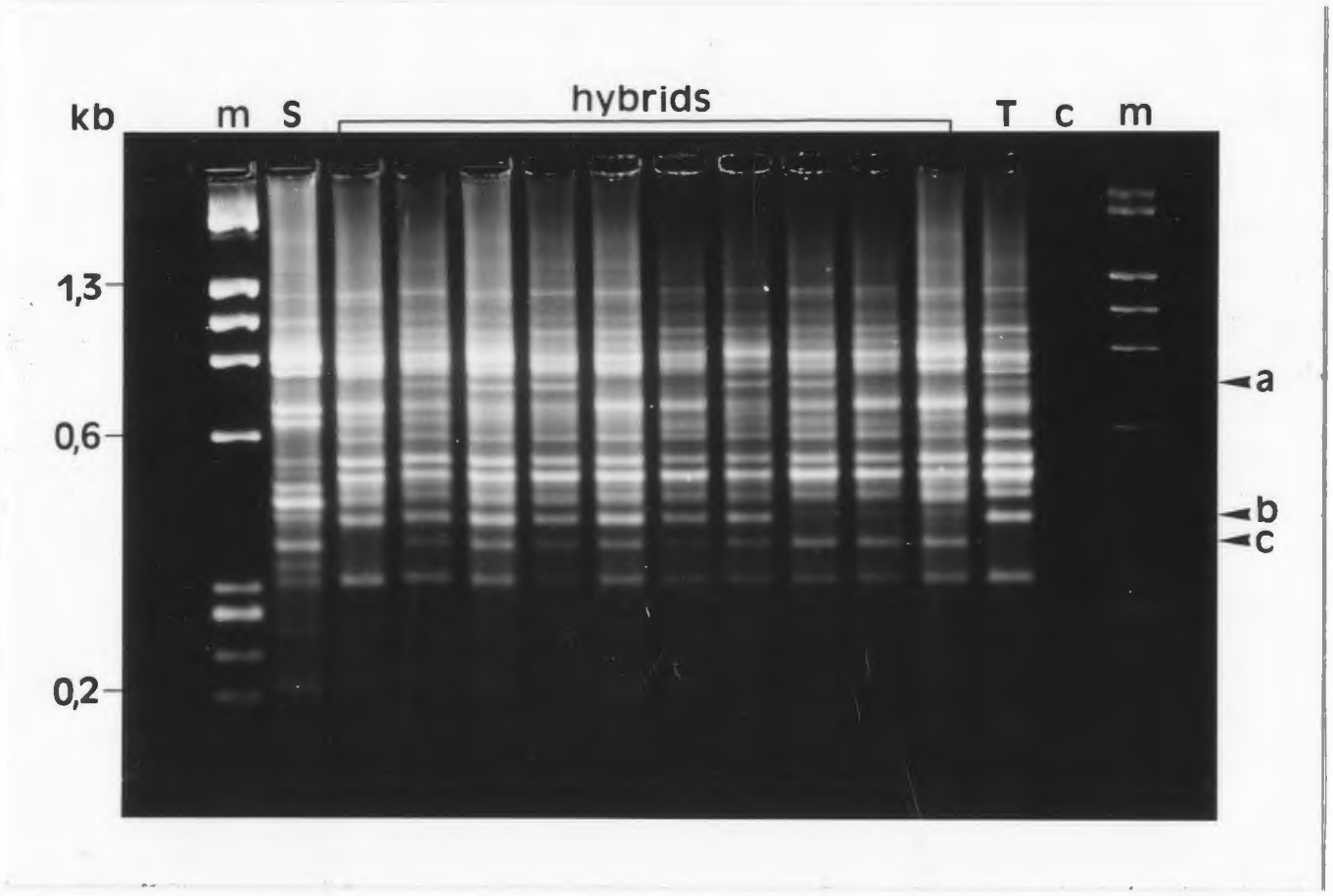

kb
m
S
hybrids
T
c
1,3
0,6
0,2

Table 2.2
Chi-square test for Mendelian (1:1) segregation of alleles in an Atlantic salmon x brown trout hybrid family (HFJ91), at thirteen polymorphic loci inherited from the brown trout parent.

RAPD marker	Primer sequence (5' to 3')	N	Progeny genotype observed (expected) A	a	χ^2 (df=1)	Probability
UBC131:268:*Str*	GAAACAGCGT	30	14(15)	16(15)	0.13	0.9>P>0.5
UBC152:825:*Str*	CGCACCGCAC	32	15(16)	17(16)	0.12	0.9>P>0.5
UBC273:530:*Str*	AATGTCGCCA	36	17(18)	19(18)	0.11	0.9>P>0.5
UBC226:316:*Str*	GGGCCTCTAT	36	15(18)	21(18)	1.0	0.5>P>0.1
UBC239:446:*Str*	CTGAAGCGGA	36	15(18)	21(18)	1.0	0.5>P>0.1
UBC280:281:*Str*	CTGGGAGTGG	36	20(18)	16(18)	0.44	0.9>P>0.5
UBC219:432:*Str*	GTGACCTCAG	33	21(16.5)	12(16.5)	2.4	0.5>P>0.1
UBC243:823:*Str*	GGGTGAACCG	36	17(18)	19(18)	0.11	0.9>P>0.5
UBC348:739:*Str*	CACGGCTGCG	35	16(17.5)	19(17.5)	0.26	0.9>P>0.5
UBC348:399:*Str*	CACGGCTGCG	35	21(17.5)	14(17.5)	1.4	0.5>P>0.1
UBC342:494:*Str*	GACATCCCTC	37	17(17)	17(17)	0.0	P>0.9
UBC354:517:*Str*	CTAGAGGCCG	35	18(17.5)	17(17.5)	0.029	0.9>P>0.5
UBC358:371:*Str*	GGTCAGGCCC	33	14(16.5)	19(16.5)	0.76	0.9>P>0.5

Figure 2.3

Amplification products generated from an Atlantic salmon (S), a brown trout (T) and their hybrid offspring (HFJ91), using random primers UBC348. Arrows **a** and **b** indicate dominant polymorphic markers attributed to the brown trout parent. Arrow **c** indicates a dominant polymorphic marker attributed to the Atlantic salmon parent. A control reaction is identified by a **c** and the molecular weight marker (λ DNA, digested with *HindIII*, and ϕ-x-174-RF DNA, digested with *HaeIII*) by an **m**. The sizes, in kilo base-pairs (kb), of marker bands are given on the left.


kb
m
S
hybrids
T
c
m
1,3
0,6
0,2
a
b
c

of hybrid progeny at RAPD loci are given in appendix C.

2.3.3 Linkage analysis of polymorphic RAPD loci

No linkage was observed between any of the three Atlantic salmon RAPD markers, with alleles at all three loci segregating independently of one another.

Of the 78 pairwise comparisons made between the 13 different brown trout markers, 13 indicated a significant ($P<0.05$) association between segregating alleles (Table 2.3). In five cases, linkage was highly significant ($P<0.001$). One tightly linked grouping involved three different RAPD loci (UBC226:316:*Str*, UBC239:446:*Str* and UBC358:371:*Str*). These markers can be ordered according to their recombinational frequencies, however, the small difference in the actual number of recombinants makes such predictions risky without a larger sample size or additional linkage analysis of other families.

2.4 Discussion

Despite the large number of RAPD loci screened, relatively few informative markers were identified in either Atlantic salmon or brown trout. It is possible that low

Table 2.3
Joint segregation analysis and recombination frequencies (r) for brown trout markers in an Atlantic salmon x brown trout hybrid family (HFJ91). Only marker pairs demonstrating significant (P<0.05) linkage have been listed.

Marker A	Marker B	N	Progeny genotypes a_1	a_2	a_3	a_4	χ^2 (df=1)	r
UBC273:530:*Str*	UBC354:517:*Str*	35	1	15	17	2	24.0***	0.086
UBC226:316:*Str*	UBC239:446:*Str*	36	14	1	1	20	28.4***	0.056
UBC226:316:*Str*	UBC358:371:*Str*	33	13	2	1	17	22.1***	0.091
UBC239:446:*Str*	UBC358:371:*Str*	33	14	1	0	18	29.1***	0.030
UBC243:823:*Str*	UBC348:739:*Str*	35	15	2	1	17	24.0***	0.086
UBC273:350:*Str*	UBC348:739:*Str*	35	12	5	4	14	8.25**	0.26
UBC280:281:*Str*	UBC348:399:*Str*	34	16	3	5	10	9.53**	0.23
UBC219:432:*Str*	UBC243:823:*Str*	33	14	7	2	10	6.8*	0.27
UBC243:823:*Str*	UBC358:371:*Str*	33	11	6	3	13	6.8*	0.27
UBC131:268:*Str*	UBC354:517:*Str*	30	10	4	5	11	4.8*	0.30
UBC273:530:*Str*	UBC243:823:*Str*	36	11	6	6	13	4.0*	0.33
UBC348:739:*Str*	UBC342:494:*Str*	33	11	4	6	12	5.1*	0.30
UBC348:739:*Str*	UBC354:517:*Str*	34	5	10	13	6	4.2*	0.32

* P<0.05 ** P<0.005 *** P<0.001

heterozygosity of the parents in this study is simply a characteristic of the populations from which they were sampled. If so, a greater number of polymorphic loci might be observed if hybrids were produced using strains of fish which have been extensively outbred. The amount of variation that can be detected across the entire range of Atlantic salmon, or brown trout, using the RAPD technique, is still unknown. Consequently, it is difficult to predict the degree of success this additional investment in time and resources would bring.

The majority of PCR products scored in the hybrid offspring were held in common by both the brown trout and Atlantic salmon parents. This restricted the number of loci at which definitive examination of both dominant and recessive (null) alleles could be made. Interference due to bands of similar molecular weight which are inherited from both parents, could be completely avoided by screening haploid progeny in which the genetic contribution of only a single parent is present. Lie et al. (1994b) have demonstrated this process in Atlantic salmon by activating eggs with irradiated sperm. Due to their developmental instability, haploid embryos do not survive beyond the hatching stage. The small amount of tissue generated by this time will limit the number of reactions that can be performed on a particular mapping population. This would be offset by a drastic increase in the number of loci at which an individual's genotype can be

confirmed.

In certain cases, potential markers were not selected for segregation analysis because they could not be consistently reproduced. This problem might be associated with the instability of an annealing reaction involving primers which are only ten nucleotides in length. Conversion of polymorphic RAPD loci into sequence tagged sites (STSs) could be accomplished by sequencing the PCR product, then constructing longer, more specific, primers. Amplification of an STS is likely to be more consistent given a more stable annealing reaction. Polymorphism could be identified not only by the presence or absence of a PCR product, but also as size polymorphims, single stranded conformational polymorphisms (SSCP), or restriction fragment length polymorphisms (RFLPs). Markers might also be less ambiguous, enabling more informative comparisons between different strains of fish or even different species. Transforming RAPD markers into STSs will require a greater technical investment, but the added reliability this would bring would considerably increase the power of these markers.

Of the markers that were acceptable, all demonstrated normal Mendelian segregation ratios in the hybrid offspring, indicating a disomic mode of inheritance, in which hybrids receive half of their genome from each parent. The absence of linkage between markers identified in the Atlantic salmon

parent is not surprising given the small number of loci tested. For the brown trout parent, 13 instances of significant ($P<0.05$) joint segregation were observed, five of which were highly significant ($P<0.001$). Given the large number of pairwise comparisons (78), and the sample size of the mapping population, it is probably safest to accept only the latter, tightly linked, marker-pairs. At $P<0.05$, four out of every 78 comparisons are expected to be significant by chance alone. This probability is reduced to less than one in 78 at $P<0.001$. The loose associations identified in this study will require further analysis, using a larger sample population, before they can be confirmed or rejected.

Chapter 3

Application of the RAPD Technique to the Search for a Sex-Linked Genetic Marker in Atlantic Salmon

3.1 Introduction

3.1.1 A sex-specific marker for Atlantic salmon

The desire to produce female-only populations of Atlantic salmon for aquaculture has generated considerable interest in finding a genetic marker which is tightly linked to the sex determining factor (Devlin et al. 1994 and 1991; Ferguson 1994; Hunter et al. 1982). These female-only populations would eliminate unwanted life-history characteristics associated with male salmon, such as sexual maturation at an early age and the subsequent reduction in growth rate, flesh quality, and market value of the fish (Wild et al. 1994). A genetic marker which can identify the sex of an individual fish would greatly improve the efficiency of distinguishing between genotypic (XX) females which have undergone hormonal sex reversal to become phenotypic males, and their normal (XY) male siblings (Feist et al. 1995; Hunter et al. 1982; Piferrer et al. 1994; Johnstone et al. 1979 and 1978). Once

identified, masculinized females can be crossed with normal females to produce all female offspring for commercial grow-out. Such markers are currently available for chinook salmon (*Oncorhynchus tshawytscha*) and coho salmon (*Oncorhynchus kisutch*) (Devlin et al. 1994 and 1991; Forbes et al. 1994; Du et al. 1993).

Although genotypic female Atlantic salmon can become phenotypic males after treatment with 17α-methyltestosterone (Johnstone and Youngson 1984; Johnstone et al. 1978), verification of masculinized individuals is problematic. Currently, sex-reversed female salmonids can be identified by morphological deformations of the reproductive system, or by performing a test cross with normal, untreated females (Devlin et al. 1994; Johnstone et al. 1979). The former method requires killing the animal and inspecting the gonads before removing them to recover the milt. The latter is time consuming and involves the maintenance of additional tanks which must be taken out of production. Both procedures involve the expense of growing both male and female treated fish to sexual maturity. A genetic marker which can unambiguously determine the sex of an individual Atlantic salmon at an early age, requiring only a small, non-invasive tissue sample, such as a piece of fin, would be economically advantageous to the aquaculture industry.

3.1.2 Evidence of male heterogamety in salmonids

Male heterogamety in salmonid fish was first established through the investigation of sex reversal in rainbow trout (*Oncorhynchus mykiss*) (Johnstone et al. 1979). In these experiments, females were masculinized through dietary administration of the androgen 17α-methyltestosterone during the first 90 days of juvenile development; the period in which gonadal differentiation occurs (Foyle 1993). When grown to maturity, sperm from the phenotypic males was used to fertilise eggs from normal, untreated females. In one case, the resulting offspring were all female, demonstrating that both parents had been homogametic and that female rainbow trout were the homogametic sex. Males used in the other crosses were all normal and heterogametic.

In similar experiments with coho salmon (*Oncorhynchus kisutch*) (Hunter et al. 1982) and rainbow trout (Chevassus et al. 1988), normal males were crossed with males that had been feminized through treatment with 17β-estradiol. The resulting families had a 3:1 male to female sex ratio indicating male heterogamety as well as the viability of YY individuals.

Further confirmation of male heterogamety in salmonids has been the description of sex chromosomes in the karyotypes of several species. Thorgaard (1983 and 1977) found that in some populations of rainbow trout, males possessed a

submetacentric Y chromosome with a shorter minor arm than the homologous X chromosome. This difference, however, was not universal among populations and was interpreted as a morphological divergence between the sex chromosomes which has not become fixed in the species.

In sockeye salmon (*Oncorhynchus nerka*), females have 58 chromosomes whereas males have only 57 (both sexes have 104 chromosome arms) (Thorgaard 1978). Here, it has been suggested that a telocentric Y chromosome has fused with a telocentric autosome to form a metacentric chromosome.

Phillips and Ihssen (1985) have reported differences in the quinacrine banding and C-banding patterns of presumed sex chromosomes in lake trout (*Salvelinus namaycush*).

3.1.3 Sex-linked genetic markers in salmonids

A variety of different sex-linked genetic markers have been reported in several salmonid species. Linkage groups involving electrophoretically detectable enzyme loci and a sex determining factor have been observed in rainbow trout (Allendorf et al. 1993; Gellman et al. 1987) and Arctic char (*Salvelinus alpinus*) (May et al. 1989); however, these loci (hexosaminidase in rainbow trout; lactate dehydrogenase and aspartate amino transferase in Arctic char) still undergo some recombination with the sex factor and do not demonstrate sex-linkage in other species of salmonid. In the case of Arctic

char, May et al. (1989) have speculated that these loci lie across the centromere from the sex factor which resides on an information-depauperate chromosome arm and that linkage is the result of a centric fusion unique to Arctic char. In Atlantic salmon, there is evidence to suggest that a gene which regulates the expression of phosphoglucose mutase (*PGM-1r**) is sex linked or under strong sex-specific selection (Pollard et al. 1994), but this association appears to be population specific and recombination between this locus and the sex determining factor can still occur.

Genetic probes associated with the sex chromosomes of other animals, such as the banded krait minor (Bkm) minisatellite from the snake *Bungarus fasciatus* or the *ZFY* and *SRY* genes from mammals, have been tested on a variety of fish species (channel catfish, *Ictalurus punctatus*, Tiersch et al. 1992; *Poecilia* and *Xiphophorus* spp., Nanda et al. 1992; coral reef fish, *Anthias squamipinnis*, Wachtel et al. 1991; sturgeon, *Acipenser transmontanus*, Ferreiro et al. 1989; rainbow trout, Lloyd et al. 1989, Ferreiro et al. 1989). With the exception of *Poecilia reticulata* (guppies) (Nanda et al. 1992), these probes have failed to detect sex-specific loci. Prodohl et al. (1994) isolated several minisatellite probes from the brown trout (*Salmo trutta*) genome, one of which

(p*Str*-A9) was found to be tightly linked with sex (recombination frequency 0.023). Although this probe detected two variable loci in Atlantic salmon, neither was sex-linked (Taggart et al. 1995).

In chinook and coho salmon, an alternate form of the growth hormone (GH) gene is male-specific and appears to be located on the Y chromosome (Forbes et al. 1994; Du et al. 1993). In this case, linkage is complete with no recombination occurring between the sex determining factor and the variant locus. Two different growth hormone genes (GH-1 and GH-2) have been characterised in salmonids (Devlin 1993). Forbes et al. (1994) have suggested that GH-2 is Y-linked and can be detected by a restriction fragment length polymorphism in the PCR product of the third intron. Du et al. (1993) have demonstrated that a third, non-functional pseudogene is found on the Y chromosome and is easily identified by using PCR to detect a 509 base pair deletion in the fifth intron. In either case, amplification of these markers allows rapid, unambiguous identification of sex in both these species. Unfortunately, variant growth hormone genes have not been observed in other salmonids (Du et al. 1993).

Devlin et al. (1991) used subtractive hybridization to develop a male specific probe for chinook salmon (*Oncorhynchus tshawytscha*). In this procedure, genomic DNA from a male was

cut with a restriction enzyme, denatured and then annealed with an excess of sheared genomic DNA from a female. Only male specific DNA could reanneal and retain the appropriate restriction sites at each end of the fragment to permit ligation into a plasmid vector. The resulting library produced a single clone which could be used as a probe to identify male DNA from a chinook salmon. Further characterization of the probe has resulted in the design of oligonucleotide primers which can be used to generate a male-specific PCR product and a simple test for Y chromosome DNA (Devlin et al. 1994). This procedure has also been attempted using rainbow trout (Nakayama et al. 1994) and Atlantic salmon (B. Devlin, pers comm.), but without success.

3.1.4 A RAPD approach to finding sex-linked markers

Despite some success with other salmonid species, a sex specific marker has yet to be found for Atlantic salmon. One approach, which has been applied with species of plants and birds, is to screen for a sex specific marker using randomly amplified polymorphic DNAs (RAPDs). Once identified, sex specific RAPD markers can be isolated and cloned, then either used directly as a probe for Southern blot hybridization, or sequenced so that longer, more discrete primers can be designed for PCR (Sabo et al. 1994; Griffiths and Tiwari 1993). Hormaza et al. (1994) tested 700 decamer primers and

scored over 4900 loci before finding a sex-linked amplification product in the pistachio plant (*Pistacia vera*). Griffiths and Tiwari (1993) found sex-specific markers for three species of birds, the great tit (*Parus major*), the jackdaw (*Corvus monedula*) and the zebra finch (*Taenopygia guttata*), by screening male and female DNA with 26 different primers using low stringency PCR. Sabo et al. (1994) found six sex-linked markers, two of which were sex-specific, after screening roseate terns (*Sterna dougallii*) at 1400 RAPD loci using 180 different primers. One of the sex-specific markers was cloned and sequenced so that longer, more specific, primers could be designed. These primers were then used to amplify the sex specific marker in both the roseate tern and the closely related common tern (*Sterna hirundo*). In all cases, bulked segregant analysis was performed on separate pools of DNA sampled from several males and several females. This was done to minimize the effects of individual variation and improve the efficiency of the screening process (Michelmore et al. 1991).

3.1.5 Objectives

This chapter describes the application of the RAPD procedure and bulked segregant analysis in the search for a sex-specific marker in Atlantic salmon.

3.2 Materials and Methods

Three adult male and eight adult female Atlantic salmon were collected from the North East Placentia River, Newfoundland (47°16'N, 53°50'W). DNA was extracted from red blood cells and quantified according to procedures given in Section 2.2.2. An equal portion of DNA from each male and each female was pooled into two bulked samples of male and female DNA. These were used for the initial screening of primers. If male specific polymorphisms were detected in the bulked DNA, further analysis was conducted using independent DNA samples to see if a marker was sex-specific or the result of individual variation.

Primers, from UBC primer series 2, 3 and 4, which had been previously tested, and were known to give reproducible amplification products, were used to screen for male specific polymorphisms. An additional 100 primers, from UBC primer series 5, were also tried, but without the *a priori* knowledge of whether or not they would work. Amplification conditions and electrophoresis have been described previously in Section 2.2.3.

3.3 Results

A total of 1152 loci were screened using 200 different primers. Bulked segregate analysis was an efficient means of

screening for male-specific amplification products. Several primers could be tested using the same gel, increasing the rate at which loci could be screened (Figure 3.1). Twenty-five of the primers from UBC primer series 5 gave no amplification product and the remaining 75 primers from this series were not tested for reproducibility of banding patterns.

Four primers, UBC134 (primer #11 in Figure 3.1), UBC269, UBC422 and UBC499, each produced single bands that were found only in the male bulked DNA and did not occur in that of the females (Figure 3.1). Further analysis of individual male and female salmon demonstrated that these loci were in fact variable and reproducible, but were the result of individual male variants and were not male-specific (Figure 3.2). Consequently, of the 1152 loci scored, none appear to be tightly linked to a sex determining factor.

3.4 Discussion

Despite the large number of RAPD loci screened, a tightly sex-linked genetic marker could not be identified. As with many other of the lower vertebrates (Nakamura 1989), it is possible that Atlantic salmon sex chromosomes are simply autosomes on which the sex determining factor is located. A general lack of heteromorphism in fish suggests that sex

Figure 3.1

RAPD patterns generated using male and female bulked DNA. Primers were tested for polymorphisms specific to the male DNA, before analysis of independent DNA samples. Numbers 1 through 15 each represent a single primer used to amplify the male DNA sample (left) and the female DNA sample (right). The arrow indicates an amplification product generated by primer 11 (UBC134) which is specific to the male DNA. Based on this result, further analysis was conducted using independent DNA samples, to see if the band was sex-specific or the result of individual variation.

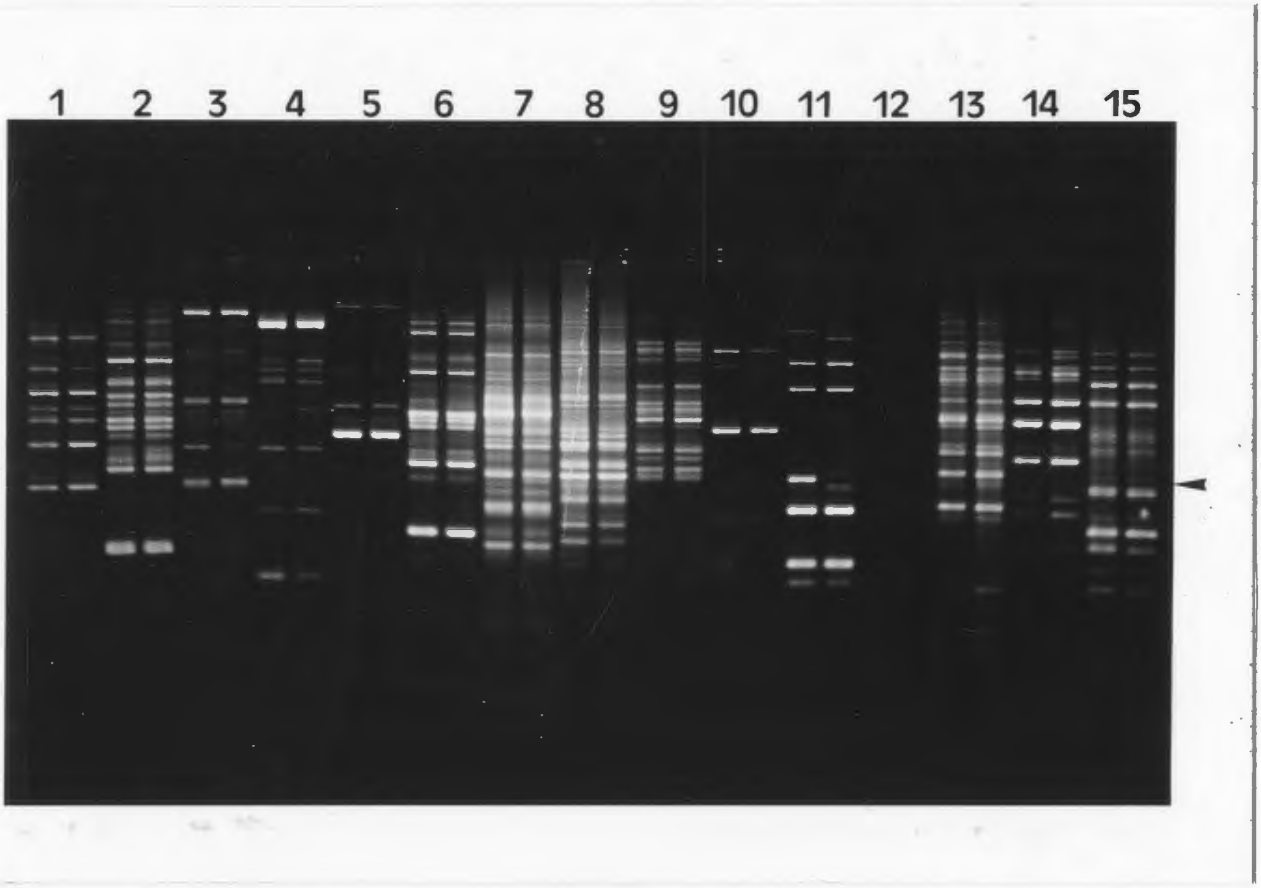
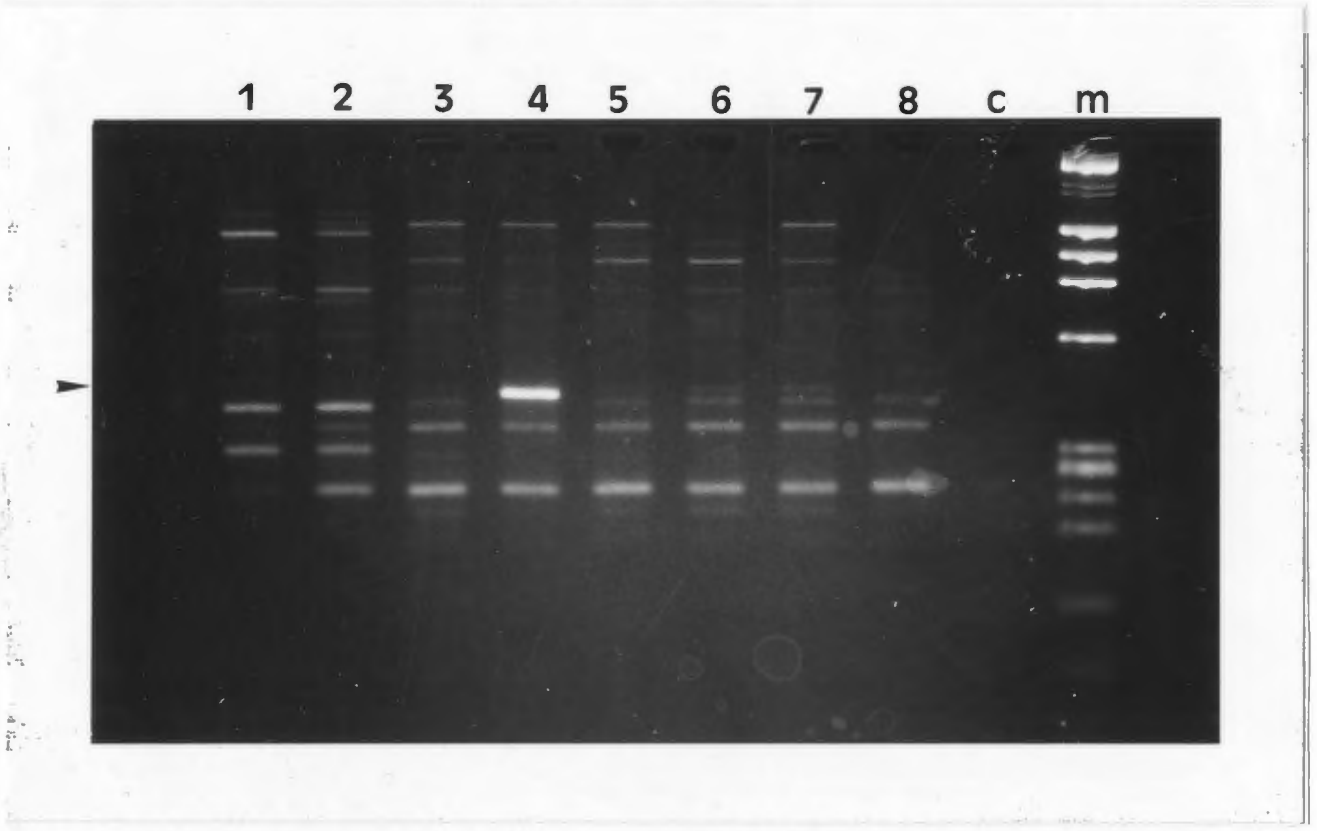

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

Figure 3.2

RAPD patterns generated by primer UBC134 using independent samples of Atlantic salmon male (3 to 5) and female (6 to 8) DNA. Lanes 1 and 2 are a male and female brown trout respectively. An arrow indicates the position of an amplification product which was male-specific for bulked DNA (Figure 3.1). Here, the band is present only for individual 4 and is absent from all other males and females. A control reaction is identified by a **c** and a molecular weight marker (λ DNA, digested with *HindIII*, and ϕ-x-174-RF DNA, digested with *HaeIII*) by an **m**.


1
2
3
4
5
6
7
8
c
m

chromosomes, where they exist, have evolved very recently (Francis 1992). Sequence divergence between the sex chromosomes is believed to result from isolation of the two homologues due to the obstruction of crossover by inversion at or near the sex determining factor (Schimid and Haaf 1989). The absence of sex specific markers in Atlantic salmon indicates that either an insufficient number of markers have been tested, or the homologous chromosomes on which the sex determining factor is found, are not extensively differentiated.

The RAPD analysis, of bulked male and female DNA, was an effective means of screening for sex-linkage at a large number of genetic loci. Other studies which have taken a similar approach to this problem, have succeeded only after extensive screening. For example, Hormaza et al. (1994) scored over 4900 loci before establishing a suitable marker for the unambiguous identification of sex in pistachio. Consequently, it may be that a much greater number of loci will have to be screened before a sex-specific marker can be found for Atlantic salmon.

Chapter 4

Linkage Mapping in an Atlantic Salmon x Brown Trout F1 Hybrid Family Using Microsatellites

4.1 Introduction

4.1.1 Microsatellite DNA

Microsatellites are tandemly repeated, short (two to five nucleotides), simple sequences, which are widely dispersed throughout the genomes of all eukaryotes (Hearne et al. 1992; Rassman et al. 1991; Weber 1990; Tautz 1989; Weber and May 1989; Hamada et al. 1982). Their function is still not fully understood, though in some cases, they have been associated with the regulation of gene expression (Stallings 1995; Lu et al. 1993; Naylor and Clark 1990). They are usually less than 300 base pairs in length (O'Reilly and Wright 1995; Rassman et al. 1991) and have been classified by Weber (1990) into three catagories. 1) Perfect microsatellites consist of repetitive elements which are not interrupted throughout their length, and are flanked on either side by non-repetitive sequences. 2) Imperfect microsatellites involve two or more perfect microsatellites which are separated by no more than three

consecutive non-repeat bases. Terminal runs must be at least three full repeats in length and internal runs must be at least 1.5 repeats long. 3) Compound microsatellites are comprised of two different types of repeats, such as $(GT)_n$ and $(CT)_n$, which are separated by no more than three non-repeat nucleotides. The second sequence must be at least five repeats in length, or ≥10 mononucleotides. Compound microsatellites can also be subclassified as perfect or imperfect depending on the structure of the repetitive elements.

The spacing of microsatellites throughout eukaryotic genomes is thought to be random and even; however, in-situ hybridization experiments have demonstrated decreased representation near centromeric and telomeric regions of the chromosome (Wintero et al. 1992; Wong et al. 1990). Their frequency may vary depending on the type of microsatellite, or the organism from which they are sampled. For example, perfect microsatellites tend to be more common than imperfect or compound microsatellites (Brooker et al. 1994; Ostrander et al. 1993; Slettan et al. 1993; Wintero et al. 1992; Weber 1990), tetranucleotide repeats occur less often than dinucleotides and $(GT)_n$ motifs tend to be more common than $(CT)_n$ sequences (O'Reilly and Wright 1995; Estoup et al. 1993; Beckman and Weber 1992). Hamada et al. (1982), using hybridization analysis, estimated the frequency of $(GT)_n$

microsatellites in salmon DNA to be twice that of humans, three times that of yeast, and eight times that of chicken. Rubinsztein et al. (1995a and b) found that microsatellite loci in humans were longer than their homologues in other primates and Brooker et al. (1994) have observed longer repeats in fish when compared with mammalian microsatellites. The evolutionary significance of these differences, if any, is still unknown.

Variability in the length of a microsatellite is common and tends to be positively correlated with the number of repeats at a particular locus (Weber 1990). Changes in size result from the loss or the gain of one or more repeated sequence. The process by which these mutations arise is poorly understood. DNA polymerase slippage during chromosomal replication or repair (Rubinsztein et al. 1995a; Richards and Sutherland 1994; Tautz 1989; Levinson and Gutman 1987) and unequal exchange of DNA between sister chromatids (O'Reilly and Wright 1995; Mahtani and Willard 1993; Weber and Wong 1993), have been proposed as possible mechanisms for change.

As genetic markers, microsatellites are both efficient and informative. Mutations in repeat number can be detected as size differences in the PCR products of loci which are amplified using primer sequences specific to the non-repetitive flanking regions of the microsatellite. Amplification products are typically separated according to

size in agarose or acrylamide gels. In some cases, nonamplifying, null alleles can result from mutations at the priming site (Pemberton et al. 1995). In addition to detecting high levels of variability, amplification of microsatellite loci is highly reproducible and in many cases, homologous loci can be found in closely related species (Rubinsztein et al. 1995a and b; Stallings 1995; Vaiman et al. 1995; Schlotterer 1991).

One drawback of working with microsatellites is the time and cost associated with developing primers to amplify specific loci. Before it can be used as a molecular marker, a microsatellite must first be cloned, isolated from a library, and sequenced. Once primers have been designed, amplification conditions for PCR must be worked out and the level of polymorphism assessed. This initial investment, however, is balanced by the information and efficiency these markers can provide.

4.1.2. Genetic linkage mapping using microsatellites

The discovery of microsatellites has led to the rapid refinement of gene maps in a number of organisms. Thousands of loci have been characterised and mapped in the human genome and many dozens of genes have been added to the map through the use of these markers (Dib et al. 1996; Dubovsky et al. 1995; Gyapay et al. 1994). Microsatellites have been used

extensively in the production of low resolution maps for the genomes of economically important animals such as sheep (*Ovis aries*) (Crawford et al. 1995), cow (Barendse et al. 1994), pig (Rohrer et al. 1994) and chicken (Cheng et al. 1995; Burt et al. 1995) as well as scientifically important organisms such as rat (*Rattus norvegicus*) (Maihara et al. 1995; Serikawa et al. 1992) and mouse (Dietrich et al. 1996 and 1994).

4.1.3. Microsatellites in fish

Microsatellites have been isolated and characterised in a variety of fish species including Atlantic salmon (*Salmo salar*) (Sanchez et al. 1996; Slettan et al. 1995 and 1993; McConnell et al. 1995a), brown trout (*Salmo trutta*) (Estoup et al. 1993), rainbow trout (*Oncorhynchus mykiss*) (Morris et al. 1996; Sakamoto et al. 1994), sockeye salmon (Scribner et al. 1996), chinook salmon (*Oncorhynchus tshawytscha*) (Leung et al. 1994), brook trout (*Salvelinus fontinalis*) (Angers et al. 1995), Atlantic cod (*Gadus morhua*) (Brooker et al. 1994), zebrafish (*Brachydanio rerio*) (Goff et al. 1992), goldfish (*Carassius auratus*) (Zheng et al. 1995), bluegill sunfish (*Lepomis macrochirus*) (Colbourne et al. 1996) and three spine stickleback (*Gasterosteus aculeatus*) (Rico et al. 1993). They are common in fish genomes, with the frequency of $(GT)_n$

microsatellites ranging from one every 7 kilobase-pairs (kb) (Atlantic cod) to one every 23 kb (brown trout), depending on the species (O'Reilly and Wright 1995). Microsatellites tend to be longer in fish than in mammals (Colbourne et al. 1996; McConnell et al. 1995a; Brooker et al. 1994) and primers developed for one species can sometimes be used to amplify presumably homologous loci in closely related fish species (Scribner et al. 1996; Angers and Bernatchez 1996; Colbourne et al. 1996; Morris et al. 1996; Rico et al. 1996; McConnell et al. 1995a; Zheng et al. 1995; Estoup et al. 1993). As with other animals, microsatellites in fish are highly polymorphic making them useful markers for population analysis (Sanchez et al. 1996; McConnell et al. 1995a and b; Angers et al. 1995; Tessier et al. 1995; Estoup et al. 1993), behavioural investigations (Zheng et al. 1995; Colbourne et al. 1996; Rico et al. 1993) and genetic mapping (Estoup et al. 1993; Kauffman et al. 1995; Postlethwait et al. 1994).

4.1.4. Gene mapping in fish with microsatellites

There are still few examples of gene mapping in fish using microsatellites. Postlethwait et al. (1994) included 13 microsatellite loci in a recombination map for the zebrafish *Danio (Brachydanio) reri*, which had been constructed primarily with RAPD markers. Also in zebrafish, Kauffman et al. (1995) determined the centromere-marker map distances for ten

microsatellite loci. Estoup et al. (1993) estimated gene-centromere distances for three microsatellite loci in brown trout using gynogenetic lines. In Atlantic salmon, segregation analysis of microsatellite markers has demonstrated normal mendelian inheritance, however linkage between loci has yet to be observed (McConnell et al. 1995b; Slettan et al. 1995).

4.1.5. Objectives

This chapter describes the application of microsatellite markers to linkage analysis in the genus *Salmo*. Hybrid families were screened for polymorphisms at several microsatellite loci which have been described in previous studies (Drover 1996; Sanchez et al. 1996; McConnell et al. 1995a; Slettan et al. 1995; Estoup et al. 1993). The goal has been to find linkage between different microsatellite loci, or between microsatellite loci and other molecular markers.

4.2 Materials and Methods

4.2.1 Brood stock, hybrid families, and DNA isolation

Collection of brood stock and the production of hybrid families have been described in Section 2.2.1. Procedures for the isolation of DNA have been provided in Section 2.2.2.

4.2.2 PCR amplification of microsatellite loci

Hybrid families were screened for polymorphisms at fourteen microsatellite loci (Table 4.1). Primer sequences and the appropriate annealing temperatures for PCR have been provided by the authors. The microsatellite locus decribed by Drover (1996), is located within the second intron of the apolipoprotein gene in *Salmo*. The others, have no known association with any functional gene.

With the exception of annealing temperatures, amplification conditions for all microsatellite loci were the same. Reactions were carried out in 12.5 μl of 20 mM Tris-acetate (pH 9), 10 mM ammonium sulfate, 75 mM potassium acetate, 0.05% Tween-20, 1.5 mM magnesium sulfate, 0.2 mM each of dATP, dTTP, dCTP and dGTP (Pharmacia), 0.5 unit of *Tfl* DNA polymerase (Promega), 0.8 μM of each primer, 0.1 μM of one primer, 5'-endlabelled with [γ-^{32}P]ATP (Amersham Life Science, Oakville, ON) using polynucleotide kinase, and 30 to 100 ng of template DNA. The following temperature profile was used: initial denaturation for 2 min at 95°C, followed by 30 cycles of 30 sec at 94°C, 30 sec at a primer-specific annealing temperature (see Table 4.1), and 30 sec at 72°C. All reactions were run using thin-walled tubes in a GeneAmp® PCR System 9600 thermal cycler (Perkin-Elmer). Temperature changes were carried out at the top speed of the thermal

Table 4.1
Primer sequences and repeat motif for microsatellite loci which have been described previously in brown trout or Atlantic salmon. Annealing temperatures are those recommended by the cited authors.

locus name	repeat motif	primer sequence (5'-3')	anneal. temp.(°C)	source species	Refrence
μF43	$(GT)_n$	AGCGGCATAACGTGCTGTCT GAGTCACTCAAAGTGAGCCC	60	*Salmo salar*	Sanchez et al. 1996
μ20,19	$(GT)_n$	TCAACCTGGTCTGCTTCGAC CTAGTTTCCCCAGCACAGCC	62		
μD30	$(CT)_n$	AGCAGTAAAGAGAGAGACTG TGTTGACTTCCTTCCCCAAG	53		
Ssa4	$(GT)_n$	ATTAGGCAGCAGCAGGCTGC TGTTCACTCACTGACACGCG	65	*Salmo salar*	McConnell et al. 1995a
Ssa14	$(CT)/(GT)_n$	CCTTTTGACAGATTTAGGATTTC CAAACCAAACATACCTAAAGCC	57		
Ssa289	$(GT)_n$	CTTTACAAATAGACAGACT TCATACAGTCACTATCATC	46		
SSOSL25	$(GT)_n$	ATCTACACAGCTCCTGGTGGCAG CATGTAATGGGTCGAGAGAAGTG	58	*Salmo salar*	Slettan et al. 1995
SSOSL85	$(GT)_n$	TGTGGATTTTTGTATTATGTTA ATACATTTCCTCCTCATTCAGT	55		
SSOSL311	$(GT)_n$	TAGATAATGGAGGAACTGCATTCT CATGCTTCATAAGAAAAAGATTGT	55		
SSOSL417	$(GT)_n$	TTGTTCAGTGTATATGTGTCCCAT GATCTTCACTGCCACCTTATGACC	53		

Table 4.1 Continued

μ15	$(GT)_n$	TGCAGGCAGACGGATCAGGC AATCCTCTACGTAAGGGATTTGC	58	*Salmo trutta*	Estoup et al. 1993
μ60	$(CT)_n$	CGGTGTGCTTGTCAGGTTTC GTCAAGTCAGCAAGCCTCAC	58		
μ73	$(GT)_n$	CCTGGAGATCCTCCAGCAGGA CTATTCTGCTTGTAACTAGACCTA	58		
ApoAI	$(GT)_n$	AGGTTGTCAAGGCTCTGGGA CCTTCTGGATGACACAGAGT	55	*Salmo trutta*	Drover 1996

cycler and completed reactions were stored at 5°C. A control(no template DNA) was run with each set of reactions.

An aliquot of each product was mixed with an equal volume of loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF) denatured at 85°C for 2 min, then subjected to electrophoresis through a 6% polyacrylamide sequencing gel (19:1 acrylamide:bis-acrylamide, 7M Urea, 1X TBE) at a constant power of 35 watts. A control M13 sequencing reaction was included on each gel as a molecular weight marker. Gels were dried without fixing, then exposed to X-ray film for approximately two days before developing.

4.2.3 Data analysis and mapping

Initial analysis of heterozygous loci was conducted using the same hybrid family described in Section 2.2.4 (HFJ91). This family was selected because it had already been screened for polymorphic RAPD markers, and could be used to test for linkage between variable microsatellite and RAPD loci. Additional analysis was performed on hybrid family HFA192 and its parents when testing for conservation of microsatellite linkage groups between brown trout and Atlantic salmon.

Tests for departure from a normal (1:1) Mendelian segregation ratio, and for linkage between loci, have been described in Section 2.2.4.

4.3 Results

4.3.1 Identification of variable microsatellite loci

Of the 14 microsatellite loci for which the hybrid family HFJ91 was tested, seven were heterozygous in the Atlantic salmon parent, and five were heterozygous in the brown trout parent (Table 4.2). Three additional polymorphic loci were identified in the brown trout parent of hybrid family HFA192, bringing to eight, the number of markers analysed in this species. Alleles at all loci were codominant and polymorphisms were identified by differences in the size of amplification products (Figure 4.1). At each locus, the Atlantic salmon and brown trout alleles differed in size, and the genetic contibution of each parent was easily identified in the hybrid offspring (Figure 4.2). Variant alleles were also observed at the ApoAI locus, a microsatellite which is associated with the second intron of the apolipoprotein AI gene in *Salmo* (Figure 4.3).

Segregation ratios of variant alleles in the hybrid offspring, were normal (1:1) in all but one case. In hybrid family HFA192, a significant (χ^2 = 4.23, $P<0.05$) deviation from normal was observed at locus μ60 when inherited from the brown trout parent. This locus, however, segregated normally (χ^2 = 0, $P>0.9$) in the same family, when inherited from the Atlantic salmon parent. The individual genotypes of hybrid

Table 4.2
Chi-square test for Mendelian (1:1) segregation of alleles at heterozygous microsatellite loci which were identified in Atlantic salmon x brown trout hybrid families HFJ91 and HFA192.

Parent	Hybrid Family	Locus	N	progeny genotype observed(expected) A	a	χ^2 (df=1)	probability
Salmon	HFJ91	Ssa4	36	19(18)	17(18)	0.11	0.9>P>0.5
		SSOSL85	32	12(16)	20(16)	2.0	0.5>P>0.1
		SSOSL417	33	33(16.5)	19(16.5)	0.76	0.5>P>0.1
		SSOSL311	32	11(16)	21(16)	3.1	0.1>P>0.05
		μ60	35	20(17.5)	15(17.5)	0.71	0.5>P>0.1
		μF43	35	19(17.5)	16(17.5)	0.26	0.9>P>0.5
		ApoAI	34	22(17)	12(17)	2.9	0.1>P>0.05
	HFA192	SSOSL85	33	15(16.5)	18(16.5)	0.14	0.9>P>0.5
		μ60	34	17(17)	17(17)	0.0	P>0.9
		μF43	34	17(17)	17(17)	0.0	P>0.9
		ApoAI	30	15(15)	15(15)	0.0	P>0.9
Trout	HFJ91	SSOSL311	32	18(16)	14(18)	0.5	0.5>P>0.1
		SSOSL417	33	17(16.5)	16(16.5)	0.03	0.9>P>0.5
		μ15	36	22(18)	14(18)	1.8	0.5>P>0.1
		μ73	33	20(16.5)	13(16.5)	0.74	0.5>P>0.1
		μF43	35	16(17.5)	19(17.5)	0.26	0.9>P>0.5
	HFA192	SSOSL85	33	15(16.5)	18(16.5)	0.27	0.9>P>0.5
		μ60	34	11(17)	23(17)	4.23*	P<0.05
		μF43	34	17(17)	17(17)	0.0	P>0.9
		ApoAI	27	16(15)	14(15)	0.13	0.9>P>0.5

Figure 4.1

Amplification products generated from an Atlantic salmon (S), a brown trout (T) and their hybrid offspring (HFJ91), at the μ73 microsatellite locus. Arrows at right indicate codominant variant alleles in the brown trout parent which differ in molecular weight. Segregation of alternate alleles can be observed in the hybrid progeny. The Atlantic salmon parent is homozygous for alleles which are greater in size than both brown trout alleles (arrow on left). A control reaction is identified by a **c** and the molecular weight marker (M13 sequencing reaction) by an **m**.

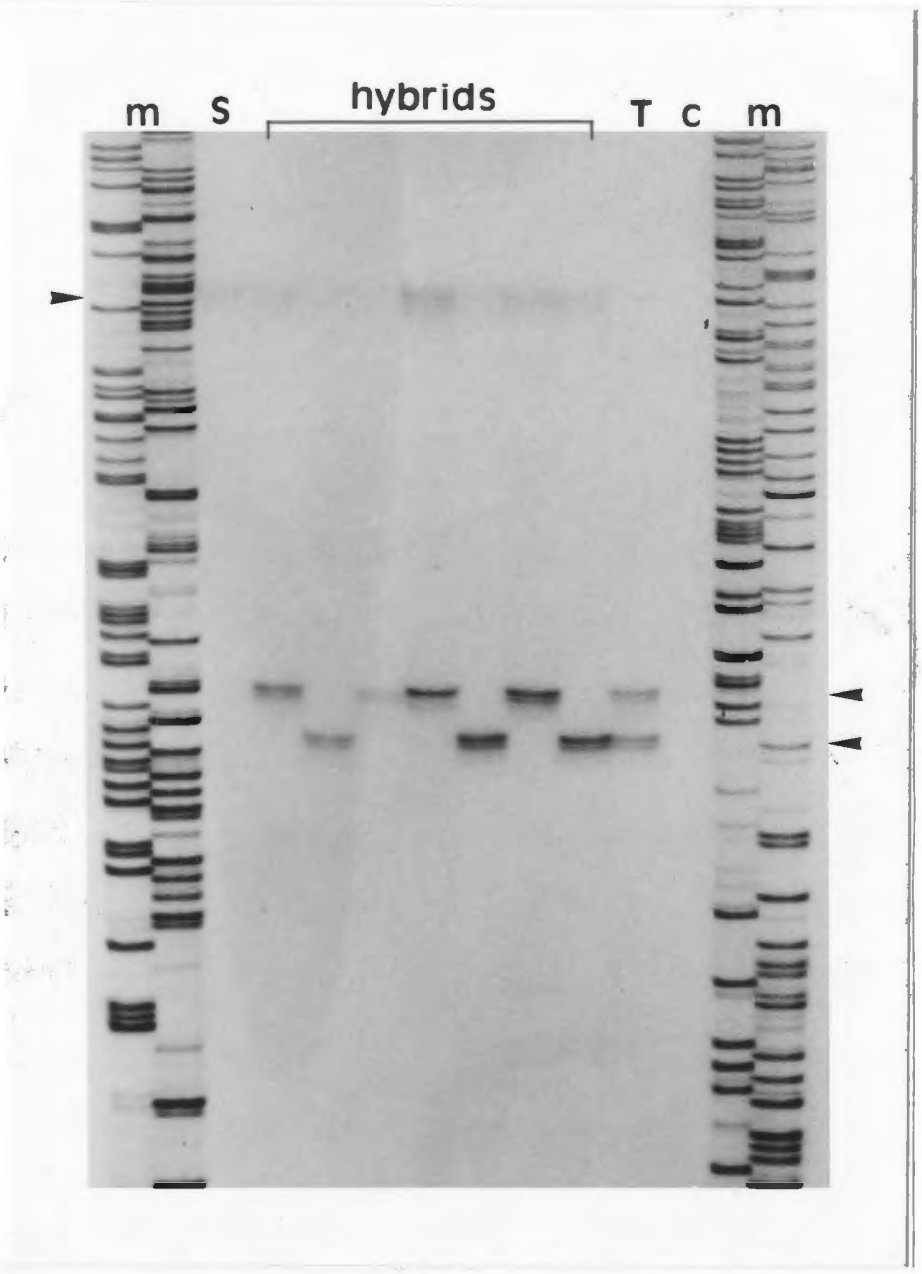

m
S
hybrids
T
c
m

Figure 4.2

Amplification products generated from an Atlantic salmon (S), a brown trout (T) and their hybrid offspring (HFJ91), at the SSOSL311 microsatellite locus. Arrows indicate codominant variant alleles in the brown trout parent (a) and the Atlantic salmon parent (b). Segregation of variant alleles inherited from both parents can be observed in the hybrid progeny.

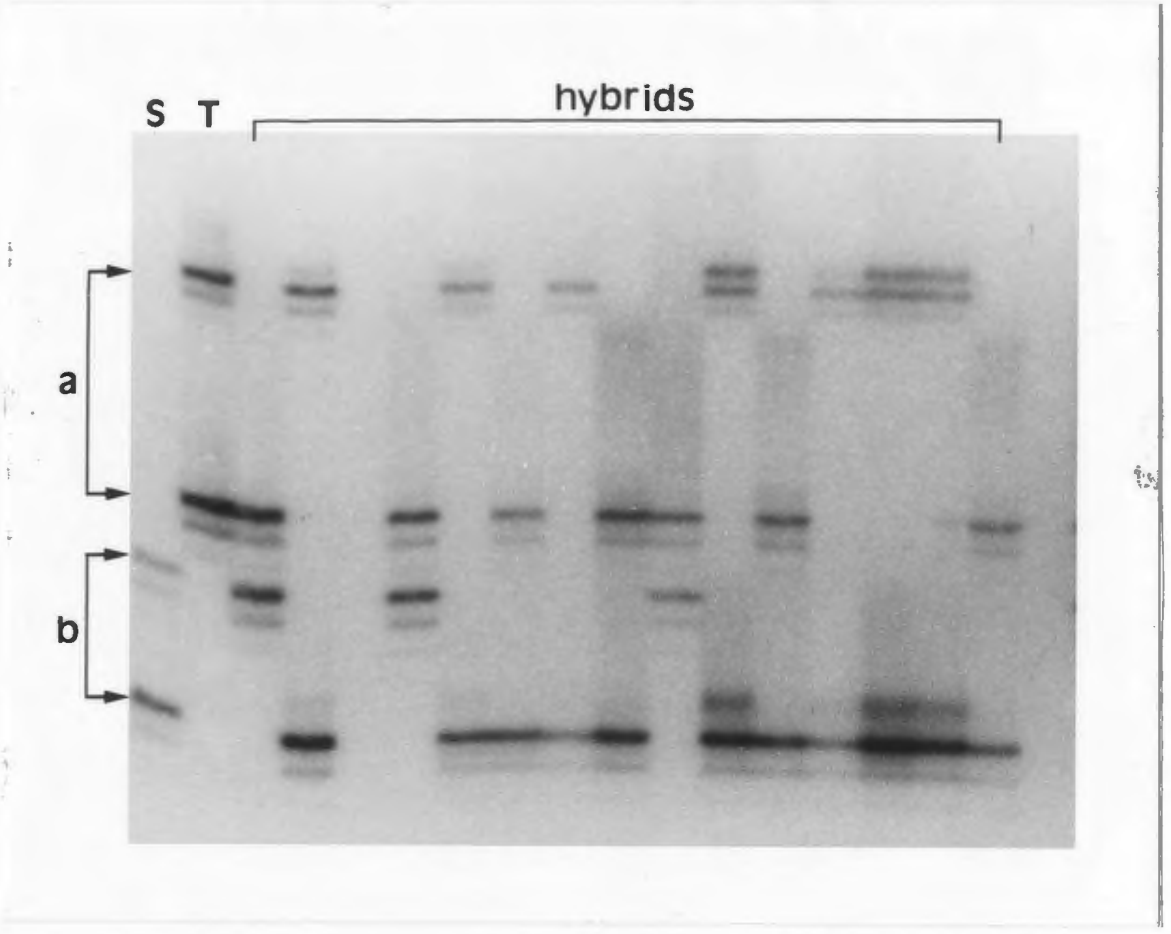
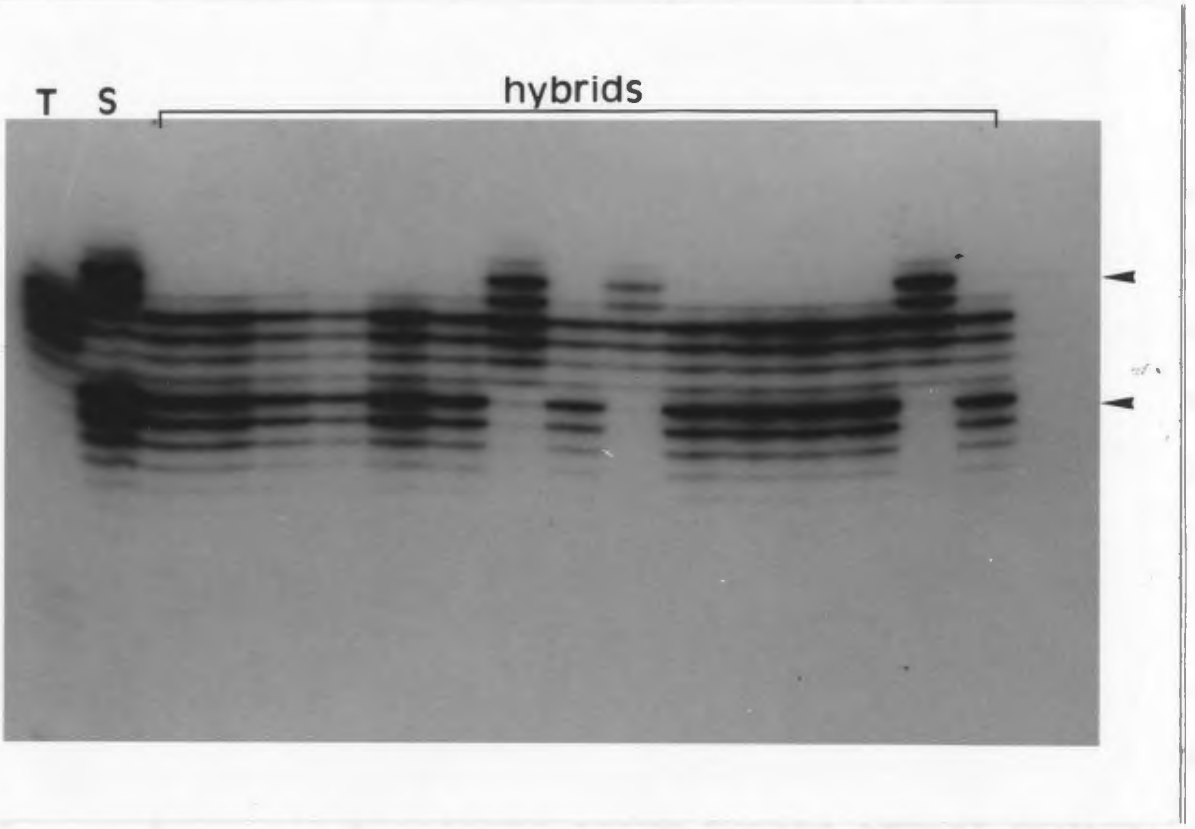

S T
hybrids
a
b

Figure 4.3

Amplification products generated from an Atlantic salmon (S), a brown trout (T) and their hybrid offspring (HFJ91), at the ApoAI microsatellite locus. Arrows indicate codominant variant alleles in the Atlantic salmon parent which differ in molecular weight. The brown trout parent is homozygous for a marker which is intermediate in size to that of the Atlantic salmon alleles.


T S
hybrids

progeny at microsatellite loci are given in appendix C.

4.3.2 Linkage analysis of polymorphic loci

In the HFJ91 brown trout parent, one highly significant ($P<0.001$) linkage was observed between microsatellites SSOSL311 and SSOSL417 (Table 4.3). Out of 65 pairwise comparisons, seven significant ($P<0.05$) linkages were detected between microsatellite and RAPD loci, three of which were highly significant ($P<0.001$).

In the HFJ91 Atlantic salmon parent, five pairs of RAPD and microsatellite markers demonstrated significant ($P<0.05$) linkage (Table 4.4). Tight linkage ($P<0.001$) was identified between three different pairs of microsatellite markers. One of these pairs (SSOSL417 and SSOSL311) was also linked ($P<0.001$) in the brown trout parent of the same family (Table 4.3).

To determine if the three linked microsatellite pairs in Atlantic salmon are conserved in brown trout, additional linkage analysis was performed using hybrid family HFA192, in which both the brown trout and Atlantic salmon parents were known to be heterozygous at all four loci (Table 4.5). In two cases, tight linkage between microsatellite loci was conserved in brown trout. One pair, μ60 and μF43, segregated independently of one another ($\chi^2 = 0.11$, $0.9>P>0.5$) and is apparently not linked in the brown trout parent.

Table 4.3
Joint segregation analysis and recombination frequencies (r) for microsatellite and RAPD markers inherited from the brown trout parent in an Atlantic salmon x brown trout hybrid family (HFJ91). Only marker pairs demonstrating significant (P<0.05) linkage have been listed.

Marker A	Marker B	N	progeny genotype a_1	a_2	a_3	a_4	χ^2 (df=1)	r
μ15	UBC219:432:*Str*	33	21	0	0	12	33.0***	0
μF43	UBC243:823:*Str*	35	14	2	3	16	17.9***	0.14
μF43	UBC348:739:*Str*	34	15	0	1	18	30.1***	0.029
μF43	UBC273:530:*Str*	35	11	5	5	15	6.43**	0.29
μ73	UBC280:281:*Str*	33	8	12	11	2	5.12**	0.30
μF43	UBC342:494:*Str*	33	10	5	6	12	3.6*	0.33
μF43	UBC354:517:*Str*	34	5	10	13	6	4.23*	0.32
SOSL417	SSOSL311	30	15	0	1	14	26.1***	0.033

* P<0.05
** P<0.025
*** P<0.001

Table 4.4
Joint segregation analysis and recombination frequencies (r) for microsatellite and RAPD markers inherited from the Atlantic salmon parent in an Atlantic salmon parent in an Atlantic salmon x brown trout hybrid family (HFJ91). Only marker pairs demonstrating significant (P<0.05) linkage have been listed.

			progeny genotype				χ^2	
Marker A	Marker B	N	a_1	a_2	a_3	a_4	(df=1)	r
Ssa4	UBC368:220:*Ssa*	30	11	5	4	10	9.97**	0.21
SSOSL311	UBC234:778:*Ssa*	32	4	16	7	5	6.12*	0.28
μ60	μF43	34	15	0	1	18	30.1***	0.029
SSOSL85	ApoAI	32	3	18	9	2	15.1***	0.16
SSOSL417	SSOSL311	30	10	3	1	16	16.1***	0.13

* P<0.025
** P<0.005
*** P<0.001

Table 4.5
Test for linkage conservation between microsatellite loci in Atlantic salmon and brown trout. Linkage analysis was performed on two different Atlantic salmon x brown trout hybrid families (HFJ91 and HFA192), each with unique male (♂) and female (♀) parents. Recombination frequencies (r) are also given.

Marker pair	hybrid family	N	informative parent (sex/id #)	progeny genotype a_1	a_2	a_3	a_4	χ^2 (df=1)	r
μ60	HFJ91	34	salmon (♀/1634)	15	0	1	18	30.1*	0.029
μF43			trout (♂/1619)	not heterozygous				-	-
	HFA192	34	salmon (♀/5211)	0	17	17	0	34.0*	0.0
			trout (♂/1105)	6	5	11	12	0.12ns	>0.5
ApoAI	HFJ91	32	salmon (♀/1634)	3	18	9	2	15.1*	0.16
SSOSL85			trout (♂/1619)	not heterozygous				-	-
	HFA192	29	salmon (♀/5211)	13	3	0	13	18.2*	0.10
			trout (♂/1105)	0	14	14	1	25.1*	0.03
SSOSL311	HFJ91	30	salmon (♀/1634)	10	3	1	16	16.1*	0.13
SSOSL417			trout (♂/1619)	15	0	1	14	26.1*	0.033

* P<0.001
ns: not significant

This last result should be accepted with caution due to non-normal segregation ratios at μ60, when inherited from the brown trout parent in family HFA192 (Section 4.3.1).

4.4 Discussion

These results demonstrate that microsatellites can be effective markers for genetic mapping in Atlantic salmon and brown trout, using a hybrid breeding strategy. Amplification products were highly reproducible, and the genetic contribution of each parent was easily identified in the offspring. Variant alleles were codominant, simplifying the identification of heterozygous loci in parents and the designation of genotype in progeny.

The high rate of heterozygosity detected at these loci make them ideal markers for linkage analysis. Cosegregating microsatellite pairs were observed in both the brown trout and Atlantic salmon parents. Linkage was also identified between microsatellite loci and several of the RAPD markers which have been described earlier in Chapter 3. As stated in Section 3.4, the large number of comparisons which are made when testing for linkage, will increase the probability of getting a significant result by chance alone. Consequently, only highly significant linkages ($P<0.001$) should be accepted.

Loosely linked pairs ($P<0.05$) have also been presented; however, confirmation, by analysis of larger mapping populations, is required. Linkage groupings, involving both RAPD and microsatellite markers, have been summarized in Chapter 6.

Conservation of microsatellite loci between brown trout and Atlantic salmon has been a major advantage. In many cases, primers which had been developed to amplify repeats in one species, have been equally effective in both. Cross species amplification will help to increase the number of markers available for mapping and demonstrates the importance of testing newly developed primers on a broad variety of species. Screening Atlantic salmon and brown trout for microsatellite loci which have been characterized in more distantly related species, such as rainbow trout (Morris et al. 1996), brook trout (Angers et al. 1995), or sockeye salmon (Scribner et al. 1996), may result in the identification of new markers, which can be used for linkage mapping.

The ability to amplify loci which are common to both brown trout and Atlantic salmon has also made it possible to test for the evolutionary conservation of linkage groups. Of the three cosegregating microsatellite pairs identified in Atlantic salmon, only two are linked in brown trout. This

difference may be the result of chromosomal rearrangements which have occurred since the divergence of these two species. It may be that these loci are located on the same chromosome in brown trout, but are separated by a distance which cannot be detected through the analysis of recombination frequencies (i.e. greater than 50 map units apart). As the marker density of genetic maps for both these species increases, the mechanisms responsible for differences between them, should become more apparent. It should be noted that one of the markers (μ60) which was linked in salmon but unlinked in brown trout, also demonstrated a non-random segregation. This deviation may have effected the outcome of the linkage analysis. In the future, it may be more prudent to utilize statistical formulas which account for segregation ratios when calculating linkage such as those proposed by Garcia-Dorado and Gallego (1992).

Comparisons of linkage maps will become more meaningful as more microsatellite loci are found in association with functional genes. For example, the ApoAI microsatellite is located in the second intervening sequence (intron) of a gene (apolipoprotein AI) which is commonly expressed in all vertebrate species. Consequently, this particular locus will be highly informative with regards to long range comparisons

between maps of *Salmo* species with those of other vertebrates, such as mammals, birds, or amphibians.

Perhaps the biggest obstacle to constructing a high resolution linkage map with microsatellites, will be the characterization of informative loci. The majority of markers screened in this study were developed to conduct genetic analysis of natural populations. A sufficient number of loci have now been described in both brown trout and Atlantic salmon to satisfy this purpose. More primers will have to be developed with the specific goal of generating a genetic map. This will require a considerable investment of time and resources. Towards this end, the sequences of 16 microsatellite loci, which were isolated from an Atlantic salmon partial geneomic library using the proceedures of (Taylor et al. 1994), are given in appendix D. Primer sequences and reaction conditions for amplification must still be determined before these loci can be tested.

Chapter 5

An Evaluation of Expressed Sequence Tags as Non-Ambiguous Markers for Genetic Mapping in Atlantic Salmon and Brown Trout

5.1. Introduction

5.1.1 Expressed sequence tags

Expressed sequence tags (ESTs) are complementary DNA (cDNA) fragments that are generated from the messenger RNA (mRNA) of different tissues (i.e. all the clones of a cDNA library). They can be used as probes to detect restriction fragment length polymorphisms (RFLPs) in the total genomic DNA of an organism (Botstein et al. 1980; Pogson et al. 1995; Bumstead and Palyga 1992), or, sequence information obtained from the clone can be used to produce PCR amplified markers (Durkin et al. 1992; Adams et al. 1991; Mazzarella and Srivastava 1994). Thousands of ESTs have been isolated from a variety of organisms and their importance to genome mapping projects is well established (Boguski 1995; Davies 1993; Adams et al. 1993 and 1991).

Identification of ESTs which correspond to genes of known

function has been simplified by the creation of electronic data bases such as GenBank or the European Molecular Biology Laboratory (EMBL). The dbEST data base was specifically created for the storage and comparison of EST sequence data (Boguski et al. 1993). Partial sequence from a cDNA clone can provide enough information for a computer assisted comparison with registered sequences within a data base (Adams et al. 1991; Palwak et al. 1995; Park et al. 1993). ESTs which are assigned to particular genes, can then be used to represent non-ambiguous loci for comparative gene mapping (O'Brien et al. 1993; Johansson et al. 1995; Dubcovsky et al. 1996). Such markers represent important anchor loci which can provide reference points between newly discovered linkage groups and existing genomic maps of the same or related species. If the corresponding gene is unknown, an EST can still be used as an anonymous genetic marker.

The advantage of using ESTs is that it enables screening for variability at gene loci which do not have electrophoretically detectable protein products or for which no protein variants are known.

Allozyme electrophoresis is usually restricted to water soluble proteins that can be detected by an enzyme coupled stain. This limits the number of loci which can be mapped and excludes many non-enzymatic gene products such as structural, regulatory, and transport proteins. In addition, protein

electrophoresis will only detect mutations in the exons of genes which result in an alteration of the size, charge or conformation of the protein product. ESTs can sample polymorphisms which result from mutations in the introns, exons or flanking regions of a particular gene. Consequently, greater levels of polymorphism are expected to be observed using ESTs than with protein electrophoresis (Pogson et al. 1995).

This approach can be demanding with regards to the time and effort required to construct cDNA libraries and determine partial sequences from a large number of clones. Southern blot analysis for RFLPs can also be labour intensive and requires large quantities of genomic DNA. These problems are mitigated, however, by the exchange of sequence information through electronic data bases and advances in PCR technology, such as the ability to detect single strand conformational polymorphisms (SSCP) (Orita et al. 1989) or heteroduplex analysis (Ganguly et al. 1993).

5.1.2 Gene mapping with expressed sequence tags

ESTs representing both known and anonymous genetic loci have been assigned to the genetic maps of many different organisms. As probes for RFLP analysis, they have been used in the construction of linkage maps for plants such as alfalfa (*Medicago sativa*) (Kiss et al. 1993), maize (Gardiner et al.

1993), barley (*Hordeum vulgare*) (Langridge et al. 1995) and rice (Causse et al. 1994) and for animals as diverse as chicken (Burt et al. 1995; Bumstead et al. 1992), pig (Ellegren et al. 1994) and mosquito (*Aedes aegypti*) (Severson et al. 1993). In advanced mapping projects, such as the human genome, ESTs can be assigned to specific chromosomes by screening a panel of somatic cell hybrids using PCR and sequence specific primers (Durkin et al. 1992). Further refinements to the arrangement of loci can be accomplished using yeast artificial chromosomes (YACs) (Das Gupta et al. 1993; Guilford et al. 1995; Mazzarella and Srivastava 1994) or linkage analysis of RFLPs (Eng et al. 1992).

5.1.3 The use of expressed sequence tags in fish

Despite their acceptance as important markers for genetic analysis (Goodier and Davidson 1993; Thorgaard 1992), application of ESTs in fish is uncommon. Pogson et al. (1995) compared patterns of variation between allozymes and nuclear RFLP loci (detected using anonymous cDNA clones) in populations of Atlantic cod (*Gadus morhua*). In this case, higher levels heterozygosity and a greater number of variant alleles were detected using ESTs than with protein electrophoresis. Other studies involving cloned cDNA sequences in fish are restricted to evolutionary or biochemical aspects of particular genes (Thorgaard 1992).

These clones, however, can still be used as genetic markers for genome mapping or population analysis. A search of GenBank (release 97) for mRNA sequences in Atlantic salmon and brown trout lists 48 and seven different clones, respectively (these include four brown trout sequences and five Atlantic salmon sequence which are reported in this study). Many other sequences from closely related species such as rainbow trout (*Oncorhynchus mykiss*) or coho salmon (*Oncorhynchus kisutch*) can also be found in the GenBank data base.

5.1.4 Objectives

The objectives of this study were to produce ESTs as non-ambiguous markers and to test their potential effectiveness for genome mapping in Atlantic salmon and brown trout. The approach taken was to randomly isolate and sequence clones from two cDNA libraries and compare them with known sequences from a computerized data base. Clones representing genes of known function were then used as probes in Southern blot analysis to screen for RFLPs in the nuclear DNA of Atlantic salmon and brown trout and their hybrids.

5.2 Materials and Methods

5.2.1 Brood stock, hybrid families, and DNA isolation

Collection of brood stock and the production of hybrid families have been described previously in Section 2.2.1.

Procedures for the isolation of DNA have been provided in Section 2.2.2.

5.2.2 Construction of cDNA libraries

Two complementary DNA (cDNA) libraries were constructed. The first library was made from the liver tissue of a juvenile brown trout and is designated BTL. The second was made from the skeletal muscle of a juvenile Atlantic salmon and is designated ASM. Proteins from both these tissues have been studied in a wide variety of vertebrates, increasing the likelihood of finding sequences which have been registered within the data bank and can be used for comparative mapping.

Ribonucleic acid (RNA) was extracted from the tissue of recently killed fish according to the protocol of Chomczynski and Sacchi (1987) with some modifications. All glassware used in the procedure, and the water used to make all of the solutions, were first treated with diethyl pyrocarbonate to inactivate any RNAse. One gram of tissue was homogenized in 10 ml of extraction buffer (4 M guanidinium thiocyanate, 25 mM sodium citrate, pH 7, 0.5% sarcosyl and 0.5 mM β-mercaptoethanol) using a polytron. An organic extraction was performed by adding 1 ml of 2 M sodium acetate (pH 4), 11 ml of water saturated phenol, and 2.2 ml of chloroform:isoamyl alcohol (24:1) to the homogenate. After gentle mixing for 2 min, the homogenate was incubated on ice for 15 min, then

subjected to centrifugation at 3000 x *g* for 30 min at 4°C. The upper aqueous phase was transferred to a clean tube, mixed with an equal volume of cold (-20°C) nucleic acid grade isopropanol, and stored at -20°C for 1 hr. The precipitate was recovered by subjecting the tube to centrifugation at 3000 x *g* for 30 min at 4°C, then discarding the supernatant. The pellet was dissolved in 3 ml of extraction buffer, then precipitated by mixing in 3 ml of cold nucleic acid grade isopropanol and storing at -20°C for 1 hr. The precipitate was subjected to centrifugation as above and the supernatant discarded. The pellet was resuspended in 5 ml of cold (-20°C) 95% nucleic acid grade ethanol, then brought to a final volume of 6.3 ml with sterile water. This last step adjusted the ethanol concentration to 75%. The suspension was subjected to centrifugation as above and the supernatant discarded. The pellet, consisting of purified total RNA, was dissolved in 1 ml of sterile water, and stored at -70°C, prior to isolation of messenger RNA (mRNA).

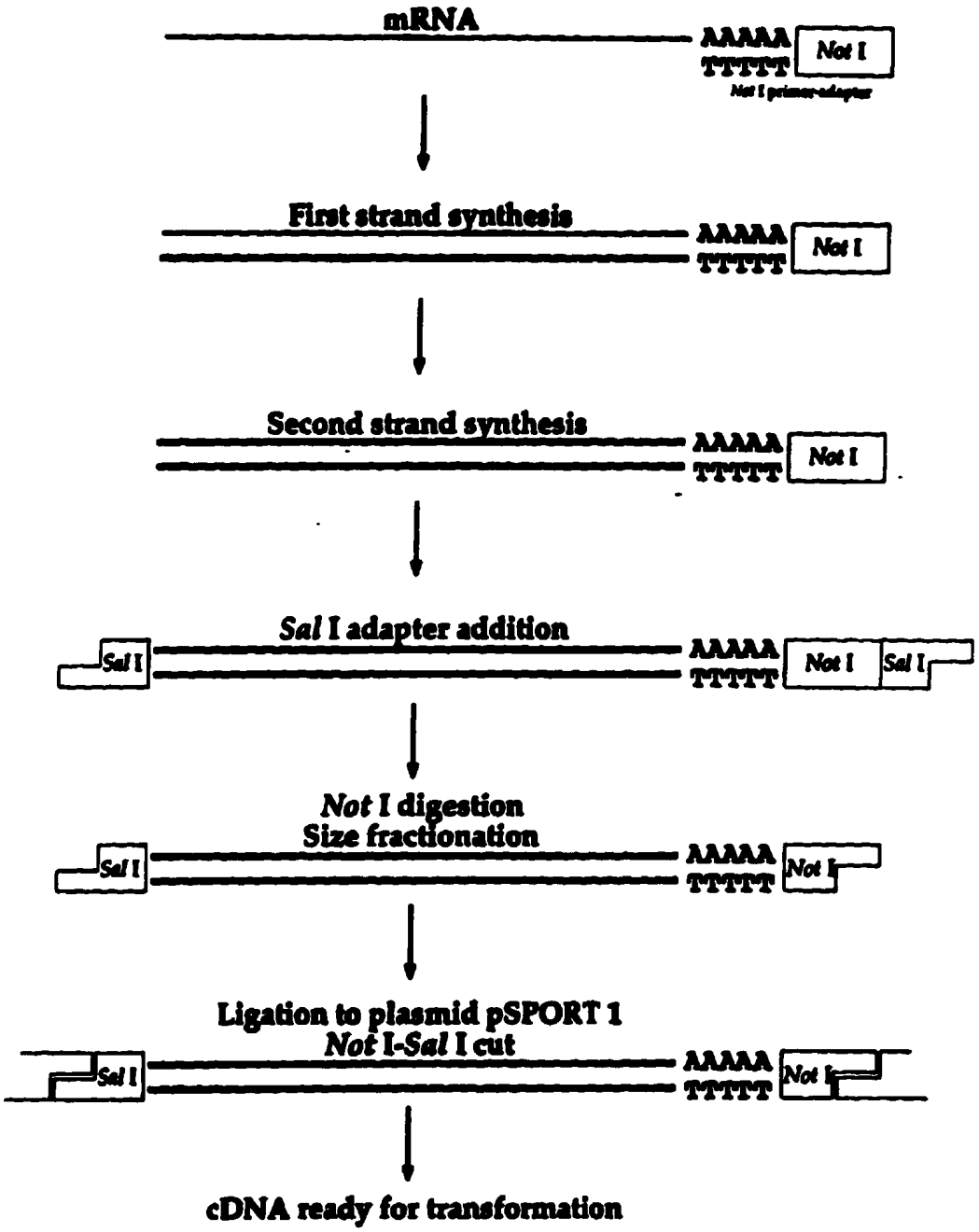
Polyadenylated RNA (mRNA) was separated from nonpolyadenylated RNA (transfer RNA and ribosomal RNA) by affinity chromatography with oligo dT cellulose (Gibco BRL, Life Technologies Inc., Gaitherburg, MD) and methods described by Davis et al. (1986). A sample of the purified mRNA was converted to complementary DNA (cDNA) and ligated into the

plasmid pSPORT 1 using the SuperScript™ plasmid system for cDNA synthesis and plasmid cloning (Gibco BRL). This kit enables directional cloning of cDNA fragments into a plasmid which has been digested at the multiple cloning site with the restriction enzymes *NotI* and *SalI* (Figure 5.1). One end of the cDNA fragment, representing the 3' polyadenylation site of the mRNA sequence, was always ligated to the *NotI* terminus of the plasmid. The 5' end of the cDNA fragment, where polypeptide template sequence is most likely to occur, was always ligated to the *SalI* terminus of the plasmid. Directional cloning simplified subsequent analysis of the cDNA clones, such that the 5' end of the cDNA fragment, provided sufficient sequence information for a data base search. It should be noted, however, that not all clones will represent the entire length of an mRNA transcript.

The ligation product was used to transform MAX Efficiency™ competent *E. coli* DH5α cells (Gibco BRL) which were grown overnight at 37°C on agar plates, in the presence carbenicillin. Plasmids could not self-ligate, therefore, only cells hosting recombinant plasmids, with inserted cDNA, formed colonies. Clones were transferred to duplicate agar plates, organised into a grid, and grown overnight at 37°C, in the presence of carbenicillin. The library was stored short-term (two weeks), at 4°C, on the agar plates. For long-term

Figure 5.1

Summary of the directional cloning procedure for the *SuperScript*™ plasmid system for cDNA synthesis and plasmid cloning. The figure is taken from the instruction manual for the kit (Gibco BRL, Life Technologies Inc.).


mRNA
AAAAA
TTTTTT
Not I
First strand synthesis
Second strand synthesis
Sal I adapter addition
Sal I
Not I digestion
Size fractionation
Ligation to plasmid pSPORT 1
Not I-Sal I cut
cDNA ready for transformation

storage, colonies were grown overnight at 37°C in 150 μl LB medium with 5.25 μg of ampicillin, in a 96 well U-Bottom Microtest III™ plate (Becton Dickinson, Lincoln Park, NJ). An equal volume of 30% glycerol in LB medium was added to each culture before storing plates at -70°C (Brown and Knudson 1991).

5.2.3 Sequencing and analysis of cDNA clones

Clones were randomly selected and cultured overnight at 37°C in 10 ml of LB medium with 350 μg of ampicillin (Sambrook et al. 1989). Plasmids were isolated using the Wizard™ MiniPreps DNA purification system (Promega). Inserts were sequenced once, at the terminus representing the 5' end of the mRNA transcript, using a T7 primer (5'-d[TTAATACGACTCACTAT]-3') and the ^{T7}Sequencing™ kit (Pharmacia). Sequences, ranging in length from 100 to 250 bp, were submitted via e-mail to the European Molecular Biology Laboratory (EMBL), for a computer assisted comparison with GenBank and EMBL data base entries, using the FASTA program of Pearson and Lipman (1988). Clones which were found to represent genes of known protein function were selected as probes to search for RFLPs in the genomic DNA of the Atlantic salmon and brown trout brood stock.

5.2.4 RFLP analysis using cDNA probes

Samples of total genomic DNA (10 μg), from the Atlantic salmon and brown trout brood stock, were digested with a

variety of restriction endonucleases. Enzymes with four base (*TaqI*, *Sau3AI*, *PalI*, *AluI*) and six base (*BglII*, *PstI*, *PvuII*) recognition sequences were selected, and used according to supplier's recommendations (Pharmacia). Digests were subjected to electrophoresis in a 1% agarose gel with a 1X TBE buffer (0.089M Tris-borate, 0.89M boric acid, 0.002M EDTA), at a constant voltage of 1.8 V/cm for 16 hr. Due to the preliminary nature of these experiments, molecular weight markers were not run on every gel since their effect on the clarity of the blot image were still uncertain.

Southern blots were prepared by soaking gels twice, for 30 min, in a 1.5M NaCl/0.5M NaOH solution, then transferring DNA fragments, by capillary action, to a nylon membrane (Biodyne® B, Pall BioSupport, East Hills, NY) using a 1.5M NaCl/0.25M NaOH solution (Sambrook et al. 1989). After transfer, membranes were baked at 80°C for 1 hr.

Probes were made from the cDNA inserts of isolated plasmids. Inserts were separated from the plasmid by digesting with two endonucleases (*PstI* and *HindIII*), each of which cut once, at the multiple cloning site, on either side of the cDNA fragment. Products were size fractionated in a 1% NuSieve® GTG® low melting point agarose gel (FMC BioProducts) with a 1X TA buffer (0.04M Tris-acetate, pH 7.0), and the cDNA fragments recovered from the gel using the GeneClean II® kit

(Bio 101 Inc., LaJolla, CA). Purified cDNA inserts were quantified by UV spectrophotometry, and samples of approximately 25 ng were labelled with [α^{32}P]dCTP (Amersham) using the Megaprime™ DNA labelling systems kit (Amersham). Unincorporated radioactive nucleotides were removed from the labelling reaction by filtration through a Sephadex® G-50 resin (Pharmacia). Probes were boiled for 5 min prior to being added to the hybridization solution.

All hybridizations were carried out in a hybridization incubator (Robbins Scientific Inc., Sunnyvale, CA). Southern blots were soaked in 15 ml of prehybridization solution (1.5X SSPE [225 mM NaCl/1.3 mM NaH_2PO_4/1.5 mM Na_2EDTA, pH 8.0], 10% polyethylene glycol [PEG], 7% SDS) for at least 1 hr at 65°C, then incubated overnight at 65°C in 15 ml of hybridization solution (1.5X SSPE, 10% PEG, 7% SDS, labelled probe). Membranes were washed at 60°C according to the following schedule: 2 x 15 min in 2X SSC [300 mM NaCl/30 mM Na_3Citrate]/0.5% SDS, 2 x 15 min in 0.2X SSC/0.5% SDS and 2 x 15 min in 0.2X SSC/0.1% SDS. Excess moisture was removed from the blots before exposing them to X-ray film, for 7 days, at -70°C, with an intensifying screen.

Southern blots were recycled by soaking them first in 0.4M NaOH at 45°C, for 30 min, twice, then in a solution of 0.2M Tris (pH 7.5), 0.1X SSC, 0.1% SDS at 45°C, for 30 min, twice. After air drying, they could be used again, in

hybridization reactions, with different probes.

5.3 Results

5.3.1 Sequencing and characterization of cDNA clones

Partial sequences of 34 cDNA clones from the BTL library, and 29 clones from the ASM library were submitted, by e-mail, for comparison with entries in the EMBL and GenBank data bases. Twenty-one of the sequences were similar to genes of known protein function (7 BTL clones and 14 ASM clones) (Table 5.1). In four cases, clones representing the same gene were sampled more than once (two serum albumin clones from the BTL library; three creatine kinase, two heavy chain myosin and two parvalbumin clones from the ASM library). Consequently, clones for 16 different genes (six from the BTL library and ten from the ASM library), were identified. All matches were made with vertebrate species. Three of the genes, apolipoprotein AI (Powell et al. 1991), creatine kinase (Garber et al. 1990) and serum albumin (Byrnes and Gannon 1990), have been described previously in salmonid fish. Two genes, fibrinogen and α-tropomyosin, have been described in lamprey (*Petromyzon marinus*) (Bohonus et al. 1986) and zebrafish (Ohara et al. 1989) respectively. The remaining ten sequences have not been reported in teleosts before now.

Table 5.1
Putative identification of cDNA clones from a brown trout liver (BTL) and an Atlantic salmon muscle (ASM) library. Also given are the species and accession numbers representing the sequence to which the greatest percent similarity was observed when using the FASTA program.

Clone	cDNA size (bp)	putative identification	species with best sequence alignment	access. number	percent alignment
BTLA40	800	transplantation antigen	human	X56932	79
BTLB27	1000	apolipoprotein AI	Atlantic salmon	X52237	97
BTLF44	1300	serum albumin	Atlantic salmon	X52397	95
BTLC22	500	serum albumin	Atlantic salmon	X52397	97
BTLF45	1500	trypsin inhibitor	human	X07173	68
BTLF39	2000	elongation factor	rat *Rattus norvegicus*	X63561	84
BTLE36	1400	fibrinogen	lamprey	M14773	64
ASM1.3	800	aldolase	rat	M12919	78
ASM8.27	650	transcription factor	human	X53281	82
ASM4.2	700	alpha-tropomyosin	zebrafish	M24635	88

Table 5.1 Continued

Clone	cDNA size (bp)	putative identification	species with best sequence alignment	access. number	percent alignment
ASM11.30	800	parvalbumin	rat	M15454	74
ASM11.35	900	parvalbumin	rat	M15454	66
ASM11.24	900	myosin heavy chain	human	X03741	78
ASM4.9	300	myosin heavy chain	chicken	M74086	73
ASM1.22	1600	creatine kinase	dog	M11660	78
ASM8.35	700	creatine kinase	rainbow trout	X53859	72
ASM11.17	900	creatine kinase	rainbow trout	X53859	83
ASM4.7	400	ribosomal protein	human	M77233	84
ASM4.4	500	beta-globin	duck *Cairina moschata*	J00926	67
ASM1.21	500	myosin light chain	rabbit *Oryctolagus cuniculus*	X54043	80
ASM8.33	500	troponin I	chicken	X03832	79

The approximate size of cloned cDNA fragments ranged between 300 and 3000 bp in length. Partial and complete cDNA sequences for all of the clones which have been submitted for comparison with the GenBank and EMBL data-base entries, are presented in appendix E and F. All of the sequences that were similar to genes of known protein function have been registered with the GenBank data base. Their accession numbers are given in appendix E.

5.3.2 RFLP analysis using cDNA probes

Eleven different ESTs were used as probes to search for RFLPs in the genomic DNA of the Atlantic salmon and brown trout brood stock (Table 5.2). All of the hybridization reactions resulted in different restriction fragment patterns. The number of bands scored for each reaction varied from one to 19, depending on which EST was used. When probes hybridized to more than one fragment, the intensity of bands often differed (Figure 5.2). Faint bands were difficult to score, especially if an inadequate quantity of DNA had been used for the Southern transfer. This problem was exacerbated when blots had been recycled more than once, and the quality of the nylon membranes had started to deteriorate.

Six different probes detected RFLPs in the genomic DNA of salmon and brown trout (Table 5.2). In all but one case, alternate alleles were not immediately apparent, and

Table 5.2
Detection of restriction fragments in the total genomic DNA of the Atlantic salmon and brown trout brood stock, using ESTs (* indicates potential RFLPs).

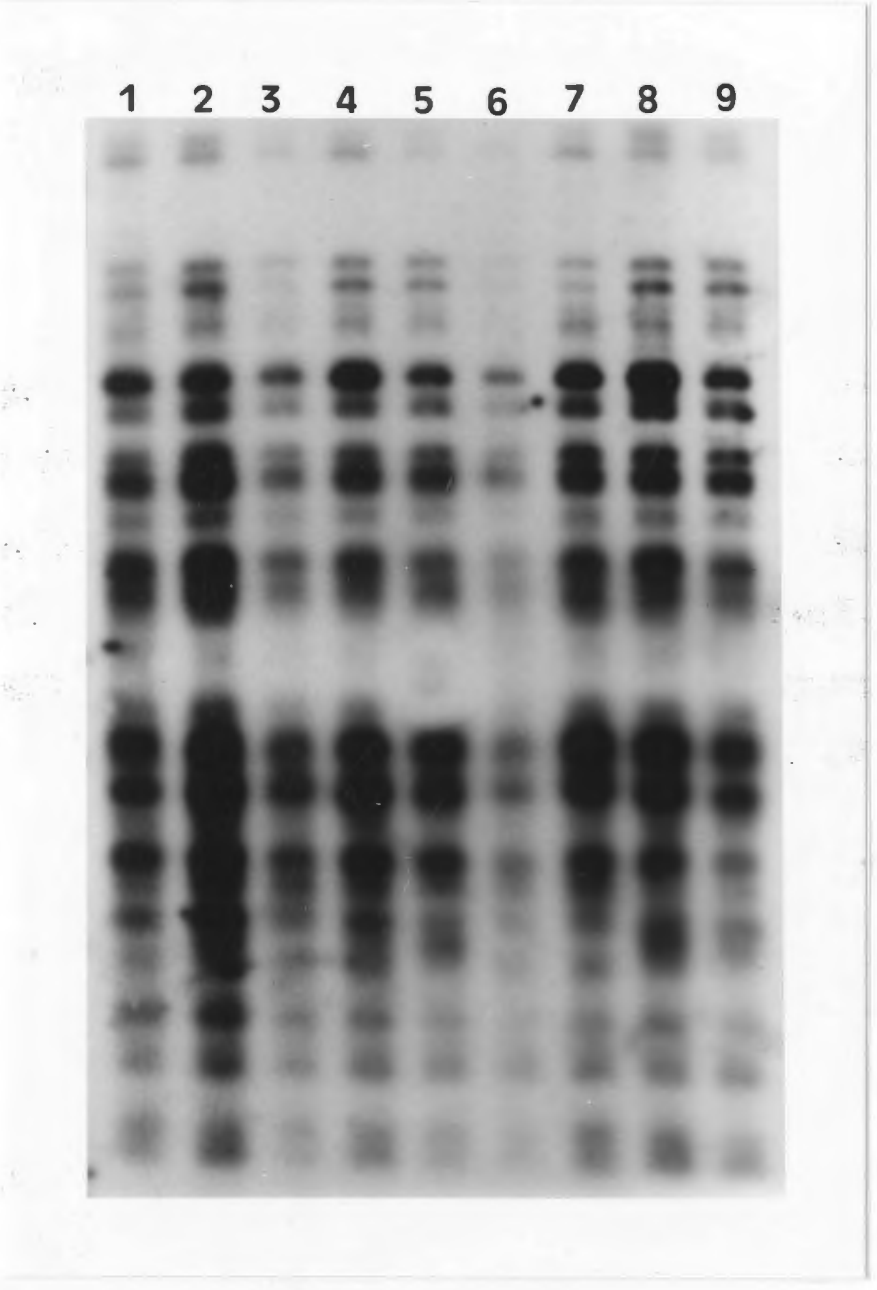
		No. obs. restriction frag.	
Probe	Enzyme	Atlantic salmon	Brown trout
Fribrinogen	Sau3AI	5*	
	PalI	4	4
	AluI	5*	5
Apolipoprotein AI	PstI	2	2
	PvuII	2	
	TaqI	3	
	BglII	2	2
Trypsin inhibitor	PstI	1	1
Beta-globin	AluI	7*	6
	PalI	5	5
	Sau3AI	5	
Myosin heavy chain	PstI	15	17*
	BglII	12	12
	TaqI	19	
	PvuII	13	
	PalI	7	
Aldolase	TaqI	2	2
	AluI	1	1
	Sau3AI	3	3

Table 5.2 continued

Probe	Enzyme	No. obs. restriction frag. Atlantic salmon	Brown trout
Troponin I	PalI	10*	
	TaqI	4	
Tropomyosin	Sau3AI	10*	
Ribosomal protein	AluI	2	
	Sau3AI	2	
	TaqI	2	
Albumin	PalI	4*	2*
	AluI	2	1
Transplant antigen	Sau3AI	2	

Figure 5.2

Southern blot of Atlantic salmon genomic DNA digested with *TaqI* and probed with the myosin heavy chain cDNA (ASM11.24). This particular probe detected multiple restriction fragments of variable intensity. Individual identification numbers for lanes 1 through 9 are: 5206♀, 1631♀, 5208♀, 5209♀, 1628♂, 5211♀, 5310♀, 1634♀, 1633♀.


1 2 3 4 5 6 7 8 9

individual genotypes were difficult to determine. For example, a possible RFLP detected in Atlantic salmon DNA which had been digested with *AluI*, and probed using the beta-globin EST, was either present or absent, making it impossible to identify heterozygous individuals (Figure 5.3). This band may represent a variable locus at which there has been the loss of a restriction site, resulting in a much larger fragment that cannot be detected on the gel. It is unlikely to be the consequence of incomplete digestion, as there was no evidence of this when the same membrane was reprobed with a different EST. Further analysis of segregating alleles in the offspring of heterozygous individuals will have to be conducted before polymorphic markers can be confirmed.

When hybridizing *AluI* digested Atlantic salmon DNA with the fibrinogen probe, variant alleles were observed at what appeared to be two different loci (Figure 5.4). For the sake of this discussion, these two loci will be reffered to as *fib*-1 (higher in molecular weight) and *fib*-2 (lower in molecular weight). At both loci, certain individuals could be identified as either heterozygous or homozygous for variant alleles. Segregation analysis of the *fib*-2 locus was performed using 24 hybrid offspring (HFJ91) of a heterozygous Atlantic salmon female. This was the same family used in the

Figure 5.3

Southern blot of Atlantic salmon genomic DNA digested with *AluI* and probed with the beta-globin cDNA (ASM4.4). Lanes 1 to 11 represent individual Atlantic salmon. The arrow indicates a potential marker which is clearly present in samples 5 and 7, but absent from samples 4, 6, 8 and 9. Individual identification numbers for lanes 1 through 11 are: 5206♀, 1631♀, 5208♀, 5209♀, 1628♂, 5211♀, 5310♀, 1634♀, 1633♀, 1630♂, 1629♂.

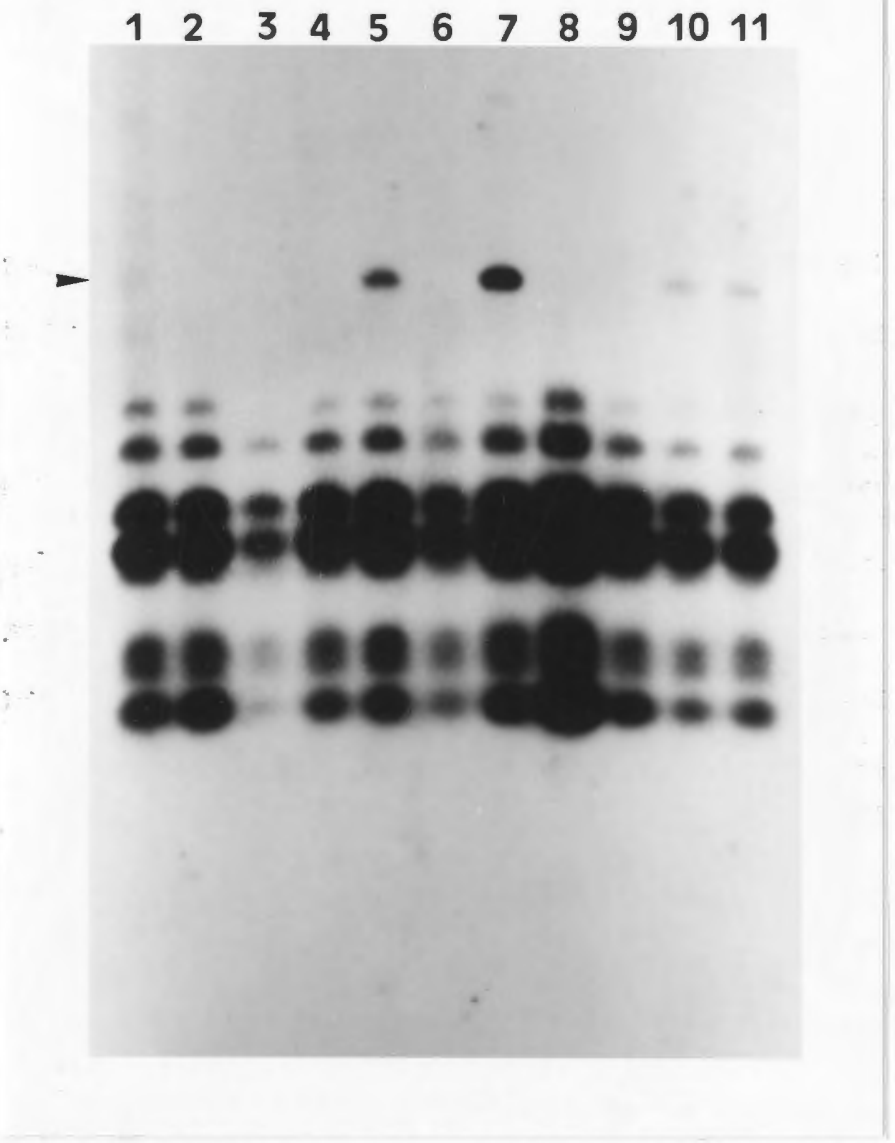
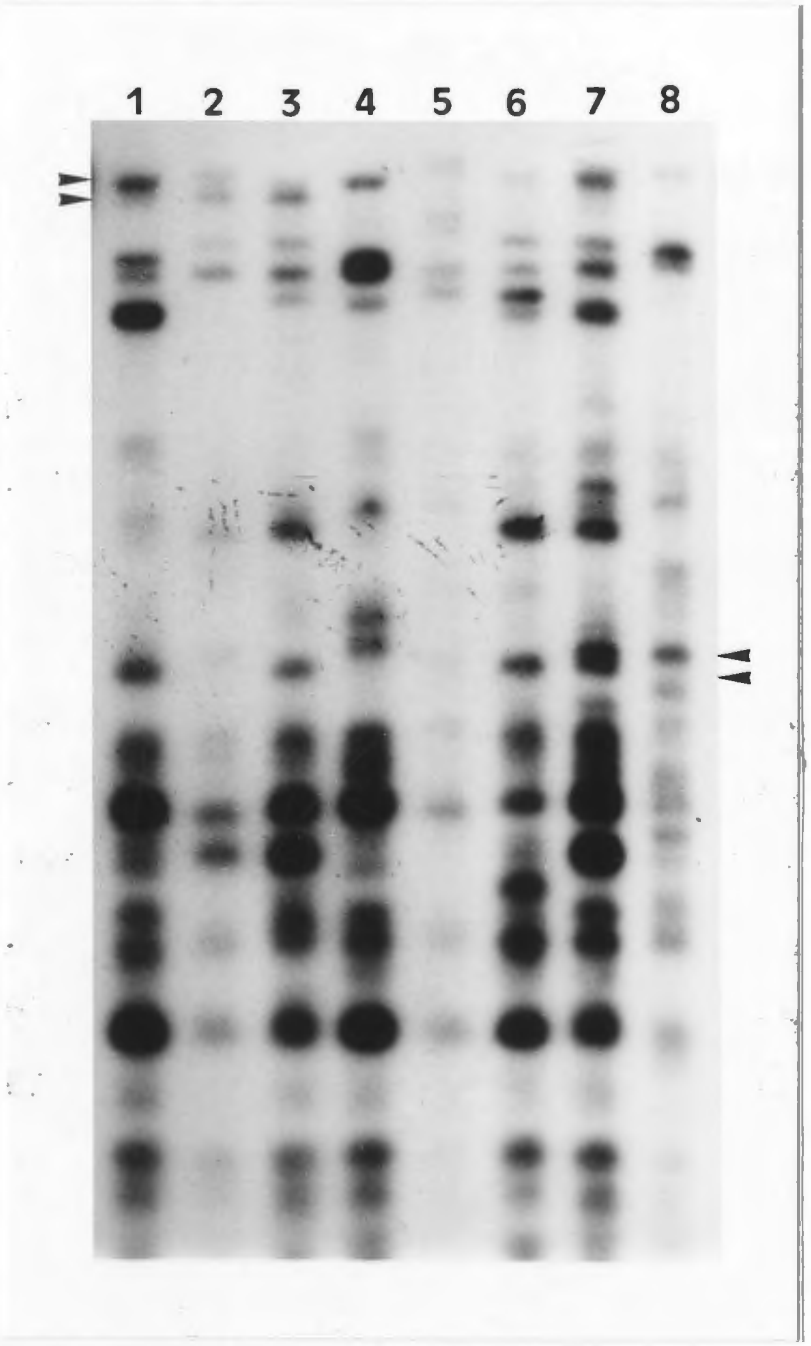

1 2 3 4 5 6 7 8 9 10 11

Figure 5.4

Southern blot of Atlantic salmon genomic DNA digested with *AluI* and probed with the fibrinogen cDNA (BTLE36). Lanes 1 to 8 represent individual Atlantic salmon. Arrows at the right indicate the variant alleles at *fib*-2. Samples 6 and 8 are homozygous for the fast and slow alleles respectively. Sample 7 is heterozygous for both alleles. Arrows at the left indicate the varient alleles at *fib*-1. Samples 1 and 3 are homozygous for the slow and fast alleles respectively. Sample 2 is heterozygous for both alleles. Individual identification numbers for lanes 1 through 8 are: 1631♀, 5208♀, 5209♀, 1628♂, 5211♀, 5310♀, 1634♀, 1633♀.


1
2
3
4
5
6
7
8

analysis of RAPD and microsatellite markers which has been presented in Chapters 2 and 4. Segregation ratios at this locus were normal (N = 24, χ^2 = 0.0, P>0.9). Linkage was not observed between the fibrinogen marker and any of the microsatellite or RAPD loci. This particular marker was faint and difficult to score in some individuals. Consequently, these results may be difficult to reproduce in other mapping populations. Individual genotypes of hybrid progeny at the *fib-2* locus are given in appendix C.

5.4 Discussion

Partial sequencing of randomly selected cDNA clones, followed by a computerised comparison with molecular data base entries, has been an effective way of collecting ESTs which can be used for mapping in Atlantic salmon and brown trout. Approximately 33% of the sequences which were submitted to the EMBL data base were similar to genes which have been described previously in Salmonids or in other vertebrate species. These clones provide a valuable resource as potential non-ambiguous markers, which can be used to compare genetic maps constructed for *Salmo*, with those of other organisms.

When used as probes to detect RFLPs in the genomic DNA of brown trout and Atlantic salmon, ESTs generated multi-banded

restriction fragment patterns. This was expected, given the tetraploid ancestry of salmonids (Johnson et al. 1987). Differences in the intensity of the bands was likely due to sequence divergence which has occurred between duplicated genes which may no longer be expressed or functional genes which belong to the same multigene family. Differences in sequence between the probe and genomic DNA would result in a lower melting temperature, less efficient hybridization and consequently bands of low intensity. Multiple banding patterns might also be the result of the cDNA probe hybridizing to several exons which are part of the same gene, but occur on different restriction fragments. This complicates the analysis of restriction fragment patterns since it is not clear whether multiple bands represent a single locus or several distinct loci. Since the focus of this experiment was to identify polymorphic markers for use in genetic mapping, less emphasis was placed on describing the restriction fragments themselves and the molecular weights of the various bands were not established. In future studies, more care should be given to this detail in order to improve the recording and transfer of information concerning these banding patterns.

Six probes detected variable banding patterns between

individual DNA samples using five different restriction enzymes. Complicated banding patterns made it difficult to interpret or confirm specific genotypes without further analysis of hybrid offspring. This problem was further confounded when polymorphisms were observed in faint bands, which were difficult to reproduce. The slow, labour intensive nature of Southern blot hybridization will make it difficult to establish a large number of genetic markers using this process.

As discussed earlier in Section 2.4, low variability may be a characteristic of the population from which the brood stock was sampled. Studies of Atlantic salmon using protein electrophoresis have generally observed a low number of polymorphic enzymes (Taggart et al. 1995; Ferguson 1994; Davidson et al. 1986). Although higher levels of variation are expected to be detected with ESTs (Pogson et al. 1995), these levels might still be low when compared with other vertebrates. Better results may have been obtained if highly outbred populations were used.

Under the current circumstances, this approach to gene mapping in *Salmo* is technically difficult. As probes for detecting single-copy RFLPs in genomic DNA, ESTs are labour intensive, time consuming and require large quantities of DNA.

These problems are amplified if variability is low and several restriction enzymes must be screened in order to detect polymorphic loci.

The sequence information obtained from the cDNA clones can still be used to investigate other techniques which may prove more productive. For example, primers to amplify the ApoAI microsatellite, which was described earlier in Section 4.2.2, were developed from the sequence information obtained from the cDNA clones BTLB27 and BTLE42 (Table 5.1). In this case, the location of an intron was inferred from the genomic sequence of the same gene in a different organism. Primers, which anneal to flanking exon sequences, were used to amplify the intron and detect length polymorphisms in the microsatellite. A similar approach can be taken with all of the cDNA clones for which functional genes have been identified. Microsatellites are not expected to be found in every intron examined, but markers such as RFLPs, may be common. Other methods for detecting variability in PCR products are to test for single stranded conformational polymorphisms (SSCPs) (Orita et al. 1989) and heteroduplex analysis (Ganguly et al. 1993), both of which can detect single nucleotide substitutions. Mutations are expected to occur more often in the non-coding, intervening sequences of

genes; however, these methods might be equally informative when amplifying exons. An added advantage to this strategy is the small quantity of DNA required for PCR analysis as well as greater efficiency and reproducibility. This will allow repeated sampling of the same mapping population during linkage analysis. ESTs might also be large enough for FISH analysis of chromosomes (Phillips and Reed 1996). Physical mapping could be performed with two probes simultaniously using different flourescent tags to determine if genes are on the same or different chromosomes.

Mapping with highly informative, non-ambiguous markers, such as ESTs will hopefully provide reference points to align new salmonid maps with those that have been constructed using protein electrophoresis. For example, in brown trout, a creatine kinase gene (*Ck-2*) has been linked to a diaphorase gene (*Dia-1*) through analysis of the variant protein products expressed by each (Johnson et al. 1987). This linkage group could be added to newly discoved groupes which are identified using a creatine kinase EST. Such alignments will be difficult in salmonids since many of the genes which have been previously mapped are multi-locus and their association with particular RFLPs cannot be directly infered. Studies in which both types of markers are examined in the same family are required to resolve this problem.

Chapter 6

Summary

6.1 Evaluation of Genetic Markers

The RAPD technique was selected for its efficiency in detecting large numbers of polymorphic loci. Unfortunately, few genetic markers have been identified using this technique. Of the 893 loci scored in this study, only three and 13 heterozygous loci were identified in the Atlantic salmon and brown trout parents, respectively. Low levels of variability may be the result of inbreeding within natural populations of Atlantic salmon and brown trout in Newfoundland. It is also possible that this technique is not sensitive enough to detect high levels of variability within these species. Added problems of poor reproducibility and the ambiguity of markers will limit the application of this technique for the construction of genetic maps in *Salmo*. It is possible, however, that the conversion of RAPD markers into STSs would improve their reproducibility and possibly allow interspecies comparisons.

Greater progress was made using microsatellites. Approximately half of the loci screened in this study were

selected as markers to perform linkage analysis on two hybrid families. Variant amplification products were reproducible, codominant and easily scored. Loci were held in common by both species, enabling direct comparisons of linkage groups to test for the conservation of linked pairs. As a greater number of microsatellites are discovered in association with genes of known protein function, such as the ApoAI locus, comparisons with more distantly related vertebrate species will be possible.

The ESTs isolated from two cDNA libraries are highly informative, non-ambiguous markers, which can be used to compare genetic maps constructed for *Salmo*, with those of distantly related vertebrates. They might also provide reference points to align new salmonid maps with those that have been constructed using protein electrophoresis. When used as probes to detect RFLPs in the genomic DNA of brown trout and Atlantic salmon, banding patterns were difficult to interpret and low in variability. This process was also time consuming, labour intensive, and required large quantities of DNA. The sequence information obtained from the cDNA clones might be better utilised in more sensitive, PCR-based techniques such as heteroduplex or SSCP analysis. ESTs might also be used as probes for FISH analysis (Moran et al. 1996, Phillips and Reed 1996).

6.2 Evaluation of Hybrid Mapping Population

The strategy behind using hybrid mapping populations was based on the assumption that brown trout and Atlantic salmon are different enough that the haploid contribution of each parent could be easily distinguished and treated separately. This was generally true when working with highly variable microsatellite markers, but was not the case for RAPDs, in which the majority of scored loci were held in common by both species. Analysis of haploid progeny would eliminate this type of interference, but the developmental instability of haploid embryos would limit the number of reactions that can be performed on a particular mapping population. Such a step might not be necessary if a commitment is made to use microsatellites as the principal type of marker for the construction of genetic maps in *Salmo*.

A major advantage of using Atlantic salmon x brown trout hybrids, is that they can be grown to a large size, providing large quantities of tissue for DNA extraction. This enables the repeated testing of individual samples, or the creation of hybrid mapping panels which can be distributed to different laboratories. The latter, will help to increase the rate at which a high density linkage map can be constructed for Atlantic salmon. In addition, by using hybrids, two genomes with divergent karyotypes can be mapped simultaneously and

Table 6.1
Summary of non-random (P<0.001) associations between segregating RAPD and microsatellite markers in Atlantic salmon and brown trout. Predicted order in brown trout groupes III and V is tentative.

I. *Salmo salar* (♀)

Linkage group	Gene order and recombination (%) among loci
I	ApoAI-(16)-SSOSL85
II	SSOSL417-(13)-SSOSL311
III	μ60-(3)-μF43

II. *Salmo trutta* (♂)

I	ApoAI-(3)-SSOSL85
II	SSOSL417-(3)-SSOSL311
III	UBC243:823:*Str*-(9)-UBC348:739:*Str*-(3)-μF43
IV	UBC273:530:*Str*-(9)-UBC354:517:*Str*
V	UBC226:316:*Str*-(6)-UBC239:446:*Str*-(3)-UBC358:371:*Str*
VI	UBC219:432:*Str*-(0)-μ15

compared for linkage conservation.

It should be noted that rates of recombination are going to differ between males and females (May et al. 1980 and 1979). Therefore, it will be important in future projects to screen the reciprocal hybrid cross so that both male and female recombination frequencies are established for each linkage group.

6.3 Summary of Linkage Analysis in *Salmo*

The results of the linkage analysis, which was performed using two Atlantic salmon x brown trout hybrid families, have been summarised in Table 6.1. Three new linkage groups have been identified in Atlantic salmon. All involve single pairs of microsatellite markers. In brown trout, six new linkages were observed which included both microsatellite and RAPD markers. Two of the groups consisted of three loci which have been arranged according to their recombination frequencies.

Of the three linkage groups that were identified in Atlantic salmon, only two have been conserved in brown trout. The third pair of loci may be unlinked in brown trout, or may be separated by a distance which cannot be detected through the analysis of recombination frequencies. Several other loose associations ($0.001<P<0.05$) have also been identified, but will require further analysis, with larger mapping

populations, before they can be confirmed or rejected. These potential linkage groups all involved RAPD markers, prohibiting direct comparisons between the two species. Conversion of RAPD markers to STSs may improve the prospect of finding corresponding linkage groups in both brown trout and Atlantic salmon. Finally, in addition to confirming all of the linkage groups in different mapping populations, it will be important to compare the map distances and recombination rates of both sexes.

REFRENCES

Adams , M.D., Kerlavage, A.R., Fields,C. and Venter, J.C. 1993. 3,400 new expressed sequence tags identify diversity of transcripts in human brain. Nature Genet. 4:256-267.

Adams, M.D., Kelly, J.M., Gocayne, J.D., Dubnick, M.,Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R., and Venter, J.C. 1991. Complementary DNA sequencing: Expressed sequence tags and the human genome project. Science 252:1651-1656.

Ahmed, F.E. and Anderson, R.D. 1994. Fisheries resources, consumption and import trends, and biotechnology developments in the USA. Fish. Res. 19:1-15.

Allendorf, F.W. and Thorgaard, G.H. 1984. Tetraploidy and the evolution of salmonid fishes. *In* Evolutionary Genetics of the Fishes (B.J. Turner ed.). Plenum Press, New York, N.Y. pp 1-53.

Allendorf, R.W., Seeb, J.E., Knudsen, K.L., Thorgaard, G.H. and Leary, R.F. 1986. Gene-centromere mapping of 25 loci in rainbow trout. J. Heredity 77:307-312.

Allendorf, F.W., Gellman, W.A. and Thorgaard, G.H. 1993. Sex-linkage of two enzyme loci in *Oncorhyncus mykiss* (rainbow trout). Heredity 72:498-507.

Andersson, L., Haley, C.S., Ellegren, H., Knott, S.A., Johansson, M., Andersson, K., Andersson-Eklund, L., Edfords-Lilija, I., Fredholm, M., Hansson, I., Hakansson, J. and Lundstrom, K. 1994. Genetic mapping of quantitative trait loci for growth and fatness in pigs. Science 263:1771-1774.

Angers, B. and Bernatchez, L. 1996. Usefulness of heterologous microsatellites obtained from brook charr, *Salvelinus fontinalis* Mitchill, in other *Salvelinus* species. Mol. Ecol.:317-319.

Angers, B., Bernatchez, L., Angers, A. and Desgroseillers, L. 1995. Specific microsatellite loci for brook charr reveal strong population subdivision on a microgeographic scale. J. Fish Biol. 47 supp. A:177-185.

Banfi, S., Borsani, G., Rossi, E., Bernard, L., Guffanti, A., Rubboli, F., Marchitiello, A., Giglio, S., Coluccia, E., Zollo, M., Zuffardi, O. and Ballabio, A. 1996. Identification and mapping of human cDNAs homologous to *Drosophila* mutant genes through EST database searching. Nature Genet. 13:167-174.

Bardakci, F. and Skibinski, D.O.F. 1994. Application of the RAPDtechnique in tilapia fish: species and subspecies identification. Heredity 73:117-123.

Barendse, W., Armitage, S.M., Kossarek, L.M., Shalom, A., Kirkpatrick, B.W., Ryan, A.M., Clayton, D., Li, L., Neibergs, H.L., Zhang, N., Grosse, W.M., Weiss, J., Creighton, P., McCarthy, F., Ron, M., Teale, A.J., Fries, R., McGraw, R.A., Moore, S.S., Georges, M., Soller, M., Womack, J.E., Hetzel, D.J.S. 1994. A genetic linkage map of the bovine genome. Nature Genet. 6:227-235.

Baxendale, S., Abdulla, S., Elgar, G., Buck, D., Berks, M., Micklem, G., Durbin, R., Bates, G., Brenner, S., Beck, S. and Lehrach, H. 1995. Comparative sequence analysis of human and pufferfish Huntington's disease genes. Nature Genet. 10:67-76.

Beckman, J.S. and Weber, J.L. 1992. Survey of human and rat microsatellites. Genomics 12:627-631.

Bitgood, J.J. and Somes, R.G. 1993. Gene map of the chicken (*Gallus gallus* or *G. Domesticus*) 2n = 78. *In* Genetic Maps: Locus Maps of Complex Genomes. 6th edition. (S.J. O'Brien ed.). Cold Spring Harbor Press, Cold Spring Harbor, N.Y. pp. 4.333-4.342.

Boguski, M.S. 1995. The turning point in genome research. TIBS 20:295-296.

Bohonus, V.L., Doolittle, R.F., Pontes, M. and Strong, D.D. 1986. Complementary DNA sequence of lamprey fibrinogen beta-chain. Biochem. 25:6512-6516.

Botstein, D., White, R.L., Skolnick, M. and Davis, R.W. 1980. Construction of a genetic map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32:314-331.

Bradshaw, H.D.Jr., Wilbert, S.M., Otto, K.G. and Schemske, D.W. 1995. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). Nature 376:762-765.

Brenner, S., Elgar, G., Sandford, R., Macrae, A., Venkatesh, B. and Aparicio, S. 1993. Characterization of the pufferfish (Fugu) genome as a compact model vertebrate genome. Nature 366:265-268.

Brooker, A.L., Cook, D., Bentzen, P., Wright, J.M. and Doyle, R.W. 1994. Organization of microsatellites differs between mammals and cold-water teleost fishes. Can. J. Fish. Aquat. Sci. 51:1959-1966.

Broun, P. and Tanksley, S.D. 1996. Characterization and genetic mapping of simple repeat sequences in the tomato genome. Mol. Gen. Genet. 250:39-49.

Brown, S.E. and Knudson, D.L. 1991. 96-well plates for recombinant library maintenance and dot-blot production. BioTechniques 10:719-722.

Bumstead, N. and Palyga, J. 1992. A preliminary linkage map of the chicken genome. Genomics 13:690-697.

Byrnes, L. and Gannon, F. 1990. Atlantic salmon (Salmo salar) serum albumin: cDNA sequence, evolution and tissue expression. DNA Cell Biol. 9:647-655.

Burt, D.W., Bumbstead, N., Bitgood, J.J., Ponce de Leon, F.A. and Crittenden, L.B. 1995. Chicken genome mapping: a new era in avian genetics. TIG 11:190-194.

Cai, Q., Guy, C.L. and Moore, G.A. 1994. Extension of the linkage map in *Citrus* using random amplified polymorphic DNA (RAPD) markers and RFLP mapping of cold-acclimation-responsive loci. Theor. Appl. Genet. 89:606-614.

Causse, M.A., Fulton, T.M., Cho, Y.G., Ahn, S.N., Chunwongse, J., Wu, K.S., Xiao, J.H., Yu, Z.H., Ronald, P.C., Harrington, S.E., Second, G., McCouch, S.R. and Tanksley, S.D. 1994. Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138:1251-1274.

Chaparro, J.X., Werner, D.J., O'Malley, D. and Sederoff, R.R. 1994. Targeted mapping and linkage analysis of morphological isozyme, and RAPD markers in peach. Theor. Appl. Genet. 87:807-815.

Cheng, H.H., Levin, I., Vallejo, R.L., Khatib, H., Dodgson, J.B., Crittenden, L.B., Hillel, J. 1995. Development of a genetic-map of the chicken with markers of high utility. Poultry Sci. 74:1855-18874.

Chevassus, B., Devaux, A., Chourrout, D. and Jalabert, B. 1988. Production of YY rainbow trout by self fertilization of induced hermaphrodites. J. Heredity 79:89-92

Chomczynski, P. and Sacchi, N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162:156-159.

Coe, E.H. and Neuffer, M.G. 1993. Gene loci and linkage map of corn (maize)(Zea mays L.)(2N = 20). *In* Genetic Maps: Locus Maps of Complex Genomes. 6th edition. (S.J. O'Brien ed.). Cold Spring Harbor Press, Cold Spring Harbor, N.Y. pp. 6.157-6.189.

Colbourne, J.K., Neff, B.D., Wright, J.M. and Gross, M.R. 1996. DNA fingerprinting of bluegill sunfish (*Lepomis macrochirus*) using (GT)n microsatellites and its potential for assessment of mating success. Can. J. Fish. Aquat. Sci. 53:342-349.

Crawford, A.M., Dodds, K.G., Ede, A.J., Pierson, C.A., Montgomery, G.W., Garmonsway, H.G., Beattie, A.E., Davies, K., Maddox, J.F., Kappes, S.W., Stone, R.T., Nguyen, T.C., Penty, J.M., Lord, E.A., Broom, J.E., Buitkamp, J., Schwaiger, W., Epplen, J.T., Matthew, P., Matthews, M.E., Hulme, D.J., Beh, K.J., McGraw, R.A. and Beattie, C.W. 1995. An autosomal genetic linkage map of the sheep genome. Genetics 140:703-724.

Crooijmans, R.P.M.A., Van Oers, P.A.M., Strijk, J.A., Van der Poel, J.J. and Groenen, M.A.M. 1996. Preliminary linkage map of the chicken (*Gallus domesticus*) genome based on microsatellite markers: 77 new markers mapped. Poultry Sci. 75:746-754.

Das Gupta, R., Morrow, B., Marondel, I., Parimoo, S., Goei, V.L., Gruen, J., Weissman, S., Skoultchi, A. and Kucherlapati, R. 1993. An integrated approach for identifying and mapping human genes. Proc. Natl. Acad. Sci. USA 90:4364-4368.

Davidson, W.S., Birt, T.P., Green, J.M. 1989. A review of genetic variation in Atlantic salmon, *Salmo salar* L., and its importance for stock identification, enhancement programmes and aquaculture. J. Fish Biol. 34:547-560.

Davisson, M.T., Wright, J.E. and Atherton, L.M. 1973. Cytogenetic analysis of pseudolinkage of LDH loci in the teleost genus *Salvelinus.* Genetics 73:645-658.

Davies, K. 1993. The EST express gathers speed. Nature Genet. 364:554.

Davis, L.G., Dibner, M.D. and Battey, J.F. 1986. Basic methods in molecular biology. Elsevier, New York.

Devlin, R.H., McNeil, B.K., Groves, T.D.D., and Donaldson, E.M. 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.

Devlin, R.H. 1993. Sequence of sockeye salmon type 1 and 2 growth hormone genes and the relationship of rainbow trout with Atlantic and Pacific salmon. Can. J. Fish. Aquat. Sci. 50:1738-1748.

Devlin, R.H., McNeil, B.K., Solar, I.I. and Donaldson, E.M. 1994. A rapid PCR-based test for Y-chromosomal DNA allows simple production of all-female strains of chinook salmon. Aquaculture 128:211-220.

Devos, K.M. and Gale, M.D. 1992. The use of random amplified polymorphic DNA markers in wheat. Theor. Appl. Genet. 84:567-572.

Dib, C., Faures, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J., Weissenbach, J. 1996. A comprehensive genetic-map of the human genome based on 5,264 microsatellites. Nature 380:152-154.

Dietrich, W.F., Miller, J.C., Steen, R.G., Merchant, M., Damron, D., Nahf, R., Gross, A., Joyce, D.C., Wessel, M., Dredge, R.D., Marquis, A., Stein, L.D., Goodman, N., Page, D.C., Lander, E.S. 1994. A genetic map of the mouse with 4,006 simple sequence length polymorphisms. Nature Genet. 7:220-245.

Dietrich, W.F., Miller, J.C., Steen, R.G., Merchant, M., Damron-Boles, D., Husain, Z., Dredge, R.D., Daly, M.J., Ingalls, K.A., O'Connor, T.J., Evans, C.A., Deangelis, M.M., Levinson, D.M., Kruglyak, L., Goodman, N., Copeland, N.G., Jenkins, N.A., Hawkins, T.L., Stein, L., Page, D.C. and Lander, E.S. 1996. A comprehensive genetic-map of the mouse genome. Nature 380:149-152.

Driever, W., Stemple, D., Schier, A., Solinca-Krezel, L. 1994. Zebrafish: genetic tools for studying vertebrate development. TIG 10:152-159.

Drover, V.A.B. 1996. Sequence and Evolution of Apolipoprotein AI in three species of salmonids. MSc. thesis, Department of Biochemistry, Memorial University of Newfoundland.

Du, S.J., Devlin, G.H. and Hew, C.L. 1993. Genomic structure of growth hormone genes in chinook salmon (*Oncorhyynchus tshawytscha*); presence of two functional genes, GH-1 and GH-II, and a male-specific pseudogene, GH-*. DNA Cell Biol. 12:739-751.

Dubcovsky, J., Luo, M., Zhong, G., Bransteitter, R., Desai, A., Kilian, A., Kleinhofs, A. and Dvorak, J. 1996. Genetic map of diploid wheat, *Triticum monococcum* L., and its comparison with maps of *Hordeum vulgare* L. Genetics 143:983-999.

Dubovsky, J., Sheffield, V.C., Duyk, G.M. and Weber, J.L. 1995. Set of short tandem repeat polymorphisms for efficient linkage screening of the human genome. Hum. Mol. Genet. 4:449-452.

Durkin, A.S., Maglott, D.R. and Nierman, W.C. 1992. Chromosomal assignment of 38 human brain expressed sequence tags (ESTs) by analyzing fluorescently labelled PCR products from hybrid cell panels. Genomics 14:808-810.

Ebert II, R.H., Cherkasova, V.A., Dennis, R.A., Wu, J.H., Ruggles, S., Perrin, T.E. and Shmookler Reis, R.J. 1993. Longevity-determining genes in Caenorhabditis elegans: chromosomal mapping of multiple noninteractive loci. Genetics 135:1003-1010.

Echard, G., Yerle, M., Milan, D., Gellin, J. and Archibald, A.L. 1993. The gene map of the pig (*Sus scrofa domestica*) 2N = 38. *In* Genetic Maps: Locus Maps of Complex Genomes. 6th edition. (S.J. O'Brien ed.). Cold Spring Harbor Press, Cold Spring Harbor, N.Y. pp. 4.241-4.249.

Elgar, G., Sanford, R., Aparicio, S., MacRae, A., Venkatesh, B. and Brenner, S. 1996. Small is beautiful: comparative genomics with the pufferfish (*Fugu rubripes*). TIG 12:145-150.

Ellegren, H., Chowdhary, B.P., Johansson, M., Marklund, L., Fredholm, M., Gustavsson, I. and Andersson, L. 1994. A primary linkage map of the porcine genome reveals a low rate of genetic recombination. Genetics 137:1089-1100.

Eng, C.M., Durtschi, B.A., Zoghbi, H.Y. and Beaudet, A.L. 1992. Isolation, mapping, and characterization of two cDNA clones expressed in the cerebellum. Genomics 14:813-815.

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.

Feist, G., Yeoh, C.G., Fitzpatrick, M.S. and Schreck, C.B. 1995. The production of functional sex-reversed male rainbow trout with 17a-methyltestosterone and 11b-hydroxyandrostenedione. Aquaculture 131:145-152.

Ferguson, A. 1989. Genetic differences among brown trout, Salmo trutta, stocks and their importance for the conservation and management of the species. Freshw. Biology 21:35-46.

Ferguson, M. 1994. The role of molecular genetic markers in the management of cultured fishes. Reviews in Fish Biology and Fisheries 4:351-373.

Ferreiro, C., Medrando, J.F. and Gall, G.A.E. 1989. Genome analysis of rainbow trout and sturgeon with restriction enzymes and hybridization with a ZFY gene probe to identify sex. Aquaculture 81:245-251.

Forbes, S.H., Knudsen, K.L., North, T.W. and Allendorf, F.W. 1994. One of two growth hormone genes in coho salmon is sex-linked. Proc. Natl. Acad. Sci. USA 91:1628-1631.

Fornzler, D., Wittbrodt, J. and Schartl, M. 1991. Analysis of an esterase linked to a locus involved in the regulation of the melanoma oncogene and isolation of polymorphic marker sequences in Xiphophorus. Biochemical Genetics 29:509-524.

Foyle, T.P. 1993. A histological description of gonadal development and sex differentiation in the coho salmon (*Oncorhynchus kisutch*) for both untreated and oestradiol immersed fry. J. Fish. Bio. 42:699-712.

Francis, R.C. 1992. Sexual lability in teleosts: developmental factors. Q. Rev. Biol. 67:1-18.

Ganguly, A., Rock, M.J. and Prockop, D.J. 1993. Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes. Proc. Natl. Acad. Sci. USA. 90:10325-10329.

Garbner, A.T., Winkfein, R.J. and Dixon, G.H. 1990. Biochim. Biophys. Acta. 1087:256-258.

Garcia-Dorado, A. and Gallego, A. 1992. On the use of classical tests for detecting linkage. J. Heredity 83:143-146.

Gardiner, J.M., Coe, E.H., Meliahancock, S., Hoisington, D.A. and Chao, S. 1993. Development of a core RFLP map in maize using an immortalized-F2 population. Genetics 134:917-930.

Gellman, W.A., Allendorf, F.W. and Thorgaard, G.H. 1987. Hexoamininase is sex linked in rainbow trout. Isozyme Bull. 20:14.

Georges, M., Nielson, D., Mackinnon, M., Mishra, A., Okimoto, R., Pasquino, A.T., Sargeant, L.S., Sorensen, A., Steele, M.R., Zhao, X., Womack, J.E. and Hoeschele, I. 1995. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. Genetics 139:907-920.

Goff, D.J., Galvin, K., Katz, H., Westerfield, M., Lander, E.S. and Tabin, C.J. 1992. Identification of polymorphic simple sequence repeats in the genome of the zebrafish. Genomics 14:200-202.

Goodier, J.L. and Davidson, W.S. 1993. Gene mapping in fish. In Biochemistry and Molecular Biology of Fishes, vol 2. (P.W. Hochachka and T.P. Mommsen eds.). Elsevier, pp 93-12.

Gottesman, M.M. and Collins, F.S. 1994. The role of the human genome project in disease prevention. Prev. Med. 23:591-594.

Grattapaglia, D., Bertolucci, F.L. and Sederoff, R.R. 1995. Genetic mapping of QTLs controlling vegetative propagation in Eucalyptus grandis and E. urophylla using a pseudo-testcross strategy and RAPD markers. Theor. Appl. Genet. 90:933-947.

Grattapaglia, D. and Sederoff, R. 1994. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137:1121-1137.

Griffiths, R. and Tiwari, B. 1993. The isolation of molecular genetic markers for the identification of sex. Proc. Natl. Acad. Sci. USA 90:8324-8326.

Grivet, L., D'hont, A., Dufour, P., Hamon, P., Roques, D. and Glaszmann, J.C. 1994. Comparative genome mapping of sugar cane with other species within the Andropogoneae tribe. Heredity 73:500-508.

Guilford, P., Dode, C., Crozet, F., Blanchard, S., Chaib, H., Levilliers, J., Levi-Acobas, F., Weil, D., Weissenbach, J., Cohen, D., Le Paslier, D., Kaplan, J. and Petit, C. 1995. A YAC contig and an EST map in the pericentromeric region of chromosome 13 surrounding the loci for neurosensory nonsyndromic deafness (DFNB1 and DFNA3) and limb-girdle muscular distrophy type 2C (LGMD2C). Genomics 29:163-169.

Guyer, M.S. and Collins, F.S. 1995. How is the human genome project doing, and what have we learned so far? Proc. Natl. Acad. Sci. USA 92:10841-10848.

Gyapay, G., Morrissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M. and Weissenbach, J. 1994. The 1993-94 Genethon human genetic linkage map. Nature Genet. 7:246-339.

Gyllensten, U. and Wilson, A.C. 1987. Mitochondrial DNA of salmonids: inter- and intraspecific variability detected with restriction enzymes. In Population Genetics and Fisheries Management (N. Ryman and F. Utter eds.). University of Washington Press, Seattle , W.A. pp. 301-317.

Hadrys, H., Balick, M. and Schierwater, B. 1992. Application of random amplified polymorphic DNA (RAPD) in molecular ecology. Mol. Ecol. 1:55-63.

Hamada, H., Petrino, M.G. and Kanunaga, T. 1982. A novel repeat element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc. Natl. Acad. Sci. USA. 79;6465-6469.

Hartley, S.E. 1987. The chromosomes of salmonid fishes. Biol. Rev. 62:197-214.

Hearne, C.M., Ghosh, S. and Todd, J.A. 1992. Microsatellites for linkage analysis of genetic traits. TIG 8:288-294.

Heeley, D.H., Bieger, T., Waddleton, D.M., Hong, C., Jackman, D.M., McGowan, C., Davidson, D. and Beavis, R.C. 1995. Characterisation of fast, slow and cardiac muscle tropomyosins from salmonid fish. Eur. J. Biochem. 232:226-234.

Hodgkin, J., Plasterk, R.H.A. and Waterston, R.H. 1995. The nematode *Caenorhabditis elegans* and its genome. Science 270:410-414.

Hoffman, E.P. 1994. The evolving genome project: current and future impact. Am. J. Hum. Genet. 54:129-136.

Hombergen, E.J. and Bachmann, K. 1995. RAPD mapping of three QTLs determining trichome formation in *Microseris* hybrid H27 (Asteraceae:Lactuceae). Theor. Appl. Genet. 90:853-858.

Hormaza, J.I., Dollo, L. and Polito, V.S. 1994. Identification of a RAPD marker linked to sex determination in *Pistacia vera* using bulked segregant analysis. Theor. Appl. Genet. 89:9-13.

Hunt, G.J., Page, R.E.Jr., Fondrk, M.K. and Dullum, C.J. 1995 Major quantitative trait loci affecting honey bee foraging behavior. Genetics 141:1537-1545.

Hunter, G.A., Donaldson, E.M., Goetz, F.W. and Edgell, P.R. 1982. Production of all-female and sterile coho salmon, and experimental evidence for male heterogamety. Trans. Am. Fish. Soc. 111:367-372.

Johansson, M., Ellegren, H. and Andersson, L. 1995. Comparative mapping reveals extensive linkage conservation - but with gene order rearrangements - between the pig and the human genomes. Genomics 25:682-690.

Johnson, K.R. and Wright, J.E.Jr. 1986. Female brown trout x Atlantic salmon hybrids produce gynogens and triploids when backcrossed to male Atlantic salmon. Aquaculture 57:345-358.

Johnson, K.R. and Wright, J.E.Jr. 1987. Linkage relationships reflecting ancestral tetraploidy in salmonid fish. Genetics 116:579-591.

Johnson, S.J., Gates, M.A., Johnson, M., Talbot, W.S., Horne, S., Baik, K., Rude, S., Wong, J.R. and Postlethwait, J.H. 1996. Centromere-linkage analysis and consolidation of the zebrafish genetic linkage map. Genetics 142:1277-1288.

Johnstone R., Simpson, T.H., Youngson, A.F. and Whitehead, C. 1979. Sex reversal in salmonid culture. Part II. The progeny of sex reversed rainbow trout. Aquaculture 18:13-19.

Johnstone, R., Simpson, T.H. and Youngson, A.F. 1978. Sex reversal in salmonid culture. Aquaculture 13:115-134.

Johnstone, R. and Youngson, A.F. 1984. The progeny of sex inverted female Atlantic salmon (*Salmo salar* L.). Aquaculture 37:179-182.

Kauffman, E.J., Gestl, E.E., Kim, D.J., Walker, C., Hite, J.M., Yan, G., Rogan, P.K., Johnson, S.L., Cheng, K.C. 1995. Microsatellite-centromere mapping in the zebrafish (*Danio rerio*). Genomics 30:337-341.

Kawaga, H., Sugimoto, K., Matsumoto, H., Inoue, T., Imadzu, H., Takuwa, K. and Sakube, Y. 1995. Genome structure, mapping and expression of the tropomyosin gene tmy-1 of *Caenorhabditis elegans*. J. Mol. Biol. 251:603-613.

Keightley, P.D., Hardge, T., May, L. and Bulfield, G. 1996. A genetic map of quantitative trait loci for body weight in the mouse. Genetics 142:227-235.

Kiss, G.B., Csanadi, G., Kalman, K., Kalo, P. and Okresz, L. 1993. Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. Mol. Gen. Genet. 238:129-137.

Kjaer, B., Jensen, J. and Giese, H. 1995. Quantitative trait loci for heading date and straw characters in barley. Genome 38:1098-1104.

Koop, B.F. and Nadeau, J.H. 1996. Pufferfish and a new paradigm for comparative genome analysis. Proc. Natl. Acad. Sci. USA 93:1363-1365.

Kubisiak, T.L., Nelson, C.D., Nance, W.L. and Stine, M. 1995. RAPD linkage mapping in a longleaf pine x slash pine F1 family. Theor. Appl. Genet. 90:1119-1127.

Lande and Thompson, R. 1990. Efficiency of marker assisted selection in the improvement of quantitative traits. Genetics 124:743-756.

Langridge, P., Karakousis, A., Collins, N., Kretschmer, J. and Mannings, S. 1995. A consensus linkage map of barley. Molecular Breeding 1:389-395.

Leung, F.C., Welt, M., Quesenberry, R.D. and Shen, X.Z. 1994. DNA-fingerprinting and cloning of hypervariable mini-satellite repeats in salmonids. Can. J. Fish. Aquat. Sci. 51 supp.1:258-266.

Levin, I., Crittenden, L.B. and Dodgson, J.B. 1993. Genetic map of the chicken Z chromosome using random amplified polymorphic DNA (RAPD) markers. Genomics 16:224-230.

Levinson, G. and Gutman, G.A. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol. and Evol. 4:203-221.

Lie , O., Slettan, A., Grimholt, U., Lundin, M., Syed, M., Olsaker, I. 1994a. Fish gene maps and their implications for aquaculture. Anim. Biotech. 5:209-218.

Lie, O., Slettan, A., Lingaas, F., Olsaker, I., Hordvik, I. and Refstie, T. 1994b. Haploid gynogenesis: a powerful strategy for linkage analysis in fish. Anim. Biotech. 5:33-45

Lin, J. and Ritland, K. 1996. Construction of a genetic linkage map in the wild plant *Mimulus* using RAPD and isozyme markers. Genome 39:63-70.

Lloyd, M.A., Fields, M.J. and Thorgaard, G.H. 1989. BKm minisatellite sequences are not sex associated but reveal DNA fingerprint polymorphisms in rainbow trout. Genome 32:865-868.

Lu, Q., Wallrath, L.L., Granok, H. and Elgin, S.C. 1993. (CT)n/(GA)n repeats and heat shock elements have distinct roles in chromatin structure and transcriptional activation of Drosophila hsp26 gene. Mol. Cell. Biol. 13:2802-2814.

Lynch, M. and Milligan, B.G. 1994. Analysis of population genetic structure with RAPD markers. Mol. Ecol. 3:91-99.

Maheshwar, M.M., Sanford, R., Nellist, M., Cheadle, J.P., Sgotto, B., Vaudin, M., Sampson, J.R. 1996. Comparative analysis and genomic structure of the tuberous sclerosis 2 (TSC2) gene in human and pufferfish. Hum. Mol. Genet. 5:131-137.

Mahtani, M.M. and Willard, H.F. 1993. A polymorphic X-linked tetranucleotide repeat locus displaying a high rate of new mutation: implications for mechanisms of mutation at short tandem repeat loci. Hum. Mol. Genet. 2:431-437.

Maihara, T., Andoh, Y., Yokoi, N., Kuramoto, T. and Serikawa, T. 1995. Fifty-six new microsatellite markers in the rat genetic map. Transplant. Proc. 27:1502-1504.

Mather, K. 1951. The Measurement of Linkage in Heredity. Methuen, London.

May, B., Johnson, K.R. and Wright, J.E. 1989. Sex linkage in Salmonids: evidence from a hybridized genome of brook trout and Arctic charr. Biochem. Gen. 27:291-301.

May, B. and Johnson, K.R. 1993. Composite linkage map of salmonid fishes (*Salvelinus, Salmo, Oncorhynchus*) *In* Genetic Maps: Locus Maps of Complex Genomes. 6th edition. (S.J. O'Brien ed.). Cold Spring Harbor Press, Cold Spring Harbor, N.Y. pp. 4.309-4.317.

May, B.,Stoneking, M. and Wright, J.E.Jr. 1980. Joint segregation of biochemical loci in Salmonidae II. Linkage associations from a hybridized *Salvelinus* genome (*S. namaycush* x *S. fontinalis*). Genetics 95:707-726.

May, B., Wright, J.E.Jr. and Stoneking, M. 1979. Joint segregation of biochemical loci in Salmonidae: results from experiments with Salvelinus and review of the literature on other species. J. Fish. Res. Board Can. 36:1114-1128.

Mazzarella, R. and Srivastava, A.K. 1994. Physical linkage of expressed sequence tags (ESTs) to polymorphic markers on the X chromosome. Hum. Mol. Genet. 3:1095-1101.

McConnell, S., O'Reilly, P., Hamilton, L., Wright, J.M. and Bentzen, P. 1995a. Polymorphic microsatellite loci from Atlantic salmon (*Salmo salar*): genetic differentiation of North America and European populations. Can. J. Fish. Aquat. Sci. 52:1863-1872.

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

McGowan, C. 1992. Aspects of Natural and Artificial Hybridization Between Brown Trout (Salmo trutta) and Atlantic Salmon (Salmo salar) in Newfoundland. M.Sc. thesis. Memorial University of Newfoundland.

McGowan, C. and Davidson, W.S. 1992. Hybridization between brown trout and Atlantic salmon; hatchability, survival and growth up to first feeding. Aquaculture 106:117-125.

McVeigh, H.P., Bartlett, S.E. and Davidson, W.S. 1991. Polymerase chain reaction/direct sequence analysis of the cytochrome b gene in *Salmo* species. Aquaculture 95:225-233.

Melo, J.A., Shendure, J., Pociask, K. and Silver, L.M. 1996. Identification of sex-specific quantitative trait loci controlling alcohol preference in C57BL/6 mice. Nature Genet. 13:147-153.

Michelmore, R.W., Paran, I. and Kesseli, R.V. 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.

Miklas, P.N., Afanador, L. and Kelly, J.D. 1996. Recombination-facilitated rapd marker-assisted selection for disease resistance in common bean. Crop Sci. 36:86-90.

Morris, D.B., Richard, K.R. and Wright, J.M. 1996. Microsatellites from rainbow trout (*Oncorhynchus mykiss*) and their use for genetic study of salmonids. Can. J. Fish. Aquat. Sci. 53:120-126.

Morrison, W.J. 1970. Non random segregation of two lactate dehydrogenase subunit loci in trout. Trans. Amer. Fish. Soc. 99:193-206.

Mitchell-Olds, T. 1996. Genetic constraints on life-history evolution: quantitative-trait loci influencing growth and flowering in *Arabidopsis thaliana*. Evolution 50:140-145.

Morizot, D.C. 1990. Use of fish gene maps to predict ancestral vertebrate genome organization. In Isozymes: Structure, Function and Use in Biology and Medicine (Z.I. Ogita and C.L. Markert eds.). Liss-Wiley, N.Y. pp. 207-234.

Morizot, D.C., Slaugenhaupt, S.A., Kailman, K.D. and Chakravarti, A. 1991. Genetic linkage map of fishes of the genus Xiphophorus (Teleostei:Poeciliidae). Genetics 127:399-410.

Morizot, D.C., Harless, J., Nairn, R.S., Kallman, K.D. and Walter, R.B. 1993. Linkage maps of non-salmonid fishes. *In* Genetic Maps: Locus Maps of Complex Genomes. 6th edition. (S.J. O'Brien ed.). Cold Spring Harbor Press, Cold Spring Harbor, N.Y. pp. 4.318-4.325.

Morizot, D.C., Schmidt, M.E. and Carmichael, G.J. 1994. Joint segregation of allozymes in catfish genetic crosses: designation of Ictalurus punctatus linkage group I. Trans. Am. Fish. Soc. 123:22-27.

Mouzeyar, S., Roeckel-Drevet, P., Gentzbittel, L., Philippon, J., Tourvieille de Labrouhe, D., Vear, F. and Nicolas, P. 1995. RFLP and RAPD mapping of the sunflower *P/1* locus for resistance to *Plasmopara halstedii* race 1. Thoer. Appl. Genet. 91:733-737.

Naish, K.A., Warren, M., Bardakci, F., Skibinski, D.O.F., Carvalho, G. R. and Mair, G.C. 1995. Multilocus DNA fingerprinting and RAPD reveal similar genetic relationships between strains of *Oreochromis niloticus* (Pisces: Cichlidae). Mol. Ecol. 4:271-274.

Nakamura, D. 1989. Evolutionary aspects of vertibrate sex determination. *In* Evolutionary Mechanisms in Sex Determination (S. S. Wachtel ed.). CRC Press, Boca Raton, FA. pp 3-14.

Nakayama, I., Foresti, F., Tewari, R., Schartl, M. and Chourrout, D. 1994. Sex chromosome polymorphism and heterogametic males revealed by two cloned DNA probes in the ZW/ZZ fish *Leporinus elongatus*. Chromos. 103:31-39.

Nairn, R.S., Coletta, L.D., McEntire, B.B., Walter, R.B. and Morizot, D.C. 1996. Assignment of the *TP53* orthologue to a new linkage group (LG XIV) in fish of the genus *Xiphophorous* (Teleostei:Poeciliidae). Cancer Genet. Cytogenet. 88:144-150.

Nanda, I., Schartl, M., Feichtinger, W., Epplen, J.T. and Schmid, M. 1992. Early stages of sex chromosome differentiation in fish as analysed by simple repeditive DNA sequences. Chromos. 101:301-310.

Narine, K.M., Kallman, K.D., Walter, R.B. and Morizot, D.C. 1992. Linkage of two enzyme loci in the fish genus *Poecilia*(Teleostei:Poeciliidae). Cytogenet. Cell Genet. 61:75-77.

Naylor, L.H. and Clark, E.M. 1990. d(TG)n/d(CA)n sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription. Nuc. Acids. Res. 18:1595-1601.

Nilsen, I.W., Bakke, I., Vader, A., Olsvik, O. and El-Gewely, M.R. 1996. Isolation of *cmr*, a novel Escherichia coli chloramphenicol resistance gene encoding a putative efflux pump. J. Bact. 178:3188-3193.

O'Brien, S. J., Womack, J.E., Lyons, L.A., Moore, K.J. Jenkins, N.A. and Copeland, N.G. 1993. Anchore reference loci for comparative genome mapping in mammals. Nature Genet. 3:103-112.

O'Brien, S.J. (ed). 1993. Genetic Maps: Locus Maps of Complex Genomes. 6th edition. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.

Ohara, O., Dorit, R.L. and Gilbert, W. 1989. One-sided polymerase chain reaction: the amplification of cDNA. Proc. Natl. Acad. Sci. USA. 86:5673-5677.

Ohno, S. 1970. Evolution by Gene Duplication. Springer-Verlag, New York, N.Y.

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

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. and Sekiya, T. 1989. Detection of polymorphisms of human DNA by gel electrophoresis and single-strand conformational polymorphisms. Proc. Natl. Acad. Sci. USA 86:2766-2770.

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

Park, Y.S., Kwak, J.M., Kwon, O., Kim, Y.S., Lee, D.S., Cho, M.J., Lee, H.H. and Nam, H.G. 1993. Generation of expressed sequence tags of random root cDNA clones of Brassica napus by single-run partial sequencing. Plant Physiol. 103:359-370.

Pawlak, A., Toussaint, C., Levy, I., Bulle, F., Poyard, M., Barouki, R. and Guellaen, G. 1995. Characterization of a large population of mRNAs from human testis. Genomics 26:151-158.

Pearson, W.R. and Lipman, D.J. 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85:2444-2448.

Peltier, D., Farcy, E., Dulieu, H. and Berville, A. 1994. Origin, distribution and mapping of RAPD markers from wild *Petunia* species in *Petunia hybrida* Hort lines. Theor. Appl. Genet. 88:637-645.

Pendas, A.M., Moran, P. and Garcia-Vazquez, E. 1995. Physical mapping of reiterated gene families in salmonids. Aquaculture 137:99-100.

Pendas, A.M., Moran, P., Freije, J.P. and Garcia-Vazquez, E. 1994. Chromosomal mapping and nucleotide sequence of two tandem repeats of Atlantic salmon 5S rDNA. Cytogenet. Cell Genet. 67:31-36.

Pendas, A.M., Moran, P. and Garcia-Vazquez, E. 1993. Replication banding patterns in Atlantic salmon *(Salmo salar)*. Genome 36:440-444.

Phillips, R.B. and Ihssen, P.E. 1985. Identification of sex chromosomes lake trout (*Salvelinus namaycush*). Cytogenet. Cell Genet. 39:14-18.

Phillips, R.B. and Reed, K.M. 1996. Application of flourescence in-situ hybridization (FISH) techniques to fish genetics - a review. Aquaculture 140:197-216.

Piferrer, F., Zanuy, S., Carrillo, M., Solar, I.I., Devlin, R.H.and Donaldson, E.M. 1994. Brief treatment with an aromatase inhibitor during sex differentiation causes chromosomally female salmon to develop as normal, functional males. J. Exp. Zoo. 270:255-262.

Plomion, C., Bahrman, N., Durel, C.E. and O'Malley, D.M. 1995. Genomic mapping in *Pinus pinaster* (maritime pine) using RAPD and protein markers. Heredity 74:661-668.

Pogson, G.H., Mesa, K.A. and Boutilier, R.G. 1995. Genetics population structure and gene flow in the Atlantic cod *Gadus morhua*: a comparison of allozyme and nuclear RFLP loci. Genetics 139:375-385.

Pollard, S.M., Danzmann, R.G. and Claytor, R.R. 1994. Association between the regulatory locus *PGM-1r** and life-history types of juvenile Atlantic salmon (*Salmo salar*). Can. J. Fish. Aquat. Sci. 51:1322-1329.

Postlewait, J.H., Johnson, S.L., Midson, C.N., Talbot, W.S., Gates, M., Ballinger, E.W., Africa, D., Andrews, R., Carl, T., Eisen, J.S., Horn, S., Kimmel, C.B., Hutchinson, M., Johnson, M. and Rodriguez, A. 1994. A genetic map for the zebrafish. Science 264:699-703.

Powell, R., Higgins, D.G., Wolff, J., Byrnes, L., Stack, M., Sharp, P.M. and Gannon, F. 1991. The salmon gene encoding apolipoprotein A-I: cDNA sequence, tissue expression and evolution. Gene 104:155-161.

Prince, J.P., Pochard, E. and Tanksley, S.D. 1993. Construction of a molecular linkage map of pepper and a comparison of synteny with tomato. Genome 36:404-417.

Prodohl, P.A., Taggart, J.B. and Ferguson, A. 1994. Single locus inheritance and joint segregation analysis of minisatellite (VNTR) DNA loci in brown trout (*Salmo trutta* L.). Heredity 73:556-566.

Promboon, A., Shimada, T., Fujiwara, H. and Kobayashi, M. 1995. Linkage map of random amplified polymorphic DNAs (RAPDs) in the silkworm, *Bombyx mori*. Gen. Res. 66:1-7.

Ogura, K., Wicky, C., Magnenat, L., Tobler, H., Mori, I., Muller, F. and Ohshima, Y. 1994. Caenorhabditis elegans unc-51 gene required for axonal elongation encodes a novel serine/threonine kinase. Gen. Dev. 8:2389-2400.

Rajapakse, S., Belthoff, L.E., He, G., Estager, A.E., Scorza, R., Verde, I., Ballard, R.E., Baird, W.V., Callahan, A., Monet, R. and Abbott, A.G. 1995. Genetic linkage mapping in peach using morphological, RFLP and RAPD markers. Theor. Appl. Genet. 90:503-510.

Rassman, K. and Schlotterer, C. and Tautz, D. 1991. Isolation of simple-sequence loci for use in polymerase chain reaction DNA fingerprinting. Electrophoresis 12:113-118.

Reiter, R.S., Williams, J.G.K., Feldmann, K.A., Rafalski, J.A. and Tingey, S.V. 1992. Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbreed lines and random amplified polymorphic DNAs. Proc. Natl. Acad. Sci. USA 89:1477-1481.

Richards, R.I. and Sutherland, G.R. 1994. Simple repeat DNA is not replicated simply. Nature Genet. 6:114-116.

Rico, C., Rico, I. and Hewitt, G. 1996. 470 million years of conservation of microsatellite loci among fish species. Proc. R. Soc. Lond. B 263:549-557.

Rico, C., Zadworny, d., Kuhnlein, U. and FitzGerald, J.G. 1993. Characterisation of microsatellite loci in the three spine stickleback *Gasterosteus aculeatus*. Mol. Ecol. 2:271-272.

Rieseberg, L.H. 1996. Homology among RAPD fragments in interspecific comparisons. Mol. Ecol. 5:99-105.

Rohrer, G.A., Alexander, L.J., Keele, J.W., Smith, T.P. and Beattie, C.W. 1994. A microsatellite linkage map of the porcine genome. Genetics 136:231-245.

Rossant, J. and Hopkins, N. 1992. Of fin and fur: mutational analysis of vertebrate embryonic development. Gen. Dev. 6:1-13.

Rubin, E.M. and Barsh, G.S. 1996. Biological insights through genomics: mouse to man. J. Clin. Invest. 97:275-280.

Rubinsztein, D.C., Amos, W., Leggo, J., Goodburn, S., Jain, S., Li, S., Margolis, R.L., Ross, C.A. and Ferguson-Smith, M.A. 1995a. Microsatellite evolution - evidence for directionality and variation in rate between species. Nature Genet. 10:337-343.

Rubinsztein, D.C., Leggo, J. and Amos, B. 1995b. Microsatellites evolve more rapidly in humans than in chimpanzees. Genomics 30:610-612.

Sabo, T.J., Kesseli, R., Halverson, J.L., Nisbet, I.C.T. and Hatch, J.J. 1994. PCR-based method for sexing roseate terns (*Sterna dougallii*). The Auk 111:1023-1027.

Sanchez, J.A., Clabby, C., Ramos, D., Blanco, G., Flavin, F., Vazquez, E. and Powell, R. 1996. Protein and microsatellite single locus variability in Atlantic salmon (*Salmo salar* L.). Heredity (in press).

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.

Sall, T. and Nilsson, N. 1994. The robustness of recombination frequency estimates in intercrosses with dominant markers. Genetics 137:589-596.

Sambrook, J., Fritsh, E.F. and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schartl, M. 1995. Platyfish and swordtails: a genetic system for the analysis of molecular mechanisms in tumor formation. TIG 11:185-189.

Schlotterer, C., Amos, B. and Tautz, D. 1991. Conservation of polymorphic simple sequence loci in cetacean species. Nature 354:63-65.

Schmid, M. and Haaf, T. 1989. Origin and evolution of sex chromosomes in amphibia: the cytogenetic data. *In* Evolutionary Mechanisms in Sex Determination (S. S. Wachtel ed.). CRC Press, Boca Raton, FA. pp 37-56.

Scribner, K.T., Gust, J.R. and Fields, R.L. 1996. Isolation and characterization of novel salmon microsatellite loci: cross-species amplification and population genetic applications. Can. J. Fish. Aquat. Sci. 53:833-841.

Seeb, J.E. and Seeb, L.W. 1986. Gene mapping of isozyme loci in chum salmon. J. Heredity 77:399-402.

Serikawa, T., Kuramoto, T., Hilbert, P., Mori, M., Yamada, J., Dubay, C.J., Lindpainter, K., Ganten, D., Guenet, J., Lathrop, G.M. and Beckmann, J.S. 1992. Rat Gene mapping using PCR-analyzed microsatellites. Genetics 131:701-721.

Severson, D.W., Mori, A., Zang, Y. and Christensen, B.M. 1993. Linkage map for Aedes aegypti using restriction fragment length polymorphisms. J. Heredity 84:241-247.

Shiels, R. 1996. Plant chromosomes unite. TIG 12:37-38.

Slettan, A., Olsaker, I. and Lie, O. 1993. Isolation and characterization of variable (GT)n repetitive sequences from atlantic salmon, *Salmo salar* L. Anim. Genet. 24:195-197.

Slettan, A., Olsaker, I. and Lie, O. 1995. Atlantic salmon, *Salmo salar*, microsatellites at the *SSOSL25, SSOSL85, SSOSL311, SSOSL417* loci. Anim. Genet. 26:277-285.

Stahl, G. 1987. Genetic population structure of Atlantic salmon. In Population Genetics and Fisheries Management (N. Ryman and F. Utter eds.). University of Washington Press, Seattle, WA. Pp. 121-144.

Stallings, R.L. 1994. Distribution of trinucleotide microsatellites in different categories of mammalian genomic sequences: Implications for human genetic diseases. Genomics 21:116-121.

Stallings, R.L. 1995. Conservation and evolution of (CT)n/(GA)n microsatellite sequences at orthologous positions in diverse mammalian genomes. Genomics 25:107-113.

Stathakis, D.G., Pentz, E.S., Freeman, M.E., Kullman, J., Hankins, G.R., Pearlson, N.J. and Wright, T.R.F. 1995. The genetic and molecular organization of the *dopa decarboxylase* gene cluster of *Drosphila melanogaster*. Genetics 141:629-655.

Stephen, A.B. and McAndrew. 1990. Distribution of genetic variation in brown trout, *Salmo trutta* L., in Scotland. Aquacult. Fish. Manage. 21:47-66.

Taggart, J.B. and Ferguson, A. 1984. Allozyme variation in the brown trout (Salmo trutta L.): single locus and joint segregation inheritance studies. Heredity 53:339-359.

Taggart, J.B., Prodohl, P.A. and Ferguson, A. 1995. Genetic markers for Atlantic salmon (*Salmo salar* L.): single locus inheritance and joint segregation analyses of minisatellite (VNTR) DNA loci. Anim. Genet. 26:13-20.

Tanksley, S.D., Fulton, T.M. and McCouch, S.R. 1993. Linkage map of rice (*Oryza sativa*) (2N = 24). *In* Genetic Maps: Locus Maps of Complex Genomes. 6th edition. (S.J. O'Brien ed.). Cold Spring Harbor Press, Cold Spring Harbor, N.Y. pp. 6.61-6.79

Tanksley, S.D. 1993. Linkage map of tomato (*Lycopersion esculentum*) (2N = 24). *In* Genetic Maps: Locus Maps of Complex Genomes. 6th edition. (S.J. O'Brien ed.). Cold Spring Harbor Press, Cold Spring Harbor, N.Y. pp. 6.39-6.60.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nuc. Acids Res. 17;6463-6470.

Taylor, A.C., Sherwin, W.B. and Wayne, R.K. 1994. Genetic variation of microsatellite loci in a bottlenecked species: the northern hairy-nosed wombat *Lasiofhinus krefftii*. Molec. Ecol. 3:277-290.

Tessier, N., Bernatchez, L., Presa, P. and Angers, B. 1995. Gene diversity analysis of mitochondrial DNA, microsatellites and allozymes in landlocked Atlantic salmon. J. Fish Biol. 47 supp. A:156-163.

Thorgaard, G.H. 1978. Sex Chromosomes in the sockeye salmon: a Y-autosome fusion. Can. J. Genet. Cytol. 20:349-354.

Thorgaard, G.H. 1992. Application of genetic technologies to rainbow trout. Aquaculture 100:85-97.

Thorgaard, G.H. 1983. Chromosomal differences among rainbow trout populations. Copia 1983(3):650-652.

Thorgaard, G.H. 1977. Heteromorphic sex chromosomes in male rainbow trout. Science 196:900-902.

Tiersch, T.R., Simco, B.A., Davis, K.B. and Wachtel, S.S. 1992. Molecular genetics of sex determination in channel catfish: studies on SRY, ZFY, Bkm, and human telomeric repeats. Biol. Reprod. 47:185-192.

Trowler, M.K., Orton, S.M., Purvis, I.J., Sanseau, P., Riley, J., Christodoulou, C., Burt, D., See, C.G., Elgar, G., Sherrington, R., Rogaev, E.I., St.George-Hyslop, P., Brenner, S. and Dykes, C.W. 1996. Conservation of synteny between the genome of the pufferfish (*Fugu rubripes*) and the region of human chromosome 14 (14q24.3) associated with familiar Alzheimer disease (*AD3* locus). Proc. Natl. Acad. Sci. USA 93:1366-1369.

Vaiman, D., Pailhoux, E., Payen, E., Saidi-Mehtar, N. and Cotinot, C. 1995. Evolutionary conservation of a microsatellite in the Wilms Tumour (WT) gene: mapping in sheep and cattle. Cytogenet. Cell Genet. 70:112-115.

Venkatesh, B., Tay, B.H., Elgar, G. and Brenner, S. 1996. Isolation, characterization and evolution of nine pufferfish (*Fugu rubripes*) actin genes. J. Mol. Biol. 259:655-665.

Wachtel, S., Demas, S., Tiersch, T. Pechan, P. and Shapiro, D. 1991. Bkm satellite DNA and ZFY in the coral reef fish *Anthias squamipinnis*. Genome 34:612-617.

Waterson, R. and Sulston, J. 1995. The genome of *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA. 92:10836-10840.

Weber, J.L. 1990. Informativeness of human (dC-dA)n (dG-dT)n polymorphisms. Genomics 7:524-530.

Weber, J.L. and Wong, C. 1993. Mutation of human short tandem repeats. Hum. Mol. Genet. 2:1123-1128.

Weber, J.L. and May, P.E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. Gen. 44:388-396.

Weeden, N.F., Timmerman, G.M., Hemmat, M., Kneen, B.E. and Lodhi, M.A. 1992. Inheritance and reliability of RAPD markers, pp. 12-17 in Proceedings of the Symposium on Applications of RAPD Technology to Plant Breeding, Minneapolis.

Weeden, N.F., Muehlbauer, F.J. and Ladizinsky, G. 1992. Extensive conservation of linkage relationships between pea and lentil genetic maps. J. Heredity 83:123-129.

Wellbrock, C., Lammers, R., Ullrich, A. and Schartl, M. 1995. Association between the melanoma-inducing receptor tryrosine kinase *Xmrk* and *src* family tyrosine kinase in *Xiphophorous*. Oncogene 10:2135-2143.

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

Whisson, S.C., Drenth, A., Maclean, D.J. and Irwin, J.A.G. 1995. Phytophthora sojae avirulence genes, rapd, and RFLP markers used to construct a detailed genetic linkage map. Mol. Plant Microbe Interact. 8:988-995.

Whittaker, J.C., Curnow, R.N., Haley, C.S. and Thompson, R. 1995. Using marker-maps in marker-assisted selection. Genet. Res., Camb. 66:255-265.

Wild, V., Simianer, H., Gjoen, H.M., Gjerde, B. 1994. Genetic parameters and genotype x environment interaction for early sexual maturity in Atlantic salmon (*Salmo salar*). Aquaculture 128:51-65.

Williams, J.G.K., Kubelik, A.R., Rafalski, J.A. and Tingey, S.V.1991. Genetic analysis with RAPD markers,in More Gene Manipulations in Fungi. (J.W. Bennett and L.L. Lasure eds.). Academic Press, San Diego. pp. 431-439

Williams, J.G.K., Hanafey, M.K., Rafalski, J.A. and Tingey, S.V. 1993. Genetic analysis using random amplified polymorphic DNA markers. Methods Enzymol. 218:704-740.

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

Wilson, I.F., Bourke, E.A. and Cross, T.F. 1995. Genetic variation at traditional and novel allozyme loci, applied to interactions between wild and reared *Salmo salar* L. (Atlantic salmon). Heredity 75:578-588.

Wintero, A.K., Fredholm, M. and Thomsen, P.D. 1992. Variable (dG-dT)n/(dC-dA)n sequences in the porcine genome. Genomics 12:281-288.

Womack, J.E., Bolch, S.N. and Fries, R. 1993. Gene map of the cow (*Bos taurus*) 2N = 60. *In* Genetic Maps: Locus Maps of Complex Genomes. 6th edition. (S.J. O'Brien ed.). Cold Spring Harbor Press, Cold Spring Harbor, N.Y. pp. 4.264-4.275.

Wong, A.K.C., Yee, H.A., van de Sande, J.H. and Rattner, J.B. 1990. Distribution of CT-rich tracts is conserved in vertebrate chromosomes. Chromos. 99:344-351.

Wright, J.E., May, B., Stoneking, M. and Lee, G. 1980. Pseudolinkage of the duplicate loci for supernatant asparate aminotransferase in brook trout (Salvelinus fontinalis). J. Heredity 71:223-228.

Wright, J.E., Johnson, K., Hollister, A. and May B. 1983. Meiotic models to explain classical linkage, pseudolinkage, and chromosome pairing in tetraploid derivative salmonid genomes. Isozymes Curr. Top. Biol. Med. Res. 10:239-260.

Yates, J.R.W. 1996. Medical genetics. BMJ 312:1021-1025.

Zheng, W., Stacey, N.E., Coffin, J. and Strobeck, C. 1995. Isolation and characterization of microsatellite loci in the goldfish *Carassius auratus*. Mol. Ecol. 4:791-792.

APPENDIX A

Size (bp) of RAPD loci which were unique to either theAtlantic salmon or the brown trout parent of hybrid family JH91.

primer	sequence (5' to 3')	Atlantic salmon	brown trout
104	GGGCAATGAT	570	355
106	CGTCTGCCCG	620, 440	355
108	GTATTGCCCT	519	
115	TTCCGCGGGC	571	
119	ATTGGGCGAT	1698, 899, 687	1227
122	GTAGACGAGC	891	645, 531
125	GCGGTTGAGG	280	533
127	ATCTGGCAGC		362, 260, 237
131	GAAACAGCGT	204, 250	520, 380, 268
132	AGGGATCTCC	483	310, 378
133	GGAAACCTCT	612	532, 476
134	AACACACGAG	504	307
136	TACGTCTTGC		994, 931, 860, 666
146	ATGTGTTGCG	795	437
147	GTGCGTCCTC	320, 258	511
148	TGTCCACCAG	459	
150	GAAGGCTCTG	426	
151	GCTGTAGTGT	843, 690, 477	519, 424
152	CGCACCGCAC	906, 373	1057, 825, 403
156	GCCTGGTTGC	452	333
157	CGTGGGCAGG	1535, 224	1012, 375
164	CCAAGATGCT	803	603
169	ACGACGTAGG		547

173	CAGGCGGCGT	340, 217	678, 520
176	CAAGGGAGGT	748, 297, 258	
181	ATGACGACGG	387	
188	GCTGGACATC	625, 552	
192	GCAAGTCACT	380	823, 429, 330
200	TCGGGATATG	459	542
204	TTCGGGCCGT	672	835, 570, 473
210	GCACCGAGAG		655
215	TCACACGTGC	509, 429	614, 310
218	CTCAGCCCAG	306	559
219	GTGACCTCAG	1141, 625	432, 310
225	CGACTCACAG	732	
226	GGGCCTCTAT	835	316
228	GCTGGGCCGA	279	576, 422
229	CCACCCAGAG	298	
230	CGTCGCCCAT	373	540
231	AGGGAGTTCC		603
234	TCCACGGACG	778	1491, 1310
237	CGACCAGAGC	439, 406	288
238	CTGTCCAGCA		371
239	CTGAAGCGGA		460
241	GCCCGACGCG	889	810, 661
243	GGGTGAACCG		823
244	CAGCCAACCG	594	843, 645
245	CGCGTGCCAG	1134, 949, 645	
250	CGACAGTCCC	460	323
253	CCGTGCAGTA	364	1033
254	CGCCCCCATT	1278	872, 603
256	TGCAGTCGAA		1786
257	CGTCACCGTT	893, 706	744, 375

260	TCTCAGCTAC	545	
262	CGCCCCCAGT	324	630, 505
264	TCCACCGAGC	1388, 244	647, 560
266	CCACTCACCG		732
268	AGGCCGCTTA		165
269	CCAGTTCGCC	373	487
270	TGCGCGCGGG		324
271	GCCATCAAGA	466, 402	614
273	AATGTCGCCA		530
275	CCGGGCAAGC	740	
278	GGTTCCAGCT	1134	
280	CTGGGAGTGG	310	281
281	GAGAGTGGAA	658	827, 566, 499, 287
283	CGGCCACCGT	310	700, 363, 281
284	CAGGCGCACA		276
285	GGGCGCCTAG	747	352, 310
287	CGAACGGCGG	485, 242	
288	CCTCCTTGAC	1174, 530	872, 781, 340
290	CCGCGAGCAC	970, 641	837, 581, 341, 289
292	AAACAGCCCG		356, 285
293	TCGTGTTGCT	1353	
302	CGGCCCACGT	630	702
303	GCGGGAGACC		630, 505
305	GCTGGTACCC		541
308	AGCGGCTAGG		393
312	ACGGCGTCAC	593	389, 363
313	ACGGCAGTGG	652	
317	CTAGGGGCTG		1123
318	CGGAGAGCGA	425, 283, 230	440, 274
319	GTGGCCGCGC	664,459	718, 575

320	CCGGCATAGA	584	
322	GCCGCTACTA	1175	577,423
324	ACAGGGAACG		1268
328	ATGGCCTTAC	552	
329	GCGAACCTCC	701	789
331	GCCTAGTCAC		766, 592
336	GCCACGGAGA	594, 550	1448, 779, 443, 172
338	CTGTGGCGGT	215, 181	547, 401
340	GAGAGGCACC	289	497, 301
342	GAGATCCCTC		1130, 1059, 494
345	GCGTGACCCG	831, 708, 482	743
348	CACGGCTGCG	352	739, 526, 483, 399
350	TGACGCGCTC	219	290
351	CTCCCGGTGG		706, 350
352	CACAACGGGT	295	
353	TGGGCTCGCT	919, 306	833, 559
354	CTAGAGGCCG	293	517, 374, 273
355	GTATGGGGCT		1041
356	GCGGCCCTCT	584	395
358	GGTCAGGCCC	639	371
359	AGGCAGACCT	593, 510	701, 530, 459
364	GGCTCTCGCG	1147	1381, 573, 279
365	TAGACAGAGG	567, 271	603
367	ACCTTTGGCT		633, 557, 397
368	ACTTGTGCGG	220	561
373	CTGAGGAGTG	675	570, 436
374	GGTCAACCCT		891
375	CCGGACACGA	1254	574, 443, 402
378	GACAACAGGA		556
381	ATGAGTCCTG	345	670, 490, 381

387	CGCTGTCGCC	914, 504	1188, 297
388	CGGTCGCGTC	872, 409	639
389	CGCCCGCAGT	377	496
392	CCTGGTGGTT	740	979, 755
399	TTGCTGGGCG		575

APPENDIX B

List of UBC primers for which no amplification product was observed in Atlantic salmon, brown trout or their hybrid offspring.

primer	sequence (5' to 3')	primer	sequence (5' to 3')
107	CTGTCCCTTT	113	ATCCCAAGAG
118	CCCGTTTTGT	121	ATACAGGGAG
123	GTCTTTCAGG	124	ACTCGAAGTC
126	CTTTCGTGCT	128	GCATATTCCG
129	GCGGTATAGT	130	GGTTATCCTC
138	GCTTCCCCTT	143	TCGCAGAACG
145	TGTCGGTTGC	160	CGATTCAGAG
161	CGTTATCTCG	162	AACTTACCGC
163	CCCCCCAGAT	166	ACTGCTACAG
167	CCAATTCACG	168	CTAGATGTGC
170	ATCTCTCCTG	172	ACCGTCGTAG
174	AACGGGCAGC	178	CCGTCATTGG
183	CGTGATTGCT	185	GTGTCTTCA
191	CGATGGCTTT	200	TCGGGATATG
201	CTGGGGATTT	202	GAGCACTTAC
205	CGGTTGGAA	206	GAGGACGTCC
207	CATATCAGGG	209	TGCACTGGAG
211	GAAGCGCGAT	212	GCTGCGTGAC
214	CATGTGCTTG	216	CATAGACTCC
217	ACAGGTAGAC	220	GTCGATGTCG
221	CCCGTCAATA	223	GATCCATTGC
224	TCTCCGGTAT	232	CGGTGACATC
233	CTATGCGCGC	236	ATCGTACGTG
240	ATGTTCCAGG	242	CACTCTTTGC

246	TATGGTCCGG	247	TACCGACGGA
249	GCATCTACCG	251	CTTGACGGGG
252	CTGGTGATGT	255	TTCCTCCGGA
258	CAGGATACCA	259	GGTACGTACT
261	CTGGCGTGAC	263	TTAGAGACGG
267	CCATCTTGTG	272	AGCGGGCCAA
274	GTTCCCGAGT	279	AGACATTAGA
282	GGGAAAGCAG	286	CGGAGCCGGC
288	CCTCCTTGAC	291	AGCTGAAGAG
294	TGATTGGCCA	298	CCGTACGGAC
299	TGTCAGCGGT	301	CGGTGGCGAA
306	GTCCTCGTAG	307	CGCATTTGCA
310	GAGCCAGAAG	311	GGTAACCGTA
314	ACTTCCTCCA	315	GGTCTCCTAG
316	CCTCACCTGT	321	ATCTAGGGAC
325	TCTAAGCTCG	326	CGGATCTCTA
327	ATACGGCGTC	330	GGTGGTTTCC
332	AACGCGTAGA	333	GAATGCGACG
334	ATGGCAAAGC	339	CTCACTTGGG
341	CTGGGGCCGT	343	TGTTAGGCTC
344	TGTTAGGCAC	346	TAGGCGAACG
347	TTGCTTGGCG	349	GGAGCCCCCT
357	AGGCCAAATG	362	CCGCCTTACA
363	ATGACGTTGA	366	CCTGATTGCC
369	GCGCATAGCA	370	TCAGCCAGCG
371	TCTCGATTGC	372	CCCACTGACG
377	GACGGAAGAG	378	GACAACAGGA
379	GGGCTAGGGT	380	AGGAGTGAGA
384	TGCGCCGCTA	385	ACCGGGAACG
386	TGTAAGCTCG	390	TCACTCAGAG

391	GCGAACCTCG	394	TCACGCAGTT
395	TCACTTGAGG	396	GAATGCGAGG
398	CAGTGCTCTT	400	GCCCTGATAT

APPENDIX C

Individual genotypes of hybrid progeny in families HFJ91 and HFA192 at marker loci inherited from either a brown trout or Atlantic salmon parent. For all RAPD markers except UBC234:778:Ssa, a, represents the null allele and A, the observed PCR product For all microsatellite markers, and RAPD marker UBC234:778:Ssa, a and A represent alleles of lower and higher molecular weight respectively. Missing data are represented by a dash (-). The sex and tag number of each parent is given in brackets.

family: HFJ91
parent: Atlantic salmon (♀/1634)

	hybrid progeny																	
Marker	2	4	5	7	10	11	6	8	12	14	20	25	32	36	51	54	3	9
UBC234:778:Ssa	a	a	A	A	a	A	A	A	a	A	a	a	A	A	A	a	a	A
UBC348:352:Ssa	A	A	A	a	a	A	A	a	a	-	A	a	a	a	a	A	a	A
UBC368:220:Ssa	a	A	a	a	A	a	A	A	a	a	A	a	a	A	A	a	a	A
Ssa4	a	a	a	A	a	A	A	a	a	A	A	a	A	A	A	a	a	A
SSOSL85	A	a	A	A	a	A	A	a	A	-	a	a	a	a	a	a	-	A
SSOSL417	a	a	a	A	a	a	a	a	A	-	a	a	a	a	A	a	A	a
SSOSL311	A	a	-	A	a	a	a	a	A	a	a	a	a	a	a	-	A	a
μ60	A	a	A	-	a	a	A	A	a	a	a	a	A	a	A	a	A	A
μF43	A	a	A	A	a	a	A	A	a	a	a	a	A	a	A	a	A	a
ApoAI	a	A	a	A	a	a	-	A	A	a	A	A	A	A	A	A	A	A
fib-2	a	a	A	A	a	a	a	a	a	A	A	a	-	-	-	-	a	A

family: HFJ91 continued
parent: Atlantic salmon (♀/1634)

hybrid progeny

Marker	13	19	22	31	38	39	41	52	15	16	18	23	24	26	27	28	34	35
UBC234:778:Ssa	A	A	A	a	A	a	a	A	A	A	A	A	A	A	a	a	A	-
UBC348:352:Ssa	A	A	A	A	A	A	A	A	A	A	A	a	a	a	a	a	A	a
UBC368:220:Ssa	a	a	a	A	a	A	A	A	a	A	-	A	-	-	-	A	-	-
Ssa4	a	a	a	A	a	A	A	A	A	A	a	a	a	a	A	A	A	A
SSOSL85	a	a	A	A	-	a	a	a	a	a	A	A	A	a	a	a	a	-
SSOSL417	a	a	A	A	a	A	A	A	-	A	a	a	a	A	A	A	A	-
SSOSL311	a	a	A	A	a	a	A	a	a	A	a	a	a	A	A	A	-	-
μ60	a	A	a	A	a	A	A	a	A	A	A	a	A	A	a	A	A	A
μF43	a	A	a	A	a	A	A	a	A	A	A	a	A	A	a	A	A	-
ApoAI	-	A	a	a	a	A	A	A	A	A	a	a	a	a	A	A	A	A
fib-2	A	A	a	-	-	-	-	-	a	A	A	A	a	A	A	-	-	-

family: HFJ91
parent: brown trout (♂/1619)

	hybrid progeny																	
Marker	2	4	5	7	10	11	6	8	12	14	20	25	32	36	51	54	3	9
UBC131:268:Str	A	A	a	a	a	-	a	a	A	-	A	A	-	-	A	a	a	a
UBC152:825:Str	a	A	a	A	a	A	A	a	a	-	a	A	-	A	a	a	a	a
UBC273:530:Str	A	a	A	A	A	a	A	A	A	a	a	a	A	A	a	a	a	A
UBC226:316:Str	a	A	a	A	A	A	A	A	a	a	A	A	a	a	A	a	A	A
UBC239:446:Str	a	A	a	A	A	A	A	A	a	a	A	A	a	a	A	a	a	A
UBC280:281:Str	A	a	A	a	a	a	a	a	A	a	A	A	A	a	a	A	a	A
UBC219:432:Str	A	-	a	A	a	a	a	A	A	-	A	-	a	a	A	a	A	A
UBC243:823:Str	A	A	a	A	A	a	a	A	A	a	A	a	a	a	A	a	a	A
UBC348:739:Str	A	A	a	A	A	a	a	A	A	-	a	a	a	A	a	a	a	A
UBC348:399:Str	A	a	a	A	A	a	a	a	A	-	A	A	A	A	A	a	A	A
UBC342:494:Str	a	a	a	A	A	a	a	a	A	a	-	a	a	-	a	a	A	A
UBC354:517:Str	a	A	a	a	a	A	a	a	a	a	A	A	a	-	A	A	a	a
UBC358:371:Str	a	A	a	A	A	A	a	A	a	-	A	A	-	a	A	-	a	A
SSOSL311	a	A	-	a	A	a	A	a	a	A	a	A	A	A	a	-	a	a
SSOSL417	a	A	a	a	A	a	A	a	a	-	a	A	A	A	a	A	a	a
μ15	A	A	a	A	a	a	a	A	A	a	A	a	a	a	A	a	A	A
μ73	a	A	A	A	A	A	A	A	-	A	A	a	a	-	A	a	A	a
μF43	A	A	a	A	A	a	a	A	A	A	a	a	a	A	a	a	a	a

family: HFJ91 continued

parent: brown trout (♂/1619)

hybrid progeny

Marker	13	19	22	31	38	39	41	52	15	16	18	23	24	26	27	28	34	35
UBC131:268:Str	-	A	a	-	A	A	a	A	a	a	A	A	A	a	a	A	a	a
UBC152:825:Str	A	-	-	A	a	a	a	A	a	a	A	a	A	A	A	a	A	A
UBC273:530:Str	a	a	a	a	A	A	a	a	a	a	A	a	a	A	A	a	A	A
UBC226:316:Str	a	a	a	a	a	A	a	a	a	A	a	a	a	A	a	a	A	a
UBC239:446:Str	a	a	a	a	A	A	a	a	a	A	a	a	a	A	a	a	A	a
UBC280:281:Str	A	A	A	A	A	A	a	a	A	a	a	a	A	A	A	a	A	A
UBC219:432:Str	A	A	a	a	A	A	A	a	A	a	A	A	A	A	A	a	A	A
UBC243:823:Str	A	A	a	a	A	A	a	a	a	A	A	a	a	A	A	a	a	.a
UBC348:739:Str	A	A	a	a	A	A	a	a	a	A	A	a	a	A	A	a	a	a
UBC348:399:Str	A	A	A	A	A	a	a	a	A	a	a	a	A	A	a	a	A	A
UBC342:494:Str	a	A	A	A	A	A	a	a	a	A	A	A	a	A	A	a	A	A
UBC354:517:Str	A	A	A	A	a	A	A	A	A	A	a	A	A	a	a	A	a	a
UBC358:371:Str	a	a	a	a	A	A	a	a	a	A	a	a	a	A	a	a	A	a
SSOSL311	a	A	A	a	a	A	A	A	A	A	A	A	a	A	A	a	-	-
SSOSL417	a	A	A	a	a	A	A	a	-	A	A	A	a	A	A	a	A	-
μ15	A	A	a	a	A	A	A	a	A	a	A	A	A	A	A	a	A	A
μ73	a	a	a	a	a	A	A	A	A	a	A	a	A	A	a	-	A	A
μF43	A	A	a	a	A	A	a	a	a	A	A	a	a	A	A	a	a	-

family: HFA192
parent: Atlantic salmon (♀/5211)

hybrid progeny

Marker	1	2	3	12	20	21	23	24	25	26	27	28	29	30	31	32	33
SSOSL85	A	A	A	A	a	a	a	A	a	A	a	a	A	a	A	a	-
μ60	A	a	a	a	A	A	a	A	A	a	A	A	a	A	A	A	a
μF43	a	A	A	A	a	a	A	a	a	A	a	a	A	a	a	a	A
ApoAI	A	A	A	-	a	a	a	A	A	A	a	-	A	a	A	a	a

family: HFA192 continued
parent: Atlantic salmon (♀/5211)

hybrid progeny

Marker	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
SSOSL85	a	A	A	a	A	a	A	a	A	A	a	a	a	a	A	a	a
μ60	a	a	a	A	a	A	A	A	A	A	a	a	a	a	a	A	a
μF43	A	A	A	a	A	a	a	a	a	a	A	A	A	A	A	a	A
ApoAI	a	A	-	a	A	a	-	a	A	A	a	a	A	a	A	a	A

family: HFA192
parent: brown trout (♂/1105)

hybrid progeny

Marker	1	2	3	12	20	21	23	24	25	26	27	28	29	30	31	32	33
SSOSL85	A	a	a	a	A	A	a	a	A	a	a	a	A	A	A	a	-
μ60	a	a	a	A	A	a	a	a	A	a	a	a	a	a	a	a	A
μF43	a	a	A	A	A	a	a	A	A	A	A	A	A	a	a	a	A
ApoAI	a	A	A	-	a	a	A	A	A	A	A	-	a	a	a	-	A

family: HFA192 continued
parent: brown trout (♂/1105)

hybrid progeny

Marker	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
SSOSL85	A	A	A	a	a	a	a	a	A	a	a	A	A	A	A	a	a
μ60	A	a	a	A	A	A	a	A	a	a	A	A	a	a	a	a	a
μF43	a	a	A	A	a	A	A	a	a	A	a	a	a	A	A	a	a
ApoAI	a	a	-	A	A	A	-	A	a	A	A	a	a	a	a	A	A

APPENDIX D

Nucleotide sequence of microsatellite loci isolated from an Atlantic salmon partial genomic library according to the protocol of Taylor et al. 1994. DNA fragments were cloned into pTZ18R and sequenced from either end using the M13 reverse (5'-d[CAGGAAACAGCTATGAC]-3') and universal (5'-d[GTAAAACGACGGCCAGT]-3') primers.

clone	frag. size (bp)	locus name	access. number	sequencing primer	sequence (5' to 3')
A97	417	SSMUN1	U79972	universal	AACTGAGGAGATTAAATCAGAACATCCTTGAAAATAATAA TTTTTTCACACCCTCACCATAGAGAGCCTTTTTATTTCTC TATTTGGTTAGGTCAGGGTGTGATTTGGGTGGGCATTCTA GTTTATGTATTTCTATGTTGGCTCTGATTTGAAAACTATG AATTCTCTGGGCT(GT)$_{64}$TGCATGGATGATCTCACTCAGA GAATCCAGGTGA
B23	462	SSMUN2	U79973	universal	CCCATAATGTCCACCAGACCAGCATAGTCCCCGTCTGAGA CCTTATTGGACAGACCACAAGATGTGTTCTGGACAAACAA CGCCAGGTCGCTCACACTGGGGGAGAGAGATGGAGAGAGA ACGTTGAAGAGGGGGTATGCAGTGTATTGGATGATGAGGG AGACACAGTAAATAGATAGGTGTTTGTGTGTCTGTTTAGA TGTA(GT)$_5$TTAATTGAGTGCACATCACATT(TG)$_6$TAGTA TATGCATTGATC(TG)$_5$TTTAATTGAGTGCACATCACATT T(GT)$_{43}$CACCTCTCATTGACATGGTTCAGTAGGTGCTCTT TGAACATC

B34	389	SSMUN3	U79974	universal	TGCAGGGGTCA(GT)$_{48}$ATCTTCAGTGCAGGGGTCA(GT)$_{8}$ ACTAATTTCAATGCAAACATTTAAACAAAATGCGCCTTGG TCATAAACAATGCATACAATACTTCAACTTAGAGCAGATT TATTACAG
B92	677	SSMUN4	U79975	reverse	GTGC(GT)$_{19}$GC(GT)$_{8}$TTGT(G)$_{25}$ACACCAGCAGGATTAA ACGATAGGCTAGGTCTGTGTTT
				universal	GGTATGGAGGTCCCTATACCCTCTATAAGTATTTAAAACC GTTTTTCAAAAATGGTTGTGGTAACCCAAAATAAAACCAG GTGCCCGTTTGGACTTAGTGCAATTTCTGAGAAGAAAAAC ATAGAATATTTTCTGAAGATTTTTTCAAACTTCCAT
C5	677	SSMUN5	U79976	reverse	TCAGCGCCACGGCATGCCGAGCACAGCCAA(CA)$_{51}$CTCTA TAAGACAGGTGGGAGGATTCAAATGTCTCACACTGTAGTA GAGCTACAATAATAACTGTGGTGAACCATGAGGGAGTGTG AAATGCA
C55	309	SSMUN6	U79977	universal	GTGTGTA(GT)$_{43}$GAGACCCTAAAGACACAGAGCAGCCACT CTGTAATGTAGCCCAGGTGGAGGCGCTAGACAAGCGTTCA CACAGTGAACCTGTGTGTGCTCAAA(CA)$_{6}$GT(CA)$_{6}$TG(C A)$_{5}$CTCTCTACTCTCTCT

D29	709	SSMUN7	U79978	universal	AAACAACTGAAATGTTTACCACTTTTCCAACAGATTAGAC AAAACATTAGAGGGTATGACATCTTGGGCTACATCTCTGC TATTTCAATCTGGAATCTTCATTGGTAAAGGTAAAATAAA TAAAAATCTCCTCTTTCACCATTTAAGCCTCCAGCGTCAA GGTCTATGGCTGTTAGAGAGTGGTTGGAAAAGGGNNNNNN ACTTTACCTACC(TG)$_{35}$
				reverse	AATAAGAAAATATTTCTCTAGGGCTAATATAAATAATTCT AAAATGCCTAGTGAACTTCAATTAATGTATAGCAATAGGT CTGAAAAAAAAACCTGTATGTTATTATTTTTAATTCCTGG GCGATTTAACTTTTTGGAGATACAGGGTTT
D72		SSMUN8	U79979	universal	TGACTCTAAGAGGAAACAGAGA(TG)$_{45}$TCAGTCTGGGTCC TAGCCCCATTGTTTATGAACTTAGTGGCTTATGCATATCA GCATGCAAACAAACCCAAA(CA)$_{10}$AACCAAA(CA)$_{8}$
E7	579	SSMUN9	U79980	reverse	ATGCGGTGTTTGATTTGTCAGGCATAACGGGACCCATGTG GTCAAAAAGTTATACTTTTGACTCATCTGTCCATAGAACA TTCTTCCAAGAGTCTTGATGATTCATCCAGGTGCTTCCAG ATGCTTTTTGGCAAACCTTGAGTCAACTTTTTGGATGAGA TGGGTCCCATATATA(AC)$_{26}$AGGTTTGCTAGCCACATAGA ACCTCTTAAAGACTTT
				universal	TAGTTTTGAGATTTCCCCCTCCCCACTCAGACCACTCCCA GACAGTCCTAGAAAAATTCTTACTCTTTGCTAAGATGCTA CTTTTTAAAAAGACAATTATATTGAGAACTGCTGCATTGG TCTTTTAACCTCAATTAATTTGGACACATAAATCAGATTT CATGTACAAAACACAATCAAGCACTCACACAAAAATGTCT GCATTCAAGGATAAGGTGG

E55	398	SSMUN10	U79981	universal	GCATTATGTGGCTGAGTGATGAGGAGACAAACCACCAGAT ATACCCTGGCATTATGTCGCTGAGGAGACAGACCACCAGA TATACCCTGGCATTATGTAGACAACATCC $(GT)_{23}$
				reverse	AGTGCGTATGTGTATGCATGTGTTTGTGTTACTTCACAGT CCCGCGTTCATAATGTGTTTNAATGTG
E72	380	SSMUN11	U79982	universal	TTTCTGATGACCCATCACTTTGCTGCAGAAGACACTCAGA AGTGTGTC $(GT)_{50}$TTGCACACGNCTGTTTTCAGTCCTCTC CAGAGATGTTCGGTCAGGTTGAAGTACGGGCTCTGGCTGG GACAATCAAGGACATTCAGAGACTT
E83	437	SSMUN12	U79984	universal	ACAACGTGGCTGAGACGATGGCACTGACAAAGGCATGTAA CAAAACACGACTATAAAATGAACAAACTACACAATTCTAC CTAAGGAGTCGACATTTTTTTTTGGTTAAACATTTTTTGT CTAAAGACCACCATTAGCTGCTGTTGAGTTTTTGCAGCAG GTTTGTGACAATACATGAGACATGAT $(TG)_{6}$A $(TG)_{14}$CCT GTGAGAG $(GT)_{20}$
F5	591	SSMUN13	U79983	reverse	GTTAATACAACAAGTCTCTCCTTGCTTGTGATTGGTGGAG ATGAAGACCTACCTCTGGAGTTGATGTCTTCATCCTCCGT GTCTGTAAAGATGAACGTCTGGAGAAGGAGAAGAGAGATA TAACACACACATACACACACACATA $(CA)_{31}$GACACCTCCT $(CA)_{21}$TACACACACATATA $(CA)_{6}$A
		SSMUN14	U79985	universal	TGCAGCCTGTTTCTGTCAGTATTAATCTGTCCTCCATTAT CTCGCTGTGTGTGTGGGGGGGGGGCTGTCT $(GT)_{22}$

F11	820	SSMUN15	U79986	reverse	TTGTATTACCAATATGAGTGGCTGAACTGAAAAAGGCTAA ATATGATTTTGAAGGTGGGCTACCATTGAGTGTGTGTGCC AGAGAGAAAGACGAAGTGTGTGTTTGCATGTAATGGATGT TCATTTTTTTTTTCTTATCATCCCCCTTCT(AC)$_{32}$CACCA CCACTGTTCTAAA
				universal	GACATTTTTTTATGAGGATCCCAGAAAAAAAAGTATGTTT TTTGTCATTTTATAAACATGGTTCATGCAATTCTACACCA CTTTACTGTCCCTGCCTGGAAACCATAGCAGAACAATTAA TACCACAAAAATGACAAGGTATTCTACTCTGCTGACACTG AAAAACTGAGAATGACATTGTCTTGAATGCATCCATCTAA ATCAAGGACAATGAGAAGGGGAGACACATA
F49	755	SSMUN16	U79987	universal	GTGTGTGTGCATGTGTGTGAGAGAGACAGAGAGAGCCTAC TTAAAACAGTGTTTGAACCTTGCTCTGTGTGTGTGTGCAT GTGTGTGAGAGAGAGAGAGCCTCATTAAAACAGTGTTTCA ACCTTGCTC(TG)$_{11}$TTTGAGATGAAGGGTCAGAGCAGTGA TTCATATGATGAGTAACACTGAATTAGCCC
F55	380	SSMUN17	U79988	universal	AGAGCAACCACTCTGATGCCCAGGCAGGCAGGCTGATCTC CCTGTTTGATCTTGTCGATAGCACCGTCATGACCCTGCTT CTAATTGAATCAATCAGAGAGAGAGAGAAAGTGTGTATAA ACAAATCATCTACCAGTTGATGTGTGTTCGCCTGTGTTGA TATTTACATGTTGTTTCAGCAGTGTGTAACATATGTTA(G T)$_{20}$

G14	803	SSMUN18		reverse	TGTTCTTTTGGTGGTGGACCATTCTTGATCACACGAGAAA CTGTTGAGCATGAAAAAAACAGCACTGATGCAGTTCTTGA CACAAACCGGTATACCTGNCTCCTACTACCATAGCCTGTT CAAAGGCTCTTAAATCTTTT
			U79989	universal	TGGAGAGGTCANCCACTCACTGGGCAGAGGCATGCCAGTG GGAACATGGACCTAATNACCTGCTCATGTCCATCGGTCAG AACCTCGCTGTTCACTGGGGCAGGTAAGACAATATG(AC)$_{5}$GCATGCATACACACATGTACACTTCT(CA)$_{58}$CTTTGTGC TGTCTTCTACTGAGTCCATGTGAGTCTGTAGTGTGGCTGT (GT)$_{4}$GATGTACACTGAGTGAACAAAACATTATG
G17	787	SSMUN19	U79990	universal	GGAGAGGTCAGCCACTCACTGGGCAGAGGCATGCCAGTGG GAACATGGACCTAATNACCTGCTCATGTCCATCGGTCAGA ACCTCGCTGTTCACTGGGGCAGGTAAGACAATATG(AC)$_{5}$ GCATGCATACACACTGT(AC)$_{50}$TTTGTGCTGTCTTCTACT GAGTCCATGTGAGATCTGTAGTGTGGCTGTTGTGTGTGGA TGTACACTGAGTGAACAAAACATTATGATGAACACCTNCT CTTTCATGACATAGCTGACCA
G25	696	SSMUN20		reverse	TCCTCTCCTTCTCCTCCTTCCTTTCTTCCTCCTTCCTCCT CTCTTCTCTCCTTCTCTCTNNNCTTCTCCTTCNCGCCTTC CTTCATCCTCCTTCCTGCTCCAGACCCCACACCCTGTACA CTGC
			U79991	universal	CACTGAGGAGGCTTTTGTGTCATTGTATCAAGTTACCTAC AA(TG)$_{5}$C(GT)$_{6}$GCTC(GT)$_{5}$GTCTC(GT)$_{33}$GCATGAGTG CCTGCCTGTGAGCTTTTCTCTGTGTTCTGTCTACCTGAGT TCATGGTCCAGTGATGTTCCTGACTGAGCCAGGCTGGACT CCACAATACATGGCCCCTCCA(GGACTGGAGCATGTGACA GGGTCTGG)$_{6}$

G36	677	SSMUN21	U79992	universal	AACAATATTTAATAGAATAGGCTCCCTCAATAAACAATGA TCTCTTATTTCGCTAG(GT)$_{57}$CAGTGTTTCATCAGGGAAG TTTT
G49	603	SSMUN22	U79993	universal	CAGGGGGTACAAAGCACAACTGAGTGTTTTCCAAACCCTA CCTGTGCTTGCTAACTCCCCAAGGCTACGGCTAC(TG)$_{5}$A ACACAAGCCATGCTTACTTGCCTAATCATGTATGATTACG TTATTTCTTTTCCTCCTTCCTCCTGA(TG)$_{85}$AAAAAAGAG GGACAGTCACCTTTCTGTGACA
				reverse	AAAAGGGGGTTTCAGTAAAGTTTGGGTAGTATATGTGTAA CTCAGGGTAGGATTGGTCCCTATNTNTCCCCTCTATCCAT ATGCAGTTTTAGTCTGACCAAACAGTCATGGCATTATCCC CTCCCACACAAAGGACATGT
I6	567	SSMUN23	U79994	reverse	TGAACAACTGATGTTTTGTAAATAAGTGACAAAGAATGTT GGAAGTGTTGACAAAATAAAAATTGAAAAGATTACAACAG ATTTTTAAAGAAGTTTTCTGTGTGATAACACTACAGTGGT GTACATACAGAAATATTAAACCCTCTTGGTCTTT(CT)$_{9}$(CA)$_{72}$AATCAT(CT)$_{6}$CACACACACCCTGGTCCCCCTCCCC CGACACACAATCCCTGCTCCTCTCCACCCCAACCAATACA CCATACTCCTTTCCCTCAGTGTCCTGTGTGAAGCAGACGT AGAGGCACTGCTGGGTGGGTGAGAGGATAGGGGTAGAGTG GTGAGACATATATCTCACACACACATACATCC

I50	787	SSMUN24	U79995	universal	TTCCTACTGAGCAAGAAAACAAGACAATANAAACATATAA GTGACATGTGTATCTGTCAGCAATATCAAACTGCTGCCTG TCCATTGATATCTTTCCATCTGTTTAACCTGCTTATGGGG AATA$(GT)_{57}(AG)_{6}$AAGGTAG
				reverse	AAACGCATTCTTAGCATCCCCGACCTTCCCCCTAACAATC GGATTGACAAGTTCCGAGCCGAGGCACGATAATTACCTAC GTCATGTAATGCATTTCATACACGATGTAAATCATGGAAA GAATGTGAGAAATTCAAGC
I52	501	SSMUN25	U79996	reverse	AACAACCCACTATCATAACAGGTGGCATATTTGCTAAGTG TTTCATCTTTGTTTAAAAGGGGAACTAGAAGAGTTGGTGC $(GT)_{20}$
				universal	AAATCGATCACTTGGCTAATTGACTTCAATACACAGGAAG AAAAAGTGTGAAACCAGCCTTTGAAGTTGCTAGGCTTTTT CCTGTGTAAAATAATAGTTAGGCTATAGTTTAC

APPENDIX E

Partial sequences of cDNA clones for which putative identifications have been made through comparison with GenBank/EMBL data base entries. Some of the sequences represent more than one clone and have been combined.
(‡Drover 1996; †Heeley et al. 1995)

Clone	access. number	putative i.d.	sequence (5' to 3')
ASM8.27	U79460	transcription factor	TCGGAAGAGGTGAACATGTTTACGAACCAGGGGACAGTGATCCATTT CAACAACCCCAAAGTGCAGGCCTCCCTGGCAGCCAACACCTTCACAA TCACTGGCCACGCGAGACCAAGCAGCTGACAGAGATGCTGCCAGCAT CCTCAACCAGCTGGGAGCAGACAGCTGACAGCTGGGAGACT
ASM8.33	U79963	troponin I	CGATCCAACCTGAAACAAGTGAAGAAAGATGAGAAGAAAGAGGATAA GGAATTGCGTGACGTTGGTGACTGGCGTAAGAACATTGATGACAAGG CCGGTATGGACGGCAGGAAGAAGATGTTCAGGGAGATAATTAGATCA TTGGGCTGTGTTGTATGATGTGACTGACTACTGTCTGTATGTCTTAC TGTA
ASM1.3	SSU18402	aldolase	ACGCGTCCGCATGAACCAGTGTCCCCTGCACAGGCCATGGGCATTGA CGTTTTCCTATGGCCGTGCCTCCAGGCATCCGCCCTGAAGGCATGGG GTGGCAAACCTGGGAATGGCAAGGCTGCCCAGGAGGAGTTCATCAAG AGAGCCTTGGCCAAACGCCTTGGCCTGCCAAGGCAGTATGTGCGGCT TCCGGAGAC

ASM1.21	U79964	myosin light chain	ACCATCCAAACCTCCATACATACCGTCTCGAGATGGCACCCAAGAAG GCCAAGAGAGGGGGGGAGCAGCAGCAGAGGCCGGTTCCTCCAACGTG TTCTCCATGTTTGAGCAAGCCAGATCCAGATCCAGGAGTACAAGGAG GCTTTCACAATTATTGACCAGAAT
ASM8.35	U79965	creatine kinase	TCCGGAGAATGTCTGTCGAGGCCCTGGACACCCTGGACGGTGAGTTC AAGGGAAAGATCATCCCCCTGAATAAGATGACCGATGCCGAGCAGGA GCAGCTGACGTCTGACCACTTCTTGTTTTGATAAGCCCGCTCTCCCT TGCTGCTGGGCGCTGGATTGGCCGTGACTGGCCAGTGCAAGAGGAAC TTGCACAACAGTGCAAAAAGCTTCTTGCTGTTGT
ASM11.17	U79966	creatine kinase	TGGCTGGTGATGAGGAGTCCTACGAGATCTTCAAGGATCTGTTGGAC CCCATCATCTCAGACCGTCATAGTGGATACAAGCCCACAGACAAGCA CAAGACCGACCTGAACTTTGAGAACCTGAAGGGAGGTGATGACCTGG A
ASM1.22	U79967	creatine kinase	TTTTTTTCTTGTCTGGTGTGCAGTGTTAGAGGCAACTATGCCTTTCG GTAACACCCACAACAACTTCAAACAACTTCAAACTCAACTTCAAAGT TGAGGAGGAGATCCCTGACCTCACCAAGCACAACAACCACATGGCCA AGGTGCTGACCAAGGACATGTACGC
ASM11.30	U79968	parvalbumin	CTGAGATCAGCAGCCTCCTCCGCCACTAATCCACTAATCCTCCAAGG ACATCGCTGCTGCCCTCCAGCATTGCGCAGCTGCTGACTCCTTCAAC CACAAGGAGTTCTTCGCCAAGGTTGGCCTGGCTGGCAAGTCTACACA GGATTGAAGAAAGCCTTCTACTTGGTTGACCATTACAAGAGTGGCTT CATTGAGGAGGATGAGCTCAAG

ASM11.35	U79969	parvalbumin	CAACCGTTCATCAAAGACGTCAGAGCTCAAAATGCCTGTGCCCATCT GTGCAAGGAGGCTGACATCAAAACCGCATTGGAAGCGTGCAAAGCCG CCGACACCTTCAACTTCAAGACCTTCTTCCACACGATTGGCTTTGCC TCGAAGTCCGCCGACGACGTCAAGAAGGCTTTCAAGGTCATTGACCA GGATGCAAGTGGCTTTATTGAGGTGAGGAGCTGAAGCTGTTCC
ASM11.24	SSU18489	myosin heavy chain	GGCCAAGAAGACTGTGGAGACAGAGAAGTCTGAGATCCAGACAGCTC TGGAGGAGGCTGAGGGAACACTGGAGCACGAGGAATCCAAGATTCTG CGTGTGCAGCTGGAGTTGAACCAGATCA
ASM4.9	SSU18489	myosin heavy chain	CCCACGCGTCCGCCCACGTCCGAAGAACGTTAACAGACTTCAGGACC TGGTAGATAAGCTGCAGATGAAAGTGAAGGCCTACAAGAGGCAGGCT GAGGAAGCTGAGGAAGCAGCAAACCAGCACATGTCTAAGTTTAGGAA GGTTCAGCATGAGCTGGAGGAGGCTGAGGA

ASM4.2 ASM8.39 ASM9.25 ASM12.29	L25609 †	tropomyosin	TCAGTTAAGCACAGACGCACAGTCTCTCTCCTCGGGTCTCTTTCGAA ACATTTGGTGGAGCCAAACAAAAAATAAGCAGCCATGGATGCCATCA AGAAGAAGATGCAGATGCTCAAGCTCGACAAGGAGAATGCCTTGGAC AGAGCTGAGGGAGCGGAGGGAGACAAGAAGGCAGCAGAGGACAAGAG CAAACAGCTCGAGGATGACTTGGTAGCTCTGCAGAAGAAGCTGAAGG GAACAGAGGATGAGTTGGACAAGTACTCTGAGTCTCTTAAGGATGCA CAGGAGAAACTTGAGGTGGCTGAGAAGACAGCCACGGACGCTGAGGC CGATGTCGCTTCCCTTAACAGACGTATCCAGCTAGTTGAGGAGGAGT TGGATCGTGCTCAGGAGCGGATGGCAACTGCCCAGACCAAGCTGGAG GAGGCTGAGAAGGCGGCTGATGAGTCTGAGAGAGGCATGAAGGTCAT TGAGAACAGGGCCTCCAAGGATGAGGAGAAGATGGAGCTGCAGGATA TCCAGCTGAAGGAGGCCAAGCACATCGCTGAGGAGGCCGACCGCAAA TACGAGGAGGTTGCCCGTAAGCTGGTCATCATTGAGAGTGATCTGGA ACGTACAGAGGAGCGCGCTGAGCTTTCAGAAGGCAAATGCTCTGAGC TTGAGGAAGAGTTGAAAACTGTGACCAACAACCTGAAGTCACTGGAG GCCCAGGCTGAGAAGTAGTCACAGAAGGAGGACAAGTACGAGGAGGA GATCAAGGTCCTCACCGACAAGCTGAAGGAGGCTGAGACTCGTGCTG AGTTTGCTGAAAGATCAGTAGCCAAACTTGAGAAGACCATGGACGAC TTGGAAGATGAGTTGTATGCCCAGAAACTGAAGTACAAGGCCATCAG CGAGGAGCTGGACAACGCCCTCAACGACATGACTTCTATATAAATTG TTTACACCTCATCAAAGACGCCTCCCGCCAGCTGAGCAGCCCTGCTC TGCTCTCTCCCGGCCTTTCTTCCATTTCCCAACCCTTCTGTCCATGC TTTCCTTTACTCCCCATGTTCTCATCCTGTCAGTGTTGTATCCCATC TCTTGTATAGAACCTCTGAATCTTTGTGTTGTATACAAATAAACCTC CTTCCTTCT
ASM4.7	SSU18487	ribosomal protein	AATTGGTCATCCTTAATATTTTCATTAAACTTGCAGAGGAGAATCCT GCCCAAACCAACCAGGAAAAGCCGTATCAAGAATAAGCAGAAGCGTC CCAGGAGCCGTACTCTGACAGCTGTGCATGACGCCATCC

ASM4.4	SSU18485	beta-globin	AAATACTGACAACATGGTTGACTGGACAGACGCCGAGAAGAGCACCA TCAGTGCTGTCTGGGGAAAAGTAGATATCAATGAGGTCGGACCACTG GCTCTGGCAAGAGTCCTGATCGTCTACCCCTGGACTCAGCCGTTATT TCGCTCTTTCGGAG
BTLC22	U79970	serum albumin	GCAGCATCTTCACCGAGGAGGCACCCAAGCTGGTTTCTGAGAGTGCA GAGCTGGTCAAGGTTTAGGATAGAATTTTGATGATTTTAATTTCAAA AGATGAACTGATGATTTAAAACAATCAAAATAAAATGAA
BTLF44	U79971	serum albumin	GACACGCAGAGTGTGTCAAAACCGCTCTGGCAGGCAGTGATGTTGAT AAGAAGATCACTGATGAGACGGACTACTACAAGAAGATGTGTGCTGC TGAGGCTGCTGTGAGCGATGACAGTTTTGAGAAGAGTATGATGGTGT ACTACACAGGATATCCCCAGGCCTCCTTCGACCAGCTCCACATGGTT CAAAC
BTLF45	STU29234	trypsin inhibitor	ACGCCGGGAGCATGAATGTTTCTGGACGGGCATTTTGTCGCATTTCG TTCGCCCCGGCCAGACCGCTCATACCCCCCTTACCTGAAAAAGATTG TGCTTCGTCGATCGACGTCGAGTGGGATCTATGTGGGGAGTCAAGAT GAAGCAGACTGTGGAGGCTATGCAGACCATATTGGATGACCTGACCA TGGATGACTACTTCA
BTLF39	STU29205	elongation factor	ATTGGCCATGTCGACTCCGGCAAGTCCACCACCACCGGCCATCTGAT CTACAAATGCGGAGGTATTGACAAGAGAACCATTGAGAACTTCGAGA AGGAGGCTGCTGAGATGGGGTAAGGCTCTTTCAAGTATGCCTGGGTG CTGGACAAGCTGAAGGCTTGAGCGTGACGTGGTATCACCATCGAGAT TCCTGTGAGTTGAGAC

BTLE36	STU18486	fibrinogen	CAGCCCTACAAGGTCTACTGTGACCAGACCACACAGAAAGGAGGCTG GACCACTATTCAGAACAGGAAGGATGGTTTCTGTTGACTTTGGCCGG CGTTGGGACAACTATCGCAGCGGTTTCGGGAATATCGCTTCGACGTT CGGAAAGGGACATTGCAACACTCCAGGGAGTACTGCTGGTACGACCG CATCAGTCAACTG

BTLB27 BTLE42	U18488 ‡	apolipo- protein AI	CAGACCACCATCATGAAATTCCTGGCTCTTGCACTCACCATCCTGCT GGCCGCAGCTACCCAGGCTGTTCCCATGCAGGCTGATGCTCCCTCTC AGCTGGAGCATGTGAAGGTAGCCATGATGGAGTACATGGCTCAGGTG AAGGAGACCGGACAGAGGTCCATCGACCTTCTGGATGACACAGAGTT CAAAGAGTACAAGGTGCAGCTGTCCCAGAGCCTTGACAACCTACAGC AGTATGCCCAGACCACCTCCCAGTCCCTGGCCCCCTACAGCGAGGCC TTCGGCGCTCAGTTGACTGATGCCGCCGCCGCCGTGCGCGCTGAGGT CATGAAGGACATGGAGGACGTGCGCACACAGCTGGAGCCAAGCGCGC CGAGCTCAAGGAAGTCCTGGACAAGCACATAGACGAGTACCGCAAGA AGCTGGAGCCCCTGATCAAGGAAATCGTTGAGCAGCGCCGCACCGAG CTGGAGGCCTTCAGGGTTAAGATGGAGCCCGTTGTGGAGGAGATGCG CGCCAAGGTGTCCACCAACGTGGAGGAGACCAAGGCCAAGCTCATGC CCATCGTGGAGACCGTCCGTGCCAAGCTGACCGAGCGTCTGGAGGAG CTGAGGACCCTGGCCGCCCCCTATGCTGAGGAGTACAAGGAGCAGAT GTTCAAGGCTGTTGGAGAGGTGCGCGAGAAGGTGGGGCCCCTGACCA ACGACTTCAAGGGCCAGGTGGGCCCCGCCGCCGAGCAGGCCAAGGAA AAGCTCATGGATTTCTACGAGACCATCAGCCAGGCCATGAAGGCATA AACACGCTCTCAACCGGACCCTCCCTCCCTCCCTTCCCGTCTCACTC ACACTGACTCACACACCATACGTACCACGCTAATGCCAAACTGATGC ACTTCCTCTGCAGTGACATGGCAGGACTCTTGCTCTCTCTAACCACC ACCACATGCGCTCAAGCGCACGCAAGCGCAGACACTAACACACTATT GCATACATTCAGTTTTGAACTGTGTTCAGGGCCTGTGTGCACAATCC TGGGCCTGCACAAATTCGACTGTACTATGAACATCAAGTGTGAATAT TCTTGTGTTGCTGCTTGTTCAAGATAGCTGTGTGCTTCGAAACAGCT CAATAAACACCATACTGTTTCATACT

BTLA40	STU18488	transplanta- tion antigen	ATCTACTTGGTCGTCTCGCTGCCATTGTGGCGAACGAAGTTCTGCTT GGCCACAAGGTGGTAGTTGTGAGATGTGAGGGTATCAACATCTCTGG AAACTTCTACCGTAACAAATTGAAGTACCTGGCTTTCCTGCGTAAGA GGATGAACACCAACCCCTCACGTGGACCATATCACTTCAGAGCGCCC AGCAGAATTTTCTGGAGGACCGTAAGGGGCATGCTGCCTCACAAAAC CAAGAGGGGACAGGCTGCACTGGAGAGGCTGAAGGTGTTCGATGGTG TCCCACCTCCTTATGACAAGAGGAAGCGCATGGTCGTACCTGCTGCC CTGAAGATTGTCCGTCTGAAGCCCACTCGCAAGTTCGCCCTCCTCGG GCGTCTGGCCCACGAGGTCGGCTGGAAGTACCAGGCCATCACAGCCA CCTTGGAGGAGAAGAGGAAGGAGAAGGCCAAGATCCGGTACGCCAAG AAAAAGACAGTGACCAAGCTGTCAAAGCTGGCAGAGAAGAACGTGGA GAGCACGATTTCAAAGTACACCGCTGTCCTGAAACAATATGGTGTCC TTGTCTGAGCTGGTTTCGTGGTGGCCAATAAAGATAAATAAATGTTT ATAGTTTTTT

APPENDIX F

Partial sequences of cDNA clones for which no putative identifications were made following comparison with GenBank/EMBL data base entries. All of the DNA inserts have been sequenced using the T7 primer (5'-d[TTAATACGACTCACTAT]-3')

Clone	size (bp)	sequence (5' to 3')
BTLA46	800	TTTTTTTATATTCAACAATTCACTTTTATTGGCTAGTTGTTCTTCCACCATGTCTCTGATCT CTCGGATGCTTCTGTGCTTGTCCAGACCGTCTGTAGTCGGGGTGTCTAGACGGCGTGAATCT TCTTCCAGCGAGCCAGAGGACAGTCGAGGGGATGTACTTCACCCCACAGGCCATCAGGAATC ACACCCTTTT
BTLB6	1200	TTTTATGTAAACGCAACTCAATGCCCACTGCAATGTGACCCTGAGCATCACAGAAGGCATTA CAGAGGTCTCTAAACACTCCTGCAACACCTGCCCCACTCTGCTGCCCCTGCATAGCCCCAGG CCTAGAGAGTGTGAAGGCAGGCCTCAGGCAGTTCAACATGGACCATAACTACACCTCCTACT TCAAACTGATGGAAGTAGGGAGGATAATATCGGTACATCATGATGAC
BTLB13	1700	AAAAGCAATAATGCTGGGCTTTAAGTTGGGTGAGGACAATCCGAAATAACGCATGATCAAAC GTCTTATCATCCGCTAAGAATGTCCCGTCTAAATCTGTTGCAATCAGCTGAATCTTTGCCAT GAGTCGTGTATTCCGCTTCTATTTTTTCGCTCTAGCAGCCGTCTATGTCTATTTTTAGTTTA CCACATTTATATCATGACAAATGTGACTAAGTACTTACGCATACGCTAGAATAGGTTGGGCA ATC
BTLB28	900	TGGTGGCCACCAGGTTGATGATCGCCAGTACCTGGCCAGGCTCCAGGGAGCGGCCGGAAAAC AGTTCAACCAACACCGAAGCACTGGGCTTGCGTTCGTCACGCACGAACACCGAGCTTTTCGG GATCGCCAGGTGCACGCGGGCGCCCTTGACGTTGTTCAGGCTGCTGATGGTGCGCGCCAGTT CGCCTTCCAGGCCACGCGATAACGGTGGCTTCATGACTGGCTGGTGCCAGGCCTGGTCTTGC GAGGATTCAAAC

BTLD6	800	CCCACGCGTCCGCTCCTGGACCCGAAGACACCATGAAGCTGCTCTCCCAGACACTGTGTCTG TGCCTGGTGCTGGCTCTCAGCATGCCCACCACCAAGCTGGACATCAAGATGGAGG(ACATGA GGG)$_7$
BTLD3	800	TTTAAGATCCACAGAGTACAGGAAGTGGGAGTTCTACAGCCAGCTGCCATCTCTGTATATGA ATACTACAACCAGAAGCCACCTGTGTGAAGTTC
BTLB39	850	AGAGTTAAAGGGACAATCTGCAGTTCAAACAACAACAAAACAGAATTGTAACCACCCTAAAA TTCATGGTCTCCGCTATAGATGCAAGGATAGTATAGATTTAACCATGTCTTGACGCTATACG GTGACAGTGTTTCTTTACGATTACATTGTTTACAAACAATAGAGTAAACAATCTATTATGGG TTCGGAAGGGCATAAACGT
BTLC18	800	CACAAACCTAATTTCACACCTAAGTGGTAAGCAACTAATAAACTGAAGACGTTTGGAGATCC AGGTAAAGAATGAAGGGTGTTGTAGCTTTGCTCCCGGT
BTLB10	200	CGCGTCCGCCCTTTTTTTTTTTTTTTTGAAGTGAACACATTTTTATTGTTGCATCAGTAGCT TGGCTTGGACGTTCTAACAAGAAGTTCTAACGTAGTTCTTAGAACCCCTTTCTGTTCTTAGA ACAGTTATTTGACCATTGGCAGTACCTGATTGGTTCAGAAGTACCTGGTTC
BTLC16	150	AACAAACAACCAGGTAACTTCTGAACCAATCAGGTACTGCCGATGGTCAAATAACTGTTCTA AGAACAGAAAGGGGTTCTAAGAACTACGTTAGAACTTCTGTTAGAACGTCCAAACCAAGCTA CTGATGCAACAATAAAAATGTGTTCACTTC(A)20
BTLD19	400	CCGAAATGTGCTTTACATTCCTTTGATAAAACAAGTGTTAAATGAATGAAGCCTTAAATTAC TACTACTGGTTCAATTCTACTTCAACGTTAACCAAACAAGAACTGGTTATGGTACTATTTCT AGATTACAGGTTTCTGAAATAATAAAATCACCT
BTLD27	500	AACTCAGATCAGGGCGGTTTGTTGACCCCACTGGCGCTACAACATGTGGACAGGAAGGAAAT CCCATGTGCATTTCCGCTAATAGCTGGTGGGCATGTGTGAACTACTACCTCTCAGTGATCCC GTTCCTGGCTGCTGTGCAGAAAGGCCTGATTGGAGATGGTCTCATTCAGGTCCAGGTACAGG CACCAGCTGAGGCAGCTGAAGACTACTGCACCTCTTACACAGACTGCTCTGCTAG

BTLD40	1500	GCGTCCGACCAGCCCTACAAGGTCTACTGTGACCAGACCACACAGAAAGGAGGCTGGACCAC TATTCAGAACAGGAAGGATGGTTCTGTTGACTTTGGCCGGCGTTGGGACAACTATCGCAGCG GTTTCGGGAATATCGCCTTCGACGTTGGAAAGGGACATTGCAACACTCCAGGGGAGTACTGG CTGGTAACGACCGCATCAGTCA
BTLF2	2000	TGGACCTTGCAGAAAGGGTTTCAGGCGCGGGGTCCAGTGGTCTGGTGGCGTCACCCTAACTC AGTAGGCAGCAAACTTTTAATACGAAAAATAGCTAACCCACTGTTAGGCGGTTCTATCGATA GTGCACAACAGTATTAAACCCGGCGTTTTATTGCGTTACGTTCTACGGGCGCCACGCTCGGC GACTGCAAAACC
BTLF1	800	AAGTTTATAGATAGGCGAGTACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTT CCTTCTCATACTCATAAGTTGTAAAATTCACAGCGGGAATAATGGAAAAAACGCCGTAAAAC GGAAGGACAAAAACGAGTGGGTCTTTGCGACCACTTTCATTTTCTACGACTTCTAGTCAACC CACGTGCTCACCATGTAGCTGACTAGAGTGTCGCATCGACTCTCAAGCACTCTGCAGTACTA TCCTGGAATCAGACGTAA
BTLF3	800	TGGAACGTGTTTGGGCCCGTCCGCTCCCCAGGTCACCAGACCACCGCAGTGGGATTGAGTCA TCCGTCGTTTGAAAATTATGCTTTTTATCGATTGGGTGACAATCCGCCAAGATAGCTATCAC GTGTTGTCATAATTTGGGCCGCGAAAATAACGCAATCGAAGATGCCGCGTGCAGCCGCTGAC GTTTTGGAGAATACGTCTGACGTCGTTCTACACGCATCTATGACCCTAA
BTLH28	1700	GAATCGTCGCCGACACGCGTTAGCCAGTCACATTGAAACCCAGGCGCGACAGGCCGATGGGA CATTGCTGTCGGCCCCGGCAATGCGCTTGTGAAACTGCGCCACCTGGGCCAGGTCGCCGGTC TGTTCGGCGAAAACATCGCCATGGTTTCACCAACGAAGGATATCGTACTCAGACATGGCATT CTCGCGCGGCTGCAGAGTCAGCGCACTCTTCTACTCACA
BTLH27	600	TCATAAGTTACTAAGGAATTACAGTTAGCAACACCAAGTCATATTACATATTCATCACAGGA TGATGCACTACGCTGCAGATCATACTCAGGTGTGTGTTGTCAGTTAAGGGCCTATTACCTGC TGAGAACCACAAGTACCTTATTGGTAGAGAATTAGCTACACAGCAACACGAGAAAACTCTGT ATGCTGTGATCATCATGACTCATTGCGTGTTATGTG

BTLH21	2800	CAAGGTCATCCTGATTTCTCTGCAGGGTGGCATCGGTTATTGGCGCTACGGCATCGAGCGCT GGTGGAGTTGGCCGAGCGCGCGGTGCAGTTGATTCTGGTGCCGGGCGATGATCGCCGACCCG GAACTCAGCGCGCTTAAGCACCGTTGGCATGGCTGAGCGCGACCGCTCTGCACTTCTGCGCA GCGTTTAA
BTLF41	1000	TGCTGGACATGATTGAACTGGACCCGGCGGTGTTCATGGAGCGCCACGCCCGTGAATTATCG GGCGGGCAAAAGCAACGTATCTGTATCGCCGCGCGCTGGCGCCGACCGCAACTGATCATTGC ATGAAGTACTG
BTLF26	1000	CCATGAAGTCGTTCTTGATCAGGCGACGGATCTTTTGCGCCTTGAGGTAGTACGCGTCGTAG TAACCGGCGGAAAGCGCGTAGGCACCGACCATGAGCCGGCGTTGTACTTCGGCACCGCCGCC TTCCGCCACGGGAGCGTTTGTACAGGTCGGTAAGGTCTTCGGGTTTTCGCAGCGGTACGCGA AGCGCACCGCGTCGAACGCGACAGT
BTLD3	800	GCCACAGAGTACAGGAAGTGGGAGTTCTACAGCCAGCTGCCATCTCTGTATATGAATACTAC AACCAGAAGCACTGTGTGAAGTTCTACCACCCACAGAGGGAGGGTGGTACTCTGAGCAGACT GTGTCTTGGAGACGTGTGCACGTGTCGAAGAGAGCTGCATTGCAAAGAAGTGGACAATTCTC
BTLD10	700	TCCTTAGGAGGGATTCCAAGGCTTTACCGGTTACCGCGTTGCATTGCGACGTCATAGCTCTT CTATAGTGTCACCTAAATTCAATTCACTGGCCGTCGTTACAACGTCGTGACTGGAAACCTGC TACCACTTATCGCTTGCAGCACACATCCTTGC
BTLC5	200	AAACATACAAACTTAGGGCCAAACAAACACGATAAAGGGCAACAGTCGAACACCAGTCGAAT TGAGAGGGACCAGACCAATAAAGTGCCACCGTTTTGGTTACGTTACGATGAATACTACGTTC ATGAGGACAGAATAGCTGATTACACAAAGATGTGATGTGGACGCTCTACACTTATAGACG
BTLC6	200	CCAGTGAATTGAATTTAGGTGACACTATAGAAGAGCTATGACGTCGCATGCACGCGTACGTA AGCTTGGATCCTCTAGA
BTLC23	200	CCACACCAGCAGCACCACGGCCACCGGGATTTCGAAGGCCACGCCAAAGGCAAGAACAGCGT CA

BTLE13	200	CGACCCACGCGTCCGAGA(AGC)$_4$A(ACA)$_5$GCA(ACA)$_{11}$GGCAGGCA(ACA)$_6$GGCA(ACA)$_{11}$A
ASM1.41	700	TTAGAGGTTGAAAAACCAATAATAAGATGGAAATGGTTAGAAGAAAGCCCACCTGTTGTCTT CATTTTTTTTTTATATGTTATTTTATCATTAACGTTCAATTTTGTAGGACCGTTAAAGGGGG GGGAAGTGAATGTTACAATGACACAACACATTGTGTTAGGCTCACTGAGAATTTGA
ASM1.13	800	TGAGGCAACGCTGACAACTGGAGTGACAGAGGATGACACTAGCATCATTGACTCTGACTCTG TGAAAAAGGATGACGCTAGCAGCACTGACTCTGGTCAAGGCAGCAATGAATATCTAGAGGAC TGATCCAATGATGACAACTGGAGCCGTTGGGGAGTGACTGACAACTCATCTAAATCTAGCAA ACAACATGTTATT
ASM1.17	500	TCTGGCTGCGGTGGCTAACAAGGAAGCTAACGTGTTATCTGCAGACCTAACGGTGTCAGGAC ATCCACTCCCCCACGGTTACTTCACCACCACCTGATAAAGAGACAGGGGGAAAATGGACTAT GAGATGTAATTTTTGTTGTTTGTTTCATTC
ASM1.2	200	CTACCCATCAAATGGCTGCTGGGCTGACGACCACTGGAGAATGGCGAAGGAAGACTTGAGAA TTGAATCTTCGTTCGTTTTTAATTCTGGCCCTTATACCATCAGCCCCCTGCACAGCTGATTT GTTGTAAAGACCATAATAAATCTTTGGTTACTAACCAC
ASM1.18	300	CACAGTCACGACAGAGCTTACATTCGCTAGTACGCGACAGAATAGTGCGAGGATGAGCAACG AGCGGT
ASM1.4	300	AGAGTTTATCTCCTGAAGATAAGGTGGGGAAATAGAAGATATTCTAAGAAGGAATGAGGCTA GAGAAAAACATGGCCTCCTTAAAGTGAACGGCTACTTTTTCTCAA
ASM1.5	400	AACGCTCGTTGCTTTCAAAACGTTTCAATTTATTTTTCTCTCTTTTTTGGTTGTTTTGATTT GTTTGTTTTAGTGTTGTAGAAACTGGACAAGGAGAAACGGTTGATAAAAAGTGGAATAAAGT GAGTGTCTCTAT

ASM1.1	300	AGGACTGGACTCCAATGGTGATGGAGAGGTGAACTTTGAGGAGTTTGTTTCTCTAGTTGTGG CCTGTCCATCGCCTGTGAACAGTGCTACCAGATGCACAAGAAGAAGATGGGGAAGTGAGGGA TGAAAAGAGGAGAGGAGGAAG
ASM1.38	300	CGGGTGTTCAAGCGCTACGTTGAGATTGGCCGTGTTGCTTACATCTCTTTCGGGCCCCACGC AGCCTGGTGGCCATCGTCGATGTCATCGACCAAAACAGAGCATTCCTCCATGGTCCTGCACA GGGGTGAAGACAGTCAATGCCTTTCAAGTGCATGCAGTTCACTGACTACGTCATTAAAGTCC ACACAGCGCTCGTAGAAGTTTGTGAGACGTGCGTGGAGA
ASM1.20	300	GTGTCAAAATGCGTTACGTTGCTGCATACCTGCTCTCTGCCCTTGGTGGTAATGCCAGCCAG CCCACAGGCTGCGGACATCAAGAAGATCCTGGAAAGTGTTGGCATTGAGGCTGACAACACAC GCATGGAGAAAGTGTGA
ASM1.40	400	GCGTCCGGCGTTCGTGGCCGCTTACCTCCTGGCTGTCGTCGGTGGCAACACTAGCCCCTCCT CAAAGGACATCAAGAACATCTGGGGAGTGTGGAACTGAGCGCGAACG
ASM4.3	800	ATGTATGCAGGAGTAAATCTGACTGAACTGACGATAAGACAGAGAGAAACACACAGAAGATA TATGTTAAGAGTGAATGTTAATGTTCTCTTCAGAATATATA
ASM4.5	750	GGTTGTTGTCCTTGATGATCTTTTAGTTTTTTTCACCACTTCACTCGTTTCTGCACGGTGAT CATTTGTTCCCTACTGTTTTATGTATTCTACATTTCTCTTCATATGGATTGGAAAGGATTGT ACTCACAGTTTTACGTAAGTATTTTTGGGGACAGACTGTTT
ASM4.11	800	AAGCGAAAGTGAATTTCATTATGGCAGGCAGGCTTTTACTCCGTACCTGTACAACCTTGCAA TTGGCAAGAAATAATGTTGTGGCAAGATCACTGCCACGCGCTATGGGCACATTAAAAGGGAT TCCAACTGATGAGGAACAGGCCACTGGACTCGAGCGACGTGCC
ASM4.12	200	GGCCGGCAATGCTGGCTGGAAAAGATTGTGGATGGGCCGCCAAGCCCCTCCCCCATCAGTTA TATTTATTTTCACTCG