IDENTIFICATION OF THE ARCTIC CHAR SEX-DETERMINING CHROMOSOME

by

Christine Kwitkowski
BSc, University College of the Fraser Valley, 2004

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In the
Department of Molecular Biology and Biochemistry

©Christine Kwitkowski

SIMON FRASER UNIVERSITY

Summer 2007

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.

APPROVAL

Name: **Christine Nichole Kwitkowski**

Degree: **Master of Science**

Title of Thesis: **Identification of the Arctic char sex-determining chromosome**

Examining Committee:

Chair: **Dr. Frederic Pio**
Assistant Professor, Department of Molecular Biology and Biochemistry

---

Dr. William S. Davidson
Senior Supervisor
Professor, Department of Molecular Biology and Biochemistry

Dr. Felix Breden
Supervisor
Associate Professor, Department of Biological Sciences

Dr. Robert H. Devlin
Supervisor
Research Scientist, Department of Fisheries and Oceans Canada

---

Dr. Lynne Quarmby
Public Examiner
Associate Professor, Department of Molecular Biology and Biochemistry

Date Defended/Approved: July 23, 2007

SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

DECLARATION OF PARTIAL COPYRIGHT LICENCE

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties as the author may desire.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

No Digital Copy for SFU Library collection:

By special permission of the Dean of Graduate Studies, the author is exempted from granting permission to Simon Fraser University Library to keep or make a digital copy for use in its circulating collection (currently available to the public at the "Institutional Repository" link of the SFU Library website). It is agreed that the author may change this decision at any time, by providing SFU Library with a signed Partial Copyright Licence in standard form.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

Partial Copyright Licence_PDF Exemption 2007

ABSTRACT

Sex-determination is the most pivotal developmental process in the life of a gonochoristic organism. The purpose of this thesis was to identify the sex chromosomes in Arctic char (*Salvelinus* alpinus). The sex-determining gene, *SEX*, was mapped to Arctic char linkage group 4 (AC-4). I have used probes designed from flanking regions of sex-linked markers to screen a fosmid library. Fosmids containing sex-linked markers were isolated and placed on Arctic char chromosomes by FISH, identifying chromosome 3 as the sex chromosome. A fosmid containing Sal05SFU was particularly informative in this analysis and was sequenced using a combination of shotgun cloning and selected primer walking. Annotation of the Sal05SFU sequence identified the first five exons of *Suppressor of fused* (*Sufu*), a candidate sex-determining gene. Additional candidate genes for sex-determination were identified through comparative genome analysis.

Keywords: Arctic char, sex-determination, *Suppressor of fused*, sex chromosome, fosmid library

DEDICATION

To Dad

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. William Davidson for giving me the opportunity to realize my goal of achieving this level of academic accomplishment, and also for his continuous guidance and support. I would also like to thank the entire "Team Davidson" for all of their help and patience.

Huge thanks are owed to Lisa Howard, for all of her tireless help and support, and a few laughs along the way.

Last but not least, special thanks are owed to my husband Joey, and all of my family, extended family, and friends who have always encouraged me and believed in me during my endless years of education. I couldn't have done it without you all!

TABLE OF CONTENTS

Approval ii
Abstract iii
Dedication iv
Acknowledgments v
Table of Contents vi
List of Tables ix
List of Figures x
CHAPTER 1 INTRODUCTION 1
1.1 Sexual Reproduction 1
1.2 Mechanisms of sex-determination 3
1.2.1 Genotypic sex determination (GSD) 4
1.2.2 Environmental sex determination (ESD) 4
1.2.3 Multiple-factor sex determination (MSD) 5
1.3 Sex-Determining Pathways 5
1.3.1 Description of sex-determining pathways in general 5
1.3.2 Sex-determination in humans 6
1.3.3 Sex-determination in *D. melanogaster* 8
1.3.4 Sex-determination in *C. elegans* 11
1.3.5 Sex-determination in *A. mellifera* 13
1.3.6 Evolutionary conservation in sex-determining pathways 15
1.4 Sex-Determination in Fishes 16
1.4.1 Teleost fishes 16
1.4.2 Sex-determination in medaka 16
1.4.3 Sex-determination in threespine stickleback 18
1.4.4 Sex-determination in Xiphophorus 18
1.5 Salmonid Fishes 19
1.5.1 Family Salmonidae 19
1.5.2 Sex-determination in Salmonids 22
1.6 Arctic Char (*Salvelinus alpinus*) 27
1.6.1 The Arctic char karyotype 27
1.6.2 Sex-determination in Arctic char 28
1.7 Identification of *SEX* 29
1.7.1 Approaches previously used to identify and characterize *SEX* 29
1.8 Aim of the Thesis 30

CHAPTER 2 MATERIALS AND METHODS **31**
2.1 Fosmid Library Screening 31
2.1.1 Probe design from Arctic char-specific marker sequences 31
2.1.2 PCR-screening of fosmid mini-libraries 34
2.1.3 Colony lifts 34
2.1.4 Hybridization/Probing 35
2.1.5 Confirmation of positive fosmids by PCR 37
2.1.6 Characterization of positive clones by *HindIII*-digestion and Southern blotting 38
2.2 Sequencing and Analysis of Clone Sal05SFU-6 39
2.2.1 Sal05SFU-6 shotgun library construction 39
2.2.2 Sequencing and analysis of Sal05SFU-6 40
2.2.3 Extension of Sal05SFU-6 contig sequences by selected primer-walking 42
2.2.4 Fosmid end-sequencing 42
2.3 Fluorescent *in situ* Hybridization (FISH) 43

CHAPTER 3 RESULTS **44**
3.1 Resources Available at the Start of the Project 44
3.1.1 Genetic linkage maps 44
3.1.2 Fosmid genomic library 46
3.1.3 Approach to identify *SEX* in Arctic char 46
3.2 Designing of Marker-Specific PCR Primers/Hybridization Probes 47
3.2.1 Primer design 47
3.2.2 Testing primers on Arctic char genomic DNA 47
3.2.3 Sequencing amplicons to get Arctic char-specific sequences 49
3.2.4 Designing hybridization probes from Arctic char-specific sequences 58
3.3 Fosmid Library Screening 60
3.3.1 PCR amplification from mini-libraries 60
3.3.2 Colony lifts and hybridization using 40-mer probes 62
3.3.3 PCR confirmation of positive fosmid clones 64
3.3.4 Characterization by *Hind*III-digestion and hybridization of Southern transfer 68
3.4 Fluorescent *in situ* Hybridization (FISH) 70
3.5 Sal05SFU-6 Sequencing 72
3.5.1 Shotgun library construction 72
3.5.2 Shotgun sequencing of Sal05SFU-6 74
3.6 Selected Primer Walking 75
3.6.1 Extension and orientation of Sal05SFU-6 contigs 77
3.7 Sequence Analyses 79
3.7.1 Sequence similarity searches 79
3.7.2 Gene descriptions 79
3.8 Characterization of Arctic char *Sufu* 80
3.8.1 Sufu amino acid alignment 80
3.8.2 Phylogenetic analysis of Sufu 84

3.8.3 Comparative genomics ... 86

CHAPTER 4 DISCUSSION ... 91
4.1 Fosmid Library Screening ... 91
4.1.1 Arctic char-specific primer and probe design ... 91
4.1.2 Arctic char Linkage Group 4 (AC-4) ... 92
4.2 Fluorescent *in situ* Hybridization of Fosmids Containing Sex-Linked Markers ... 93
4.2.1 Identification of the Arctic char sex chromosomes ... 93
4.2.2 The Arctic char sex-determining region ... 94
4.2.3 Possible locations of *SEX* on the Arctic char sex chromosome ... 94
4.3 Sequence Analyses ... 98
4.3.1 Genes identified in Sal05SFU-6 ... 98
4.3.2 Candidate Arctic char sex-determining genes ... 101

REFERENCES ... 103

LIST OF TABLES

Table 3.1: Primers designed from sex-linked markers used to screen fosmid library 48

Table 3.2: Hybridization probes designed from Arctic char-specific marker sequences 59

Table 3.3: Summary of roles/functions of genes identified through comparative genome analysis 89

LIST OF FIGURES

Figure 1.1: Schematic diagram of the human sex-determining pathway 7
Figure 1.2: The *Drosophila melanogaster* sex-determining pathway 10
Figure 1.3: Models for somatic sex-determining pathways in *Caenorhabditis elegans* 12
Figure 1.4: Model for sex-determination in *Apis mellifera* 14
Figure 1.5: Phylogenetic tree of Salmonid fishes 21
Figure 3.1: Arctic char Linkage Group 4 (AC-4) 45
Figure 3.2: Arctic char-specific marker sequences 50
Figure 3.3: 1.5% agarose gel showing amplification products from a Sal05 SFU mini-library PCR-screen 61
Figure 3.4: A phosphor-image of a membrane hybridized with the Sal05SFU 40-mer probe 63
Figure 3.5: 1.5% agarose gel showing PCR products of Sal05SFU clones 65
Figure 3.6: PCR results of SsaBHMS431 amplified from 22 potentially positive fosmid clones using R primer only 67
Figure 3.7: 0.8% agarose gel of *Hind*III digested Sal05SFU PCR positive clones and phosphor image of hybridized Southern membrane 69
Figure 3.8: Results of Fluorescent *in situ* Hybridizations 71
Figure 3.9: 0.8% agarose gel showing *PVUII* digest of Sal05SFU-6 shotgun clones 73
Figure 3.10: An illustrative diagram of the approach used to extend sequence coverage of Sal05SFU-6 76
Figure 3.11: Illustrative diagram summarizing sequence coverage and contig orientation of Sal05SFU-6 78
Figure 3.12: Amino acid alignment of *Sufu* 82
Figure 3.13: Neighbor-joining tree based on the amino acid alignment of Sufu 85
Figure 3.14: Comparative analysis of *Sufu* across other fish species 87

CHAPTER 1 INTRODUCTION

1.1 Sexual Reproduction

Sex-determination is one of the most pivotal developmental processes in the life of a gonochoristic organism. This biological decision not only affects reproductive fate, but also influences subsequent behavior, morphology, and physiology (Beye et al., 2003; Brennan and Capel, 2004).

Sexual reproduction involves the fusion of haploid gametes to form a diploid zygote with genetic contributions from both parents (Charlesworth, 2006). For centuries, evolutionary biology has sought to understand why sexual reproduction is so common (Kondrashov, 1993). Searching for a mate, courtship and the act of mating make sexual reproduction energetically expensive, and significant energy is also required to retain two of each haploid genome within each cell (Lively and Loyd, 1990). In addition, there are risks of predation and infection by sexually transmitted diseases (Agrawal, 2006). Lastly, there is the 'two-fold cost' of sex: sexual organisms must produce twice as many offspring as asexual organisms to retain the same genetic contribution per capita. Surprisingly, despite these drawbacks, sexual reproduction is nearly ubiquitous among eukaryotes (Otto and Gerstein, 2006).

There are many hypotheses for why sexual reproduction is so prevalent, and how it has been maintained. It is often assumed that an advantage to sex is generation of increased genetic variability in fitness; however, population genetic models show that sex and recombination can actually reduce variability in fitness

when other forces (such as selection and migration) promote genetic diversity (Charlesworth, 2006; Otto and Gerstein, 2006).

The 'Mutual Deterministic' (MD) hypothesis, put forth by Kondrashov, suggests that sexual reproduction is actually maintained by a high genomic rate of deleterious mutations (U) (Kondrashov, 1993). The idea is that if accumulated mistakes can be removed by recombination later on, it may be energetically favourable to allow those mistakes to accumulate in the first place. Kondrashov predicted that sexual reproduction would be favoured over asexual reproduction when U exceeds one event per generation, assuming the newly arising asexual populations are free of mutations (Charlesworth, 1990). However, when U is below one event per generation, an asexual population would be more productive, and thus more favourable.

This hypothesis has been challenged by Keightley and Eyre-Walker (2000, 2001) who compared genomes of closely related organisms to see if U was always one or above per generation. While they found that species with long generation times had U values above two, they also found species with shorter generation times had U values as low as 0.05. Based on these observations, Keightley and Eyre-Walker conclude that the MD hypothesis cannot be the sole explanation of maintenance of sexual reproduction, and that there must be other advantages not yet explained.

The 'Red Queen' (RQ) hypothesis is based on the tightly co-evolving interactions between organisms and their parasites, predators and competitors,

and the idea that rapid evolutionary change in one may lead to extinction of another (Otto and Gerstein, 2006). The hypothesis explains that sexual organisms must constantly produce variability through sex and recombination in order to keep up to the pace of co-evolving species (Bell, 1982). It has been shown, however, that the RQ hypothesis only remains legitimate under conditions of strong selection on a small number of fitness-associated loci. Under conditions of weak selection, particularly acting on multiple fitness-associated loci, recombination during sexual reproduction is more likely to break apart more favourable gene combinations(Otto and Nuismer, 2004).

Although it is clear that none of the outlined hypotheses can completely explain the prevalence and maintenance of sexual reproduction, this does not mean they should be discounted. Together with other factors, such as mutation and random genetic drift, they may still play significant roles (Howard and Lively, 1998). It is most likely adaptive evolution that principally drives the evolution of sex, probably in combination with a large number of other mechanisms.

1.2 Mechanisms of Sex-Determination

The remarkably diverse mechanisms of sex-determination can be divided into broad but distinct categories: genotypic sex determination (GSD), environmental sex determination (ESD) (Bull, 1983) and multiple-factor sex-determination (MSD) (Leberg, 1998).

1.2.1 Genotypic sex determination (GSD)

In any genetic sex-determination mechanism, certain factors at the genetic level are inherited in the zygote, directing its development to a particular gender. Male heterogamety is a GSD mechanism in which all zygotes are either XX (female) or XY (male). Sex is determined at conception based on inheritance of the X and Y chromosomes. In the reverse of this mechanism, female heterogamety, gender depends on inheritance of Z and W chromosomes; females are ZW and males are ZZ. A unique GSD mechanism is haplo-diploidy. In this mechanism, males are the result of unfertilized eggs, and females from fertilized eggs (Bull, 1983).

1.2.2 Environmental sex determination (ESD)

In ESD organisms, sex is determined after fertilization by environmental factors (Bulmer, 1987). These factors can include temperature, salinity, pH, and various population dynamic influences (Conover and Heins, 1987). In a sub-group of ESD vertebrates, sex is determined specifically by incubation temperature during a distinct larval phase. This mechanism is known as temperature-dependant sex determination, or TSD. In organisms exhibiting ESD or TSD, sex cannot be predicted by zygotic genotype because there is very little, if any, genetic difference between the sexes (Valenzuela et al., 2003; Bulmer, 1987). It should be noted that the underlying assumption of ESD mechanisms is that they involve genes that are sensitive to the environment, not that they lack a genetic contribution completely.

1.2.3 Multiple-factor sex determination (MSD)

In some organisms, three or more sex factors act together to decide gender. This is known as multiple-factor sex-determination, or MSD. A diverse set of taxa, including fish, exhibits MSD. It is thought to be extremely rare, but a lack of sex-linked markers has made the true frequency of MSD difficult to assess (Bull, 1983).

1.3 Sex-Determining Pathways

1.3.1 Description of sex-determining pathways in general

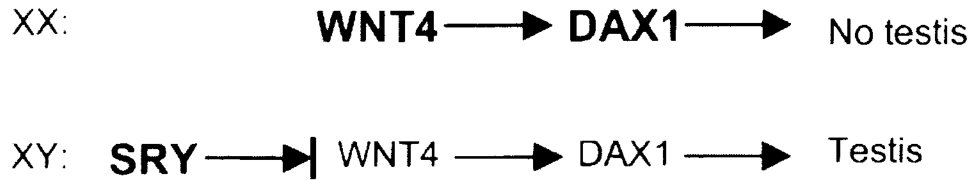
In general, a sex-determining pathway can be defined as an intricate series of biochemical reactions resulting in specification of sex-cells and subsequent differentiation of these cells (Devlin and Nagahama, 2002). Substantial sex-determining pathway characterization has only been performed on a small group of multi-cellular animals: the fruit fly (*Drosophila melanogaster*), the nematode (*Caenorhabditis elegans*), the Japanese medakafish (*Oryzias latipes*), the honeybee (*Apis melllifera*) and most mammals, including humans. However, the downstream components of their sex-determining pathways have not all been elucidated and our understanding of how these pathways evolved is very limited (Beye et al., 2003; for review see Stothard and Pilgrim, 2003). For most other species, the sex-determining pathways remain poorly understood, despite tremendous efforts over decades of study.

1.3.2 Sex-determination in humans

Humans and most other mammals have a strictly chromosomal system of sex-determination (Marshall Graves, 2002). In this system, males are the heterogametic sex (XY) and females are homogametic (XX). It is that smaller Y chromosome that bears the master testis-determining factor, *SRY* (Scherer, 2002). While it is clearly established that *SRY* presence is what induces male differentiation, its direct target gene(s) remain unknown (Vilain, 2002). Being a member of the HMG domain family of transcription factors, *SRY* most likely acts as a transcriptional regulator, activating or inhibiting the production of elements downstream in the pathway.

It has been observed that not all XX male individuals express *SRY*, suggesting that they lack proper expression of a downstream male-repressor, the elusive target of *SRY.* Based on this, it is a longstanding thought that *SRY* acts in a "double-repressor" system, in which it inhibits other factors that inhibit masculinization (e.g. DAX1) (McElreavey et al., 1993; Vilain, 2002; Scherer, 2002). One current model of this system, born from expression studies, suggests that DAX1 is inhibited by *SRY* through the action of WNT4 (Figure 1.1) (Vilain, 2002). It has also been suggested that *SRY* acts only as a switch to initiate the male-development pathway, and contributes very little to male differentiation itself (Pask and Marshall Graves, 1999). It is most likely that there are a number of other genes downstream of *SRY* that act in the human sex-determining pathway, and have yet to be placed in a hierarchy of action.

Figure 1.1 **Schematic diagram of the human sex-determining pathway.** In XX individuals, *WNT4* expression up-regulates the expression of *DAX1*, which subsequently prevents formation of testes, and ovaries develop. In XY individuals, expression of *SRY* inhibits the action of WNT4, leading to DAX1 inhibition and testes formation. Genes written in bold are "ON", while genes not written in bold are "OFF".



1.3.3 Sex-determination in *D. melanogaster*

It is well established that sex is determined genetically in the fruit fly, *Drosophila melanogaster* (Figure 1.2) (Schütt and Nöthinger, 2002; Pomiankowski et al., 2004). The primary genetic signal for sex-determination is the ratio of X chromosomes to sets of autosomes (X:A). A ratio of 0.5 (1X:2A) results in male development, while a ratio of 1.0 (2X:2A) results in female development (Schütt and Nöthinger, 2002). The ratio is ultimately formed by three X-linked numerator *sisterless* genes (*sisA,B,C*) and *runt* (*run*) with the autosomal denominator gene *deadpan* (*dpn*). The products of these genes are transcription factors that regulate the expression of the gene *Sxl* (*Sex-lethal*). Downstream of *Sxl* is a cascade of regulatory genes which branches into three pathways, one of which controls sexual differentiation in the soma (Schütt and Nöthinger, 2002).

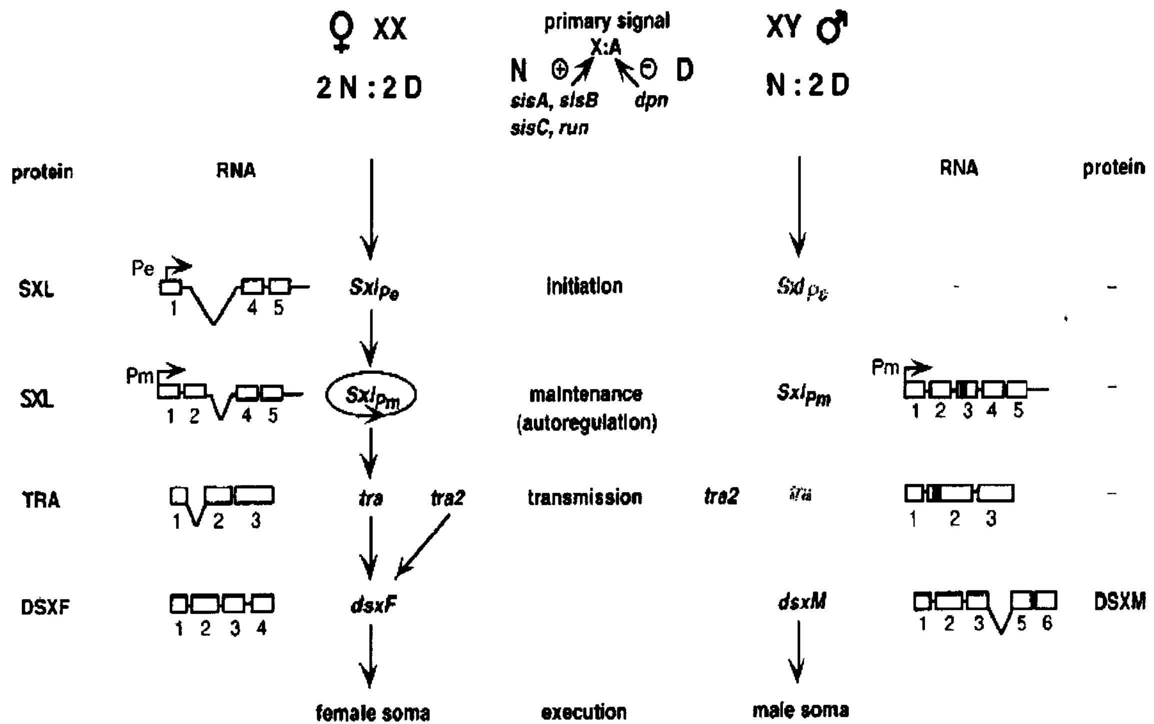
Sxl is unique in that it contains a stop codon (UAG) embedded in exon 3. If this is not removed during RNA processing, a non-functional SXL protein is produced. The double dose of X-linked numerator genes in females activates *Pe*, an early promoter of *Sxl,* which produces transcripts in which exons 2 and 3 are spliced out. This results in an active SXL protein, which also acts as a splice enhancer by binding its own pre-mRNA and ensuring the removal of exon 3. Thus, an auto-regulatory loop is activated in females by active SXL, ensuring continuous production of the protein. Functional SXL protein is not produced in males due to *Sxl* being prematurely truncated. The auto-regulatory loop is never

activated in males as a direct result of insufficient dosage of X-linked numerator genes.

The active form of SXL is an RNA binding protein that regulates production of *transformer* (*tra*), the next gene in the sex pathway. A series of premature stop codons exists in *tra* at the start of exon 2. In females, SXL protein blocks the canonical splice site, forcing the use of an alternate splice site just downstream of the stop codons. This action prevents premature termination, and results in production of active TRA protein. The lack of active SXL protein in males prevents removal of the stop codons, so no functional TRA is produced.

TRA also binds RNA, and this action forces alternative splicing of the gene *doublesex* (*dsx*), which is the next major downstream factor in the sex pathway. In females, TRA acts together with TRA-2 to initiate the alternative splicing pattern including exons 1-4, the product of which is DSXF. In males, *dsx* is spliced in its default form (removal of exon 4) due to absence of TRA. Hence, male *dsx* mRNA contains exons 1-3 and 5-6, producing the male-specific DSXM isoform. Both DSXM and DSXF are transcription factors that sex-specifically enhance or repress a number of downstream sex-specific developmental genes, which in turn initiate one of two routes of sexual differentiation.

Figure 1.2 **The *Drosophila melanogaster* sex-determining pathway.** The female developmental pathway is on the left; the male pathway is on the right. Open gene symbols indicate functional inactivity. Vertical arrows indicate the cascade of information. Boxes represent exons; only the first 5 exons (of 10) are shown for *Sxl*. Bars in exon 3 of *Sxl* and in exon 2 of *tra* represent stop codons. Naming abbreviations are as follows: sis, sisterless; run, runt; dpn, deadpan; sxl, sex-lethal; tra, transformer; dsx, doublesex. Faded italics indicate absence of functional gene product.

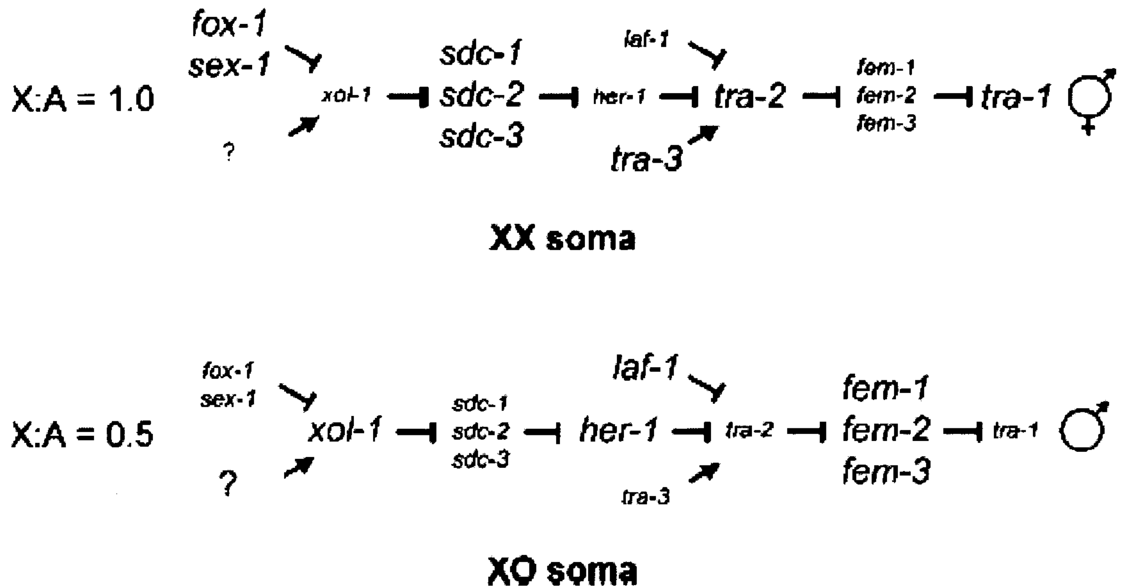


1.3.4 Sex-determination in *C. elegans*

The primary genetic signal in the sex pathway of the nematode *Caenorhabditis elegans* is the ratio of X chromosomes to autosomes, much like in the *Drosophila* sex pathway (Figure 1.3) (for review see Stothard and Pilgrim, 2003). The difference between the two mechanisms is that *C. elegans* exists as either a male or a hermaphrodite; males inherit one X chromosome, hermaphrodites inherit two X chromosomes. In hermaphrodites (Figure 1.3A), specific dosage of the X-linked genes *fox-1* and *sex-1* results in reduced expression of *xol-1*. This results in prevention of xol-1's inhibitory actions on the hermaphrodite promoting *sdc* genes, *sdc-1, 2* and *3*. These genes are transcriptional repressors of the gene *her-1*, which is a repressor of *tra-2*. Thus, if tra-2 is not able to inhibit the *fem* genes (*fem 1, 2* and *3*), they are prevented from inhibiting *tra-1*, and hermaphrodite development is promoted. Conversely, an insufficient initial dosage of *ser-1* and *fox-1* results in opposite repression effects, ultimately leading to the inactivation of *tra-1*. When this is the case, the default male pathway is initiated (Figure 1.3).

Despite the extensive characterization of the *C. elegans* sex-pathway and its major sex-determining genes, many other trans-acting cofactors have been identified. Mutations in these cofactors result in improper sexual development, and characterization of these genes is needed for our overall understanding of how they contribute to the sex-pathway.

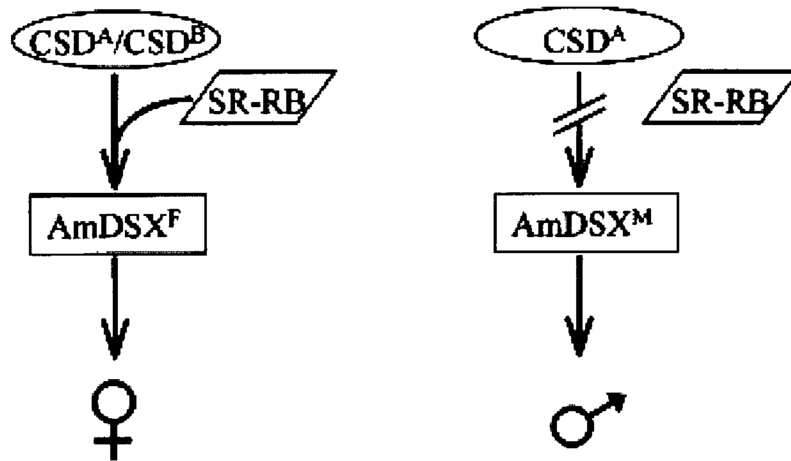
Figure 1.3 Models for somatic sex-determining pathways in *Caenorhabditis elegans.* X= copies of X chromosomes, A=copies of autosomes. Genes printed in large font represent active forms, while genes printed in small font represent inactive forms.



1.3.5 Sex-determination in *A. mellifera*

The honeybee, *Apis mellifera*, employs a haplo-diploidy system of sex-determination; that is males develop from unfertilized eggs, and females from fertilized eggs. In this system, a single locus with several alleles directs sexual development (Beye et al., 2003). In this so-called complementary sex-determining mechanism (Figure 1.4), sex is determined by the allelic composition of a single locus, *complementary sex determiner* (*csd*). The *csd* transcript is essentially the same in both sexes, but the decisive difference is that females carry two different alleles while males have only one. CSD is active only if the polypeptides are derived from two different alleles, so the single copy in males is inactive. In females, the heterodimeric CSD protein co-operatively binds an SR-type protein containing an RNS recognition motif (SR-RB protein). This complex then splices *Apis mellifera dsx* pre-mRNA. The resulting female-specific isoform, $AmDSX^F$ subsequently induces female development. In males, *Amdsx* is expressed by default to produce the $AmDSX^M$ isoform.

Figure 1.4 **Model for sex-determination in *Apis mellifera*.** The interrupted arrow indicates lack of SR-RB binding during male development.



1.3.6 Evolutionary conservation in sex-determining pathways

A number of studies have been conducted in order to aid our understanding of the origin of sex-determination, and have revealed sex-related genes common across a wide variety of organisms (Raymond et al., 1998; Ottolenghi and McElreavey, 2000). The gene *DMY*, which is expressed solely in male embryos of medaka, belongs to the family of genes that contain DM-domains. The domain was named from a related DNA binding motif found in *dsx* and *mab*-3 genes, involved in sex-determination in *D. melanogaster* and *C. elegans*, respectively. DM-domain containing genes, including *DMRT1*, have been found in many vertebrates. In mammals, this gene has been implicated in XY female sex reversal syndromes (Raymond et al., 1998). However, no point mutations in *DMRT1* correlate with a sex reversal phenotype (Swain, 2002). *DMRT1* homologues have also been identified in mice (Raymond et al., 2000) birds (Nanda et al., 1999) and reptilian species (Kettlewell et al., 2000). This evident conservation of *DMRT1* suggests that it plays a role in early testis development, and that its function has been maintained throughout evolution.

Since it is the initial signal at the start of the pathway that varies the most across species (Graham et al., 2002), identification of this signal in other organisms is the first step crucial to understanding their sex-determining pathways overall. In accomplishing this, our understanding of the evolution of sex-determining mechanisms will greatly increase.

1.4 Sex-Determination in Fishes

During the course of evolution in the fish lineage, there has been frequent switching between different sex-determination systems. This was made apparent by plotting all known mechanisms of sex-determination in fish against teleost phylogeny (Devlin and Nagahama, 2002; Veith et al., 2003).

1.4.1 Teleost fishes

Fish in the class *teleostei* display a very large range of sex-determination mechanisms. These include, but are not limited to, strict environmental influence (e.g., temperature or group dynamics), primary genetic sex-determination modulated by secondary environmental factors, and strict genetic sex-determination (Woram et al., 2003). Sex reversal has also been observed in several species either spontaneously, or after treatment with steroid hormones (for review see Baroiller et al., 1999). It is this wide variety of mechanisms that make teleosts an ideal group of organisms for use in sex-determination studies. Thus far, both the molecular and evolutionary mechanisms that drive sex-determination and its variability in teleost fish remain poorly understood. However, it is strongly believed that parts of teleost genomes are evolving very rapidly, especially their sex chromosomes (Charlesworth, 2004).

1.4.2 Sex-determination in medaka

The Japanese medaka fish, *Oryzias latipes*, is the first non-mammalian vertebrate for which the master sex-determining gene has been elucidated

(Matsuda et al., 2002; Zhang, 2004; for review see Matsuda, 2005). While this recent elucidation of a master sex-determiner in fish is promising, many studies have shown that the sex-determining gene of medaka, *DMY,* is absent from other fish species (Kondo et al., 2003).

Medaka fish employ an XX/XY sex determination system, although the Y chromosome is not cytogenetically distinct. The central sex-signal is *dmrt1bY* (*DMY*), a Y chromosome-specific duplicate of an autosomal gene called *dmrt1* (doublesex/Mab-3-related transcription factor 1) (Volff, 2005). Dmrt1 is a member of a family of proteins containing a conserved DNA-binding motif called the DM domain, and is a transcription factor ubiquitously involved in vertebrate sex-determination and differentiation (Kondo et al., 2004).

The heterogametic males possess two types of *dmrt1* genes: the autosomal *dmrt1* and the Y-specific *DMY.* Apparently, *DMY* is the only functional gene in the Y-specific part of the sex chromosome, solidifying the belief that it corresponds to the master sex-determining gene (Nanda et al., 2002). Natural mutations in *DMY* result in XY sex-reversed females (Matsuda at al., 2002). However, spontaneous sex-reversed XX males have also been observed, indicating that a true male phenotype can also occasionally develop in the absence of *DMY* (Nanda et al., 2003; Swain, 2002). Regardless, continuous functional and expression analyses of *DMY* have consistently shown it to be the master gene for male sex-determination in the medaka.

1.4.3 Sex-determination in threespine stickleback

At this point, little is known about the mechanism of sex-determination in the threespine stickleback, *Gasterosteus aculeatus*. Cytological studies have shown that threespine sticklebacks lack sexually dimorphic chromosomes (Cunado et al., 2002). Some evidence has suggested that temperature and density of rearing may affect sex ratios in threespine sticklebacks, while other evidence, such as identification of DNA markers with sex-specific alleles suggest a genetic basis. In fact, recent research supporting the genetic mechanism hypothesis successfully showed that a single major chromosome region controls sex determination in *G. aculeatus* (Peichel et al., 2004). This region shows heterozygosity in males and strong suppression of recombination, much like a nascent Y chromosome. However, candidate sex-determining genes have yet to be identified in this species.

1.4.4 Sex-determination in Xiphophorus

The platyfish, *Xiphophorus maculatus*, has three sex chromosomes: W, X, and Y. In this system, males can either be XY or YY, while females can be WX, WY, or XX (Kallman, 1984). The X and Y chromosomes cannot be distinguished cytologically, but the master sex-determining locus, *SD*, has been identified on both the X and Y chromosomes (Froschauer et al., 2002). Since platyfish are closely related to medaka, recent research has tested for the presence of *dmrt1Y* in the *SD* region of platyfish. However, the results of the experiments show that

DMY is not the master sex-determining gene in the platyfish (Veith et al., 2003), and other candidate genes have yet to be identified in this species.

1.5 Salmonid Fishes

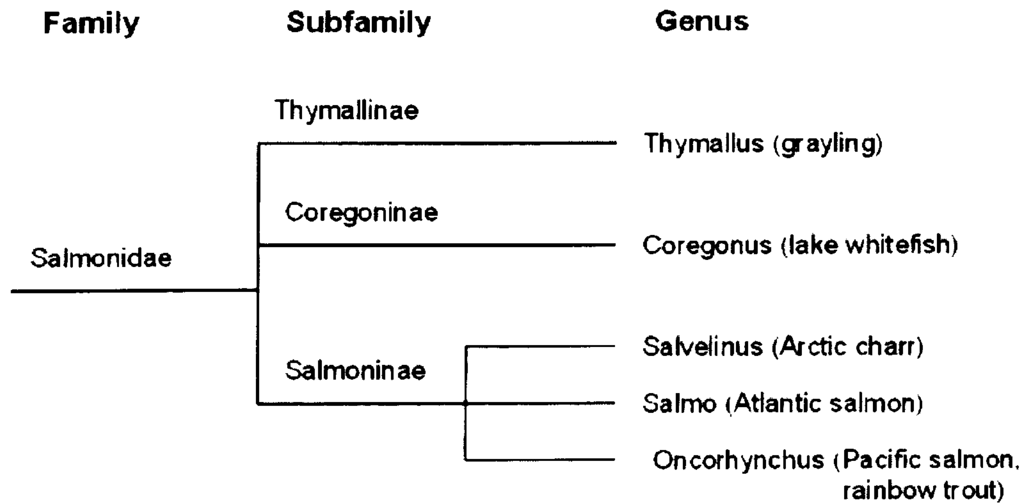
1.5.1 Family Salmonidae

There are three subfamilies in the Salmonidae family. These are Coregoninae (whitefishes and ciscos), Thymallinae (graylings), and Salmoninae (trout, salmon and char) (Figure 1.5) (Nelson, 1984). The sub-family Salmoninae has been extensively studied, as its species have long been popular in both sport and commercial fishing, as well as in the field of aquaculture. Species belonging to Salmonidae are native to the Northern hemisphere, but have been introduced to many other parts of the world as sources of game fish and farming stock. Species of salmonids pursue both anadromous and freshwater life cycles, but in either case, spawning always takes place in freshwater. Some species, such as Arctic char (*Salvelinus alpinus*), Atlantic salmon (*Salmo salar*) and brown trout (*Salmo trutta*), are known to be repeat spawners, while sockeye salmon (*Oncorhynchus nerka*) and chinook salmon (*Oncorhynchus tshawytscha*) die after spawning (Nelson, 1984).

The radiation of teleost species is most likely due to a genome duplication event that occurred at the time of their divergence. This is supported by the observation that ray-finned fishes, *Actinoptergii*, including teleosts, have two paralogous copies of many genes, while tetrapods only have one (Allendorf and

Thorgaard, 1984; Hoegg et al., 2004; Christoffels et al., 2004). It is believed that, in an event separate from the teleost genome duplication, the common ancestor of salmonids underwent a genome tetraploidization (Allendorf and Thorgaard, 1984; Hoeg et al., 2004). This prediction is based on several unique observations of salmonid DNA and chromosomes. Salmonids have about twice as much DNA per cell as their closest relatives, and also have twice as many chromosome arms. They also display a high incidence of duplicated enzyme loci, and multivalents are a common observation in meiotic preparations from many male salmonid species (Ohno et al., 1968; Allendorf and Thorgaard, 1984). Since the tetraploidization event, the genome of salmonids has been returning to a diploid state. In other words, the four homologous chromosomes have been diverging into two distinct sets of homologous chromosomes, and evidence suggests that this diploidization is not yet complete in salmonid fishes (Woram et al. 2004).

Figure 1.5 **Phylogenetic tree of Salmonid fishes.** The phylogenetic relationship between the three subfamilies of the Salmonidae family is shown, with species examples for each genus.



1.5.2 Sex-determination in Salmonids

For many years it was thought that salmonid sex determination was strictly under genetic control. However, it has since been proven that extrinsic factors, such as temperature, pH, salinity, hormones, and pollutants, can influence sex in the majority of salmonid species (Piferrer et al., 1993; Metcalf and Gemmel, 2006). Salmonids are particularly sensitive to these factors during early development (between hatching and first feeding) and it is during this stage that phenotypic sex may change (Devlin and Nagahama, 2002; Metcalf and Gemmel, 2006).

1.5.2.1 Evidence of male heterogamety in salmonids

Although sex chromosomes remain to be identified in most salmonids, male heterogamety has been a longstanding general rule for salmonids. Primary evidence of this comes from observing the sex ratios of the progeny of hormonally sex-reversed fish (reviewed in Phillips and Ráb, 2001).

When estradiol-17ß (induces female development) or methyltestosterone (induces male development) is introduced to breeding tanks during early developmental stages, fish will spontaneously reverse sex (reviewed in Devlin and Nagahama, 2002). Subsequently, when sex-reversed female salmonids (being phenotypically male while genotypically female) are crossed with normal females (phenotypically and genotypically female), all offspring are female. It then follows that females must be homogametic (XX), and males heterogametic

(XY) (Johnstone et al., 1979; Hunter et al., 1982, 1983; Johnstone and Youngson, 1984).

The identification of sex-specific markers has increased support of male heterogamety in many species including Arctic char (*Salvelinus alpinus*), rainbow trout (*Oncorhynchus mykiss*), lake trout (*Salvelinus namaycush*), brown trout (*Salmo trutta*), Masu salmon (*Oncorhynchus masou*), pink salmon (*Oncorhynchus gorbuscha*), coho salmon (*Oncorhynchus kisutch*) and chinook salmon (*Oncorhynchus tsawytscha*) (May et al., 1989; Du et al., 1993; Forbes et al., 1994; Prodöhl et al., 1994; Young et al., 1998; Nakayama et al., 1999; Sakamoto et al., 2000; Devlin et al., 2001; Zhang et al., 2001; Stein et al., 2002). For each species, marker PCR product length polymorphisms have been observed such that males were heterozygous at the allele, and females were homozygous. Genetic linkage mapping of the markers placed them within the sex-linkage group, suggesting that the polymorphisms observed are differences between alleles located on the Y and X chromosomes, respectively (Young et al., 1998; Sakamoto et al., 2000; Devlin et al., 2001).

1.5.2.2 Salmonid sex chromosomes

Although morphologically distinct sex chromosomes are generally not visible in salmonids, heteromorphic sex chromosomes have been identified in rainbow trout (*O. mykiss*), lake trout (*S. namaycush*), least sisco (*Coregonus sardinella*) and sockeye salmon (*O. nerka*) (Phillips and Ráb, 2001; Stein at al.,

2002). The sex chromosomes have small sex-specific regions and large pseudoautosomal regions (Devlin and Nagahama, 2002).

Heteromorphic chromosome formation often includes addition of heterochromatin (Phillips and Ráb, 2001). In both brook trout (*Salvelinus fontinalis*) and lake trout, it has been observed that the largest pair of submetacentrics have an X-specific heterochromatin block at the end of their short arms, thus identifying them as the sex chromosome pair (Phillips and Ihssen, 1985; Phillips et al., 2002). Significant size-differences have been observed in comparing sex chromosomes of different rainbow trout (*O. mykiss*) populations, suggesting an early stage of visible sex-chromosome differentiation. Rainbow trout populations lacking heteromorphisms have also been observed, indicating that chromosomal rearrangements differentiating the sex chromosomes of this species are still in the process of fixation (Thorgaard, 1977, 1983).

Sex-specific markers have been identified in several species, which facilitated identification of sex chromosomes by fluorescent *in situ* hybridization (FISH) using probes designed from those markers (Phillips, 2001; Phillips et al., 2001, 2002; Stein et al., 2001). When these markers were used as hybridization probes across species, they hybridized to one or more pairs of autosomes, strongly suggesting that sex-chromosomes differ among salmonid species (Phillips et al., 2001). One example is the male specific Oty1 marker in chinook salmon (Devlin et al., 1991, 1994, 1998). Two other closely linked male-specific

markers, GH-Y and OtY2, are shared across all Pacific salmon, except sockeye salmon and amago salmon (*Oncorhynchus rhodurus*) (Brunelli et al. 2004; Phillips et al., 2007). The GH-Y marker was recently used to identify the sex chromosome pair in coho salmon, and has also been shown to be closely linked with *SEX* in pink salmon (*Oncorhynchus gorbuscha*) and chum salmon (*Oncorhynchus keta*). However, it has been shown that the sex chromosome pairs in coho salmon (*Oncorhynchus kisutch*) and chinook salmon (*Oncorhynchus tshawytsha*) are not homologous. Cytogenetic and genetic mapping of rainbow trout microsatellite loci in these two species has shown that chinook and coho sex chromosome pairs correspond to rainbow trout linkage groups 7 and 23, respectively (Phillips et al., 2007).

Taken together, current data support the hypothesis that salmonids represent the earliest stages in sex chromosome differentiation (Phillips et al., 2001). Solidifying this hypothesis is the observation that YY male coho and chinook salmon are both viable and fertile, suggesting that the X and Y chromosomes still retain similar gene complements (Hunter et al., 1982; Chevassus, 1988; Onozato, 1989; Devlin et al., 2001).

1.5.2.3 Sex-linkage in Salmonids

Microsatellite linkage maps are now available for many economically important salmonid species such as rainbow trout (*O. mykiss*) (Sakamoto et al., 2000, Young at al., 1998; Guyomard et al., 2006), brown trout (*S. trutta*) (Gharbi, 2001; Gharbi et al., 2006), Atlantic salmon (*S. salar*) (Gilbey et

al., 2000; Moen et al., 2004; R. Danzmann, unpublished results; B. Hoyheim, unpublished results), and Arctic char (*S. alpinus*) (Woram et al., 2004). So far, linkage data indicate a lack of conservation among phenotypic sex-determining loci (hereafter noted as *SEX*) among salmonid species (Woram et al., 2003). One study conducted within the genus *Salvelinus* leant evidence to this when it found that sex-linked allozyme markers in Arctic char (*S. alpinus*) were not linked to *SEX* in either lake trout (*S. namaycush*) or brook trout (*S. fontinalis*) (May et al., 1989). A separate study of the genus *Oncorhynchus* showed that a growth hormone marker, known to be linked to *SEX* in coho salmon (*O. kisutch*), chinook salmon (*O. tshawytscha*) and Masu salmon (*O. masou*), is not linked in rainbow trout (*O. mykiss*) or amago salmon (*Oncorhynchus rhodurus*) (Du, et al., 1993; Forbes et al., 1994; Nakayama et al., 1999; Zhang et al., 2001). In studies of the genus *Salmo*, a minisatellite known to be in tight association with *SEX* in brown trout (*S. trutta*) mapped to an autosomal pair in Atlantic salmon (*S. salar*) (Taggart et al., 1995).

The sex linkage groups of Arctic char (*S. alpinus*), Atlantic salmon (*S. salar*), rainbow trout (*O. mykiss*) and brown trout (*S. trutta*) were compared in a recent study by Woram et al. (2003). It was observed that the position of the sex-determining locus was not conserved with respect to synteny among sex-linkage groups between species. The sex-linked markers that amplified across species were found to be in autosomal, homologous linkage groups in other species. It was concluded that *SEX* is likely located at the end of the sex linkage groups of Atlantic salmon, brown trout and Arctic char. Although distal ends of linkage

groups are not necessarily coincidental with the telomeric ends of chromosomes, large recombination distances between *SEX* and its associated markers is suggestive of terminal placement. In rainbow trout, *SEX* has been mapped close to putative centromeric markers, rather than the distal end of the linkage group, suggesting that the sex chromosome in rainbow trout was created through translocation of an ancestral sex chromosomal segment to an autosome (Thorgaard, 1977; Woram et al., 2003). A previous hypothesis that the sex-determining locus has been transposing within the genomes of some salmonids without relocating adjacent markers, and thus causing disruption of sex-linkage among species, supports the assumption that *SEX* is placed telomerically in Arctic char, Atlantic salmon and brown trout (Phillips et al., 2001).

1.6 Arctic Char (*Salvelinus alpinus*)

1.6.1 The Arctic char karyotype

As indicated previously, the extant salmonid ancestor (and thus the Arctic char ancestor) experienced a doubling of chromosomes roughly 25-100 million years ago, resulting in tetraploidy (Allendorf and Thorgaard, Hoeg et al., 2004). Previous cytogenetic studies have revealed diploid chromosome numbers (2n) between 78 and 84, with chromosome arm numbers (NF) of 98 or 100 (reviewed in Hartley et al., 1989). In those that have NF=98, the large acrocentric chromosome they possess appears to be the result of a tandem fusion (Phillips et al., 2002). *Salvelinus alpinus* has two more metacentrics not found in other

Salvelinus species, a difference likely due to differences in Robertsonian fusions (May, et al., 1989).

Although lake trout (also known as lake char, *Salvelinus namaycush*) have heteromorphic sex chromosomes that can be easily identified cytogenetically (Stein at al., 2002), no sex chromosomes analogous to the X and Y in humans can be observed by karyotype analysis of Arctic char. C bands are commonly located at all centromeres, as well as at the telomeres of many chromosomes, while Q bands are faint at some centromeres and prominent at some telomeres. Although the distribution of Q bands is similar to that obtained for C banding, there are far fewer Q bands. This suggests that subsets of heterochromatin exist, and that there may be an individual variation in DNA content (Hartley, 1989).

1.6.2 Sex-determination in Arctic char

To date, nothing is known about the molecular basis of sex-determination in Arctic char, nor has the sex chromosome pair been identified. However, as in other salmonids, it is believed that males are the heterogametic sex. Linkage analysis has identified several microsatellite and AFLP markers that are associated with *SEX* (Woram et al., 2004). Such markers are currently facilitating studies that seek to identify the sex-determining chromosome, and identify *SEX* in Arctic char.

1.7 Identification of *SEX*

1.7.1 Approaches previously used to identify and characterize *SEX*

The integration of genetic linkage maps and large-insert genomic libraries has greatly assisted the identification of *SEX* in other species (Beye et al., 2003; Matsuda et al., 2002). In the honeybee (*A. mellifera*), *SEX* was mapped between a marker obtained by bulk segregation analysis using multilocus fingerprinting and a RAPD based genetic marker. Probes were designed from the flanking regions of these markers, and were used to screen a 20 fold genomic coverage cosmid library and a 5 fold genomic coverage BAC (bacterial artificial chromosome) library. Probes were then designed from the end sequences of positive clones, and used to perform chromosome walking towards the sex-determining factor. When this technique was coupled with positional cloning of *SEX*, the region containing *SEX* was narrowed down to a 24 kb segment. This segment was shotgun sequenced to a 12-fold coverage, and subsequent sequence analysis eventually revealed that *csd* is the primary genetic signal of the sex-determining pathway in the *A. mellifera*.

DMY, the primary genetic signal in the sex-determining pathway of medakafish, was discovered in a similar manner (Matsuda et al., 2002; Matsuda, 2005). Matsuda and his colleagues isolated sex-linked DNA markers, made genetic and cytogenetic maps, and constructed a BAC genomic library (reviewed in Matsuda et al., 2005). Chromosome walking techniques were then used to identify a BAC clone containing the *SEX* locus.

1.8 Aim of the Thesis

The purpose of this thesis is to integrate the *SEX*-containing linkage group AC-4 with the Arctic char karyotype in order to characterize the sex-determining chromosome and identify the sex-determining region of the genome.

CHAPTER 2 MATERIALS AND METHODS

2.1 Fosmid Library Screening

2.1.1 Probe design from Arctic char-specific marker sequences

Prior to screening the fosmid library, it was necessary to design Arctic char-specific hybridization probes from the sex-linked markers that were derived from other salmonid species.

2.1.1.1 Primer design and testing viability on Arctic char DNA

Oligonucleotide primer pairs were manually designed to the flanking regions of 9 sex-linked microsatellite markers and one SCAR marker using the guidelines in "DNA Probes" (Keller, 1989). Since some primers were designed from sequences from salmonid species other than Arctic char, primer viability was tested on Arctic char genomic DNA. PCR amplifications were performed in a T3 Thermocycler (Biometra) in 28 µl containing 100 ng genomic DNA, 1X PCR reaction buffer (1.5 mM $MgCl_2$, Invitrogen), 50 nM dNTP (Roche), 0.5 U of Taq DNA polymerase (Invitrogen) and 0.5 µM forward and reverse PCR primers. PCR products were separated by electrophoresis on a 1.5% agarose gel containing 1X TBE and 0.5 µg/ml ethidium bromide. The DNA fragments were visualized using a UV trans-illuminator (Ultra-Violet Products). When necessary, PCR conditions were optimized by altering the annealing temperature (T_m).

2.1.1.2 Sequencing Arctic char-specific PCR products

To obtain Arctic char-specific marker sequences, all PCR products were first cloned into the pSTBlue-1 vector (AccepTor™ Vector Kit) in a 10 µl volume consisting of 50 ng sPTBlue-1 vector, 50 ng of DNA fragment, 5 µl Clonables™ 2X Ligation Premix, and dH_2O to top up to 10 µl. The ligation reaction was allowed to incubate at 16°C overnight (~16 hours). The ligation reactions were used to transform 50 µl of NovaBlue Singles Competent Cells (Novagen). The cells were first allowed to thaw on ice briefly. 1 µl of the ligation reaction was mixed with 50 µl of cells and then incubated for 5 minutes on ice. The cells were then heat-shocked for exactly 30 seconds at 42°C, and then incubated on ice for 2 minutes. 250 µl of room temperature SOC medium was added, and 20-100 µl of each sample was plated onto an LB agar plate inoculated with 200 µg/ml carbenicillin. For blue/white screening of recombinants, plates were pre-spread with X-gal and IPTG for a final concentration of 70 µg/ml and 80 µM per plate, respectively, and allowed to soak in for at least 30 minutes prior to plating cells. Plates were incubated overnight (~16 hours), or until colonies were sufficiently grown. After incubation, plates were incubated at 4°C to allow the blue coloration to intensify. Ten to fifteen white colonies were selected for each marker, and were grown in 5ml of LB medium inoculated with 25 µg/ml carbenicillin for 16-18 hours at 37°C on a shaker. DNA was isolated from 3ml of each culture using the QIAprep® Miniprep Kit (Quiagen). The volume of reagents was adjusted accordingly for 5ml overnight cultures.

PCR amplifications were performed on each clone in a T3 Thermocycler (Biometra) in 28 µl containing 2 µl (30 ng) clone DNA, 1X PCR reaction buffer (1.5 mM $MgCl_2$, Invitrogen), 50 nM dNTP (Roche), 0.5 U of Taq DNA polymerase (Invitrogen) and 0.5 µM of T7 and SP6 primers. Positive controls were also done using 50 ng Arctic char genomic DNA as a template. The PCR temperature profile consisted of an initial denaturation step of 94°C for 5 minutes; 34 cycles of 94°C for 30 seconds, 45 seconds at 50°C, and 72°C for 1 minute; and a final extension step of 72°C for 10 minutes. PCR products were separated by electrophoresis on a 1.5% agarose gel containing 1X TBE and 0.5 µg/ml ethidium bromide. The DNA fragments were visualized using a UV trans-illuminator (Ultra-Violet Products).

A representative clone for each different band size was then sequenced in a 10 µl reaction containing 4 µl Sequencing reagent premix (DYEnamic ET terminator Cycle Sequencing Kit), 0.5 µl of either T7 or SP6 primer (10 µM), 3 µl template DNA (200-500ng), and 2.5 µl dH_2O. Sequencing reaction were performed in a T3 Thermocycler (Biometra) with the following temperature profile: 25 cycles of 95°C for 20 seconds, 50°C for 15 seconds, and 60°C for 60 seconds. For post-reaction cleanup, 40 µl of 95% ethanol and 1 µl of sodium acetate/EDTA buffer (1.5M sodium acetate (pH>8.0) and 250 mM EDTA) were added to each sample, and mixed well with a vortex mixer. Samples were centrifuged in a microcentrifuge for 15 minutes at ~13,000 rpm. The supernatant was discarded, and DNA pellets were washed briefly with 70% ethanol. The supernatant was again discarded, DNA pellets were air-dried for ~1 hour, and

were resuspended in appropriate loading buffer. Sequencing reactions were run on either a 310 Genetic Analyzer (Applied Biosystems ABI Prism™) or a 377 DNA Sequencer (Perkin Elmer ANI Prism™). Sequences were viewed and edited using Sequencher 4.6 software. Hybridization probes were designed from these Arctic char sequences.

2.1.2 PCR-screening of fosmid mini-libraries

To screen the fosmid library for clones containing sex-linked markers using the designed primer pairs, PCR amplifications were performed in a T3 Thermocycler (Biometra) in 28 µl reactions containing 2 µl (30 ng) mini-library DNA, 1X PCR reaction buffer (1.5 mM $MgCl_2$, Invitrogen), 50 nM dNTP (Roche), 0.5 U of Taq DNA polymerase (Invitrogen) and 0.5 µM forward and reverse PCR primers. Positive controls were also done using 50 ng Arctic char genomic DNA as a template. The PCR temperature profile consisted of an initial denaturation step of 94°C for 5 minutes; 34 cycles of 94°C for 30 seconds, 45 seconds at 55°C, and 72°C for 1 minute; and a final extension step of 72°C for 10 minutes. PCR products were separated by electrophoresis on a 1.5% agarose gel containing 1X TBE and 0.5 µg/ml ethidium bromide. The DNA fragments were visualized using a UV trans-illuminator (Ultra-Violet Products).

2.1.3 Colony lifts

Mini-libraries containing the marker of interest were optimally titred so as to achieve a colony density of approximately 100-200 colonies per plate. The

optimal volume of minilibrary was then spread evenly onto each LB agar plate containing 20 µg/ml chloramphenicol. The plates were inverted and colonies were allowed to grow overnight (~16 hours) at 37°C, or until colonies had reached 0.5-1 mm diameter. Before performing colony lifts, the plates were allowed to cool at 4°C for at least 30 minutes. For each plate, a nylon membrane (Hybond-XL) the same size (diameter) as the plate was placed on the agar/colony surface and marked in several positions using a pin to ensure correct orientation in subsequent analyses. After 30 seconds, the membrane was removed from the plate using blunt-ended forceps grasping the outermost edge only. The DNA was liberated from the bacteria, denatured and fixed to the membrane following a neutralization step. This was achieved by placing the membranes colony uppermost on a series of solution-saturated 3MM paper pads: a) Lysis step, lysis buffer (10% (w/v) SDS) 1-3 minutes, b) Denaturation step, denaturation buffer (87.66% NaCl, 20% NaOH) 2-5 minutes, c) Neutralization step twice, neutralization buffer (87.66% NaCl, 60.5% Trizma base). Membranes were then washed vigorously in 2X SSC, and air-dried on dry 3MM paper pads. DNA was fixed to the membranes by baking at 80°C for 2 hours. Membranes were stored in airtight bags between sheets of 3MM paper until used.

2.1.4 Hybridization/Probing

Eight 40mer oligonucleotide probes were manually designed to the flanking regions of sex-linked markers using the guidelines/suggestions in the

reference book (Keller, 1989). For probing both colony-lift membranes and Southern membranes, the 40mers were labeled at the 5'end with ^{32}P at 37°C in 10µl reaction consisting of 2 µl oligonucleotide probe (10 µM), 4 µl (0.37 MBq) γ ^{32}P-ATP (Perkin Elmer), 10 U T4 polyoligonucleotide kinase (Invitrogen), 2 µl Invitrogen 5X Forward Reaction buffer (350 mM Tris-HCL (pH 7.6), 50mM $MgCl_2$, 500 mM KCL and 5 mM 2-mercaptoethanol) and 1 µl dH_2O. Southern membranes required an additional separate hybridization mixture for labeling the 1Kb ladder (Invitrogen), prepared in the same manner as outlined above. Each previously prepared membrane (Hybond-XL; see 2.1.3) was pre-hybridized at 55°C for two hours in 90 µl modified Church-Gilbert buffer (0.5 M phosphate buffer ($NaHPO_4$, pH 7.2), 7% SDS and 10 mM EDTA). To each hybridization tube containing six pre-hybridized colony-lift membranes, 5 µl radioactively labeled probe mixture was added, and membranes were allowed to hybridize at 55°C overnight (~16 hours). For hybridization tubes containing a single Southern membrane, the entire 10 µl of radioactively labeled probe mixture and the entire 10 µl of radioactively labeled 1Kb ladder mixture were added. Membranes were then washed as follows: a brief wash in low stringency buffer (2X SSC and 0.1% (w/v) SDS), two 5 minute washes in low stringency buffer, two 5 minute washes in medium stringency buffer (1X SSC, 0.1% (w/v) SDS), and two 5 minute washes in high stringency buffer (0.1% SSC, 0.1% (w/v) SDS). Membranes were then wrapped in plastic film and expose to either X-ray film (Kodak) for 48 hours, or storage phosphor screens (Molecular Dynamics) overnight (~16 hours).

Hybridization signals were detected by developing the X-ray film or using a STORM 820 Phosphoimager (Amersham Biosciences) for the phosphor-screens.

2.1.5 Confirmation of positive fosmids by PCR

Fosmid clones positive by hybridization were located on and picked from the original plates from which colony lifts were performed. Each was grown in 5ml of LB medium inoculated with 25 µg/ml chloramphenicol for 16-18 hours at 37°C on a shaker. DNA was isolated from 3ml of each culture using the QIAprep® Miniprep Kit (Quiagen). The volume of reagents was adjusted accordingly for 5ml overnight cultures. Each fosmid that was positive by hybridization for a particular marker's probe was tested for the presence of that marker by PCR using appropriate primers and annealing temperatures. PCR amplifications were performed in a T3 Thermocycler (Biometra) in 28 µl containing 20 ng fosmid DNA, 1X PCR reaction buffer (1.5 mM $MgCl_2$, Invitrogen), 50 nM dNTP (Roche), 0.5 U of Taq DNA polymerase (Invitrogen) and 0.5 µM forward and reverse PCR primers. Positive controls were also done using 50 ng Arctic char genomic DNA as a template. The PCR temperature profile consisted of an initial denaturation step of 94°C for 5 minutes; 34 cycles of 94°C for 30 seconds, 45 seconds at the appropriate T_m, and 72°C for 1 minute; and a final extension step of 72°C for 10 minutes. PCR products were separated by electrophoresis on a 1.5% agarose gel containing 1X TBE and 0.5 µg/ml ethidium bromide. The DNA fragments were visualized using a UV trans-illuminator (Ultra-Violet Products).

2.1.6 Characterization of positive clones by *Hind*III-digestion and Southern blotting

Those fosmids that were PCR-positive for the marker of interest were subjected to a *Hind*III digestion using the following recipe: 10 units of *Hind*III enzyme (Invitrogen), 2.5 μl React2 10X buffer (Invitrogen), 1 μg fosmid DNA template, and dH_2O to top up the volume to 23 μl. Clones were digested overnight (~16 hours). 20 μl from each of the *Hind*III--digested samples plus 6 μl 6X loading dye (Invitrogen) were run on a 0.8% TBE/agarose gel at 100V until the blue dye had migrated the full length of the gel. The resulting gels were stained for 30 minutes in a bath containing 0.5 μg/ml of ethidium bromide. The DNA fragments were visualized using a UV trans-illuminator (Ultra-Violet Products). DNA in the gel was denatured by soaking in a denaturation solution (0.5M NaOH and 1M NaCl) for two 20 minute intervals on a shaker. The gel was transferred to fresh solution after the first 20 minute interval. The denatured DNA was then Southern-transferred to a nylon membrane (Hybond-XL) cut to the same size of the gel. The Southern transfer was allowed to proceed overnight (~16 hours). The Southern membrane was rinsed briefly in a 2X SSC solution and allowed to air-dry. The DNA was fixed to the membrane by baking at 80°C for 2 hours. Membranes were stored between clean, dry sheets of 3MM paper in airtight plastic bags until they were probed (see 2.1.4).

2.2 Sequencing and Analysis of Clone Sal05SFU-6

2.2.1 Sal05SFU-6 shotgun library construction

The Sal05SFU-6 clone was prepared using the QIAprep® Miniprep Kit (Quiagen). Approximately 15 µg of isolated fosmid DNA was sheared by sonication. Sonication efficiency was verified by agarose gel electrophoresis. Sonicated DNA was end-repaired using the Epicentre End-It DNA-Repair Kit. The end-repaired reactions were carried out at room temperature in a 50 µl final volume containing 5 µl each of dNTP mix, ATP, and 10X buffer from the kit, as well as 1 µl of End-Repair Enzyme Mix and 34 µl of DNA, and were incubated for 45 minutes.

Sonicated, end-repaired fosmid DNA was size fractionated by gel electrophoresis using a 1% agarose gel made with 1X TAE. Upon completion of electrophoresis, the lane containing the 1Kb ladder (Invitrogen) was cut off from the rest of the gel and stained in a 0.5 µg/ml ethidium bromide bath for 45 minutes. The ladder was visualized under UV light (Ultra-Violet Products), and notches were made in the gel at the 2 and 5 Kb standard sizes. The remainder of the gel was then compared to the notched ladder portion, and DNA between 2 and 5Kb was cut out from the gel and subsequently extracted using the Quiagen Gel Extraction Kit. DNA concentration was then measured by spectrophotometry and the sample was diluted down to 25 ng/µl.

The 2 to 5 Kb size fractionated portion of the DNA was then ligated into a *Sma*I-digested, phosphatase-treated pUC19 vector. Positive (standard insert

DNA) as well as negative (no insert) controls were also performed. The ligation reactions were carried out using T4 DNA Ligase (Invitrogen) in 20 µl reaction volumes containing 10U of ligase, 100 ng of insert DNA, 20 ng of *Sma*I-digested, phosphatase-treated pUC19 vector, and 4 µl of 5X reaction buffer. The sample was incubated overnight (~16 hours) at 14°C.

The ligation reaction was used to transform 50 µl of NovaBlue Singles Competent Cells (Novagen). The cells were first allowed to thaw on ice briefly. 1 µl of the ligation reaction was mixed with 50 µl of cells and then incubated for 5 minutes on ice. The cells were then heat-shocked for exactly 30 seconds at 42°C, and then incubated on ice for 2 minutes. 250 µl of room temperature SOC medium was added, and 20-100 µl of each sample was plated onto an LB agar plate inoculated with 200 µg/ml carbenicillin. For blue/white screening of recombinants, plates were pre-spread with X-gal and IPTG for a final concentration of 70 µg/ml and 80 µM per plate, respectively, and allowed to soak in for at least 30 minutes prior to plating cells. Plates were incubated overnight (~16 hours), or until colonies were sufficiently grown. After incubation, plates were incubated at 4°C to allow the blue coloration to intensify.

2.2.2 Sequencing and analysis of Sal05SFU-6

Sal05SFU-6 shotgun clone DNA was prepared using the Sambrook alkali method (Sambrook et al., 1989). All shotgun clone DNA samples were subjected to a *Pvu*II-digest in a 25 µl reaction containing 300-500 ng DNA, 10U *Pvu*II enzyme (Invitrogen), 2.5 µl React 6X Buffer (Invitrogen) and dH_2O to top the

volume up to 25 µl. Shotgun clones were digested overnight (~16 hours) at 37°C. Before sequencing the shotgun clones, it was necessary to determine how much DNA should be used in the sequencing reactions based on the sizes of the shotgun clones. This was done by running 20 µl of each *Pvu*II -digested shotgun clone on a 0.8% TBE/agarose gel to separate the vector fragment from the cloned insert fragment. PCR products were separated by electrophoresis on a 1.5% agarose gel containing 1X TBE and 0.5 µg/ml ethidium bromide, along with a 1Kb standard ladder (Invitrogen). The DNA fragments were visualized using a UV trans-illuminator (Ultra-Violet Products). Once the size of the shotgun clone fragment was determined, it was added to the size of the vector, and the total size was used to determine how many ng of DNA template should be used in subsequent sequencing reactions. This was determined using the charts found in the DYEnamic ET terminator Cycle Sequencing Kit handbook.

Shotgun sequencing reactions were performed in a T3 Thermocycler (Biometra) in 10 µl containing 4 µl Sequencing reagent premix (DYEnamic ET terminator Cycle Sequencing Kit), 0.5 µl of either M13F or M13R primer (10 µM), the appropriate amount of template DNA, and dH_2O if necessary to top up to 10 µl. The sequencing temperature profile was as follows: 25 cycles of 95°C for 20 seconds, 50°C for 15 seconds, and 60°C for 60 seconds. For post-reaction cleanup, 40 µl of 95% ethanol and 1 µl of sodium acetate/EDTA buffer (1.5M sodium acetate (pH>8.0) and 250 mM EDTA) was added to each sample, and mixed well with a vortex mixer. Samples were centrifuged in a microcentrifuge for 15 minutes at ~13,000 rpm. The supernatant was discarded, and DNA pellets

were washed briefly with 70% ethanol. The supernatant was again discarded, DNA pellets were air-dried for ~1 hour, and pellets were resuspended in appropriate loading buffer. Sequencing reactions were run on either a 310 Genetic Analyzer (Applied Biosystems ABI Prism™) or a 377 DNA Sequencer (Perkin Elmer ANI Prism™). Sequences were viewed and edited using Sequencher 4.6 software.

2.2.3 Extension of Sal05SFU-6 contig sequences by selected primer-walking

20-mer oligonucleotide primers were manually designed approximately 100-200 nucleotides in from the ends of end-joined contigs such that they would extend the sequence outward. The primers were used to sequence from clones from which the primers were designed (i.e., using the appropriate clone DNA as template), and sequence reactions were performed in the same manner as during shotgun sequencing. Sequencing reactions were run on either a 310 Genetic Analyzer (Applied Biosystems ABI Prism™) or a 377 DNA Sequencer (Perkin Elmer ANI Prism™). Sequences were viewed and edited using Sequencher 4.6 software.

2.2.4 Fosmid end-sequencing

Sal05SFU-6 end-sequencing reactions were performed in a T3 Thermocycler (Biometra) in 10 µl containing 4 µl Sequencing reagent premix (DYEnamic ET terminator Cycle Sequencing Kit), 0.5 µl of either pCC1™/pEpiFOS™ Forward or Reverse sequencing primers (10 µM), 500 ng

Sal05SFU-6 template DNA, and dH_2O if necessary to top up to 10 µl. The sequencing temperature profile was as follows: 35 cycles of 95°C for 20 seconds, 50°C for 15 seconds, and 60°C for 60 seconds. For post-reaction cleanup, 40 µl of 95% ethanol and 1 µl of sodium acetate/EDTA buffer (1.5M sodium acetate (pH>8.0) and 250 mM EDTA) was added to each sample, and mixed well with a vortex mixer. Samples were centrifuged in a microcentrifuge for 15 minutes at ~13,000 rpm. The supernatant was discarded, and DNA pellets were washed briefly with 70% ethanol. The supernatant was again discarded, DNA pellets were air-dried for ~1 hour, and pellets were resuspended in appropriate loading buffer. Sequencing reactions were run on either a 310 Genetic Analyzer (Applied Biosystems ABI Prism™) or a 377 DNA Sequencer (Perkin Elmer ANI Prism™). Sequences were viewed and edited using Sequencher 4.6 software.

2.3 Fluorescent *in situ* Hybridization (FISH)

Fosmid DNA was induced to high copy number by growing 500 µl of overnight culture in 5 ml of LB medium inoculated with 25 µg/ml chloramphenicol and 4 µl *CopyControl*™ Fosmid Induction Solution (Epicentre) for 5 hours at 37°C on a shaker. DNA was isolated from 5 ml of each induced culture using the QIAprep® Miniprep Kit (Quiagen). The volume of reagents was adjusted accordingly for 5ml overnight cultures. The fosmid DNA was quantified by spectrophotometry, and approximately 10 µg of each clone was sent by courier to Ruth Phillips at Washington State University in Vancouver, Washington, where the FISH analysis was performed (Phillips et al., 2002).

CHAPTER 3 RESULTS

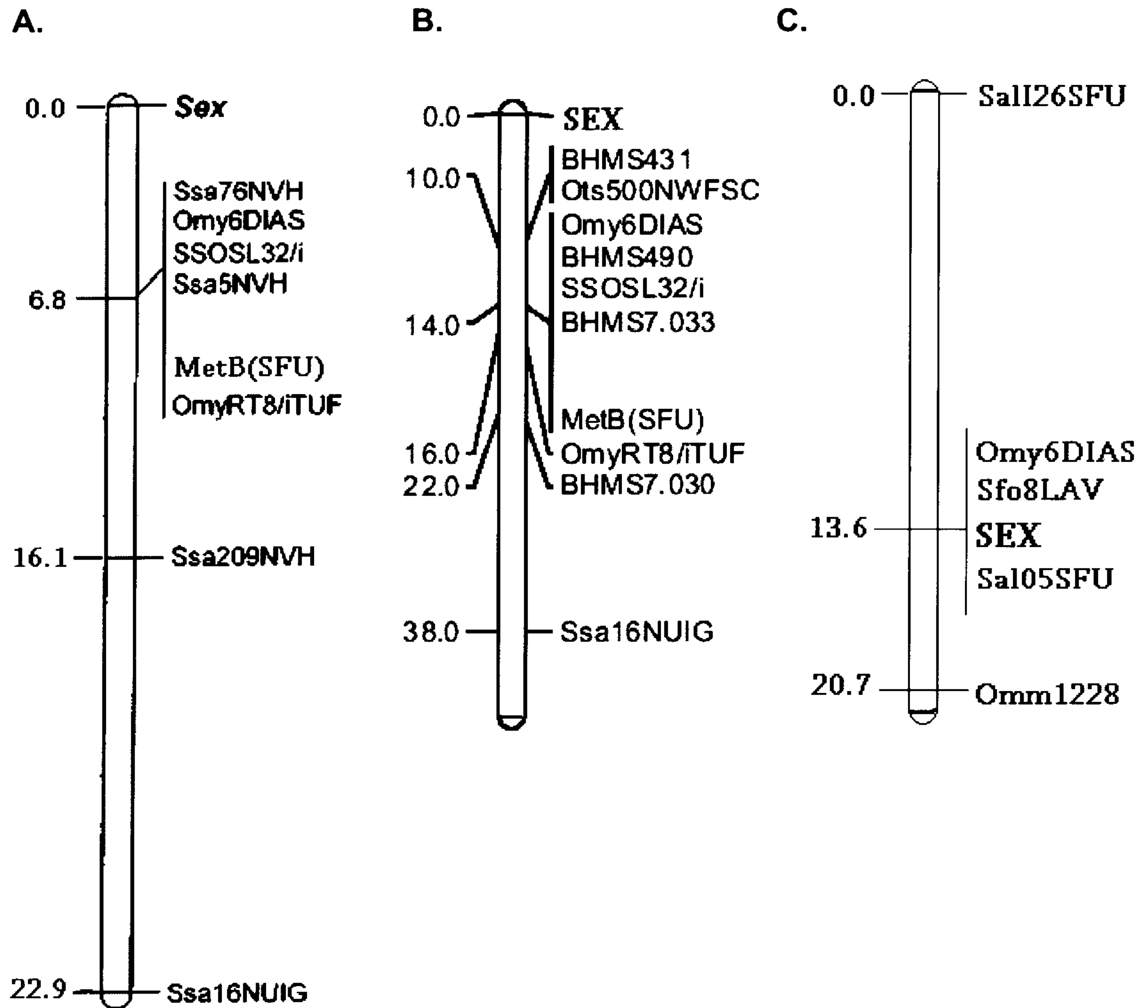
3.1 Resources Available at the Start of the Project

3.1.1 Genetic linkage maps

To identify *SEX* in Arctic char, I took as my starting point consolidated genetic linkage maps that had been generated in the Davidson lab at Simon Fraser University (SFU), and by Roy Danzmann's group at the University of Guelph (Woram et al., 2003; Woram et al., 2004) (Figure 3.1). The map generated by the Davidson lab is based on crosses involving two strains of Arctic char maintained at the Icy Waters facility in Whitehorse, Yukon: Nauyuk Lake (Northwest Territories) and Tree River (Northwest Territories) (McCarthy, 2004).The University of Guelph genetic maps, with a total of 46 linkage groups, were developed using crosses between two strains of Arctic char: Fraser River (Labrador) and Nauyuk Lake. Arctic char linkage group 4 (AC-4) has been designated the linkage group containing *SEX*, based on its association with the male phenotype (Woram et al., 2004).

Also available was the marker Sal05SFU, which had recently been developed into a SCAR (sequence characterized amplified region) from the AFLP marker CAC/ATC-465 (Ou, 2005). Sal05SFU has 0% recombination with the male phenotype in the Arctic char broodstock at the Icy Waters facility in Whitehorse, Yukon.

Figure 3.1 **Arctic char Linkage Group 4 (AC-4).** A. AC-4 from Woram et al. (2003) B. AC-4 from Woram et al. (2004) C. AC-4 developed in the Davidson lab for a Tree River male. SEX represents phenotypic sex. Estimates of map distances are indicated in centiMorgans.



3.1.2 Fosmid genomic library

Also prior to the start of this project, a fosmid library, characterized by genomic insert sizes of approximately 40Kb, was constructed from DNA from a male Tree River strain of Arctic char (Mosher et al., 2005). The library was partitioned into several smaller mini-libraries of approximately 5,000 clones each, facilitating rapid screening for a variety of markers positioned on the Arctic char linkage maps. Although the fosmid library has insert sizes smaller than a BAC library, fosmid clones are easier to handle (Ohtsuka et al., 2002). The library had been previously screened for the sex-linked markers Omy6DIAS and Sfo8LAV, and clones positive for these markers were identified, confirming the success of the library construction as well as its usefulness as a resource for this project.

3.1.3 Approach to identifying *SEX* in Arctic char

The approach used to identify the sex chromosome in Arctic char in this project was to screen the Arctic char fosmid library for clones containing sex-linked markers that appear on the Arctic char genetic linkage maps. Once isolated, fosmid clones positive for the marker of interest could be used as probes in FISH analysis to identify the sex chromosome. In addition, sequencing the fosmids containing markers that exhibited 0% recombination with *SEX* might identify candidate sex-determining genes.

3.2 Designing of Marker-Specific PCR Primers/Hybridization Probes

3.2.1 Primer design

Oligonucleotide primers were manually designed to the flanking regions of seven sex-linked microsatellite markers and one SCAR marker such that each pair consisted of a forward primer of approximately 20 nucleotides, and a complementary reverse primer also approximately 20 nucleotides in length. Source sequences used for designing the remaining primer pairs were derived from the following members of Salmonidae: Atlantic salmon (*Salmo salar*), brook trout (*Salvelinus fontinalis*), and rainbow trout (*Onchorhynchus mykiss*) (except for SalI26SFU and Sal05SFU, which were Arctic char-specific sequences already). The primer sequences are shown in Table 3.1.

3.2.2 Testing primers on Arctic char genomic DNA

Each primer pair was tested in PCR reactions using DNA extracted from a Nauyuk Lake strain of Arctic char as a template. Amplification products were successfully produced from Arctic char DNA for all designed primer pairs. The PCR annealing temperature (T_m) was optimized for each primer pair. A summary of the results of these PCR reactions is shown in Table 3.1.

Table 3.1 Primers designed from sex-linked markers used to screen fosmid library. (F-) Forward primer, (R-) Reverse primer. "Multiple"= 4 or more bands.

Marker Name	Primer Sequences (5'-3')	GenBank accession	Number of bands amplified	Approximate size of product(s) (bp)	PCR T_m (°C)
SsaBHMS431	F- GCAGTCAGTGGGTTTTAG R- TTGGCTTCCATTCGCAGC	AF256720	1	130	55.0
SsaBHMS490	F- TCAAATCTGTGTGACTGC R- ATGGAACAAATCACCCTC	AF256724	2	110, 410	55.0
SsaBHMS7.030	F- ATGGTGAGTAGATCCCCC R- ATGACTCCAGTTCCACCC	AF256838	Multiple	Variable	55.0
SsaBHMS7.033	F- AGACACGAGGACAGAAAC R- AGCTGTGACTCTCAATGTC	AF256658	1	150	55.0
SSOSL32	F- TTATGTCAGACCAGGTGGCTA R- CTGTATACTATGGTGGTGGCT	Z69642	Multiple	Variable	53.0
Sall26SFU	F- ACATACTCTGCCTTATTTCATACG R- CATTGGTGATCCTTTCTTCAG	Unpublished SFU [1]	1	290	55.0
Omy6DIAS	F- CCACCAACTTCTTACATGAT R- CTATGGGGACAGCCGAATAA	AF239042	3	150, 180, 290	55.0
Omm1228	F- CCCTTCCTGTGTGTCGTTGTT R- CAGGAGTCACTTGGCAGTAGGAG	AF470009	1	270	58.0
Sfo8LAV	F- CAACGAGCACAGAACAGG R- CTTCCCCTGGAGAGGAAA	U50305	1	280	52.0
Sal05SFU	F- TTATTATTATTGTTGTTGTTGTATG R- TGCCGCAGAAAAATCTAAC	Unpublished SFU [2]	1	320	52.0

1-Colin McGowan and Evelyn Davidson, Davidson Lab, MBB
2-George Ou, Davidson Lab, MBB

3.2.3 Sequencing amplicons to get Arctic char-specific sequences

The PCR products from markers that amplified a single band from Arctic char genomic DNA were sequenced directly using either the forward or the reverse PCR primer as a sequencing primer. For those that amplified two or more bands, it was necessary to sequence each band to identify which, if any, contained the marker. This was done by first cloning the multi-banded PCR products into the pSTBlue-1 vector. Ten to fifteen clones were selected for each marker and subjected to PCR using the T7 and SP6 primers. A representative clone for each different band size was then sequenced using either the T7 or SP6 primers. Arctic char-specific sequences were obtained for the markers SsaBHMS 431, SsaBHMS490, SsaBHMS7.033, SSOSL32, Sall26SFU, Omm1228, and Sal05SFU. A sequence containing the microsatellite marker for BHMS7.033 was not obtained, even though all different sizes of amplicons were cloned and sequenced. The Arctic char-specific marker sequences are shown in Figure 3.2.

Figure 3.2 **Arctic char-specific marker sequences.** Where possible, alignment with source-species sequence (on bottom) is shown; in these cases, the Arctic char-specific sequence appears on top of the pair-wise-aligned sequences. The PCR primers are highlighted in green. Regions where probes were designed are highlighted in yellow. The numbering refers to the sequence in GenBank for the marker.

SsaBHMS431

```
               ....|....| ....|....| ....|....| ....|....| ....|....|
                       10         20         30         40         50
A.charBHMS     ---------- ---------- ---------- ---------G ATTGCAGTCA
S.salarBHM     CCCAGTGACT CAGCAGTCAG TGGGTTTTAG AGTGNGTGTG TGTGT-GTGT

               ....|....| ....|....| ....|....| ....|....| ....|....|
                       60         70         80         90        100
A.charBHMS     GTGGGTTT-- TAGAGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGGCTAC
S.salarBHM     GTGTGTTTGT TTGTGTGTGT GTGTTTGTGT GTGTGTGTGT GTGTGGCTAC

               ....|....| ....|....| ....|....| ....|....| ....|....|
                      110        120        130        140        150
A.charBHMS     TGACTCAGTC ATAAGTAGGT TTTAGAGAGA GCCACACCCC AAAACGGGGT
S.salarBHM     TGACTCAGTC GTAAGTAGGT TTTGGAGAGA GCCGCACCCC AAAGCAGGGT

               ....|....| ....|....| ....|....| ....|....| ....|....|
                      160        170        180        190        200
A.charBHMS     CCAGCTGCGA ATGGAAGCC- ---------- ---------- ----------
S.salarBHM     CCAGCTGCGA ATGGAAGCCA ACAATGAGTT TTCGGTGGAC AGAGAGGAAA
```

SsaBHMS490

```
              ....|....| ....|....| ....|....| ....|....| ....|....|
                     310        320        330        340        350
A.charBHMS    ---------- ---------- ---------- ---------- --TCAAATCT
S.salarBHM    CCTGGCTGTG GCGACAGAAG GTCTCTGGCT GTGCCTGAAA ACTCAAATCT

              ....|....| ....|....| ....|....| ....|....| ....|....|
                     360        370        380        390        400
A.charBHMS    GTGTGACTGC AGGCTTTGTT CTGAGCATGA GCCTGAGA-- ----------
S.salarBHM    GTGTGACTGC AGGCTTTGTT CTGAGCATGA GCCTGAGTGT GTGTGTGTGT

              ....|....| ....|....| ....|....| ....|....| ....|....|
                     410        420        430        440        450
A.charBHMS    ---------T GTGTGTGTCT CGGGGGTTGT GTGTTTGAGA GGGGAAATGT
S.salarBHM    GTGTGTGTGT GTGTGTGTGT GGGGGGTTGT GTGTT-GAG- GGGGAAATGT

              ....|....| ....|....| ....|....| ....|....| ....|....|
                     460        470        480        490        500
A.charBHMS    CAAAAGAGGG TGATTTGTTC CAT------- ---------- ----------
S.salarBHM    CAAA-GAGGG TGATTTGTTC CATCATCCAG CATTCCAATA CCGCGTCCCT
```

SsaBHMS7.033

```
            ....|....| ....|....| ....|....| ....|....| ....|....|
                   10         20         30         40         50
A.charBHMS  ---------- ---------- ---------- ---------- ----------
S.salarBHM  CAAATGACAT ATTAAATGTT TCAATACCAG GGACGCTTTT AGAGGTGTGT

            ....|....| ....|....| ....|....| ....|....| ....|....|
                   60         70         80         90        100
A.charBHMS  ---------- --------AG ACACGAGGAC AGAAACACAC CACACACACA
S.salarBHM  GCTCCTCCTG TGATACACAG ACACGAGGAC AGAAACACAC CACACACACA

            ....|....| ....|....| ....|....| ....|....| ....|....|
                  110        120        130        140        150
A.charBHMS  CACACACACA CACACACACA CAC------- ---------- -TGAACAGTG
S.salarBHM  CACACACACA CACACACACA CACACACACA CACACACACA CTGAACAGTG

            ....|....| ....|....| ....|....| ....|....| ....|....|
                  160        170        180        190        200
A.charBHMS  GAGCGCTTTC ATCTCCACAG AGACATTCCG TCGCGCAGAC ATTGAGAGTC
S.salarBHM  GTGCGCTTTC ATCTCCACAG AGACATTCCG TCGCGCAGAC ATTGAGAGTC

            ....|....| ....|....| ....|....| ....|....| ....|....|
                  210        220        230        240        250
A.charBHMS  ACA------- ---------- ---------- ---------- ----------
S.salarBHM  ACAGCTACAC TAACACAGAC ACTCAGAGAC GTTGCACACA GACATGCAAG
```

SSOSL32

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                        10         20         30         40         50
A.charSSOS      TTATGTCAGA CCAGGTGGGT ATCTCTCACA GGTGAGCCCA CAGACAGAGC
S.salarSSO      TTATGTCAGA CCAGGTGGCT ATCTCTCACA G-TGAGCCCA CAG-C-GCGC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                        60         70         80         90        100
A.charSSOS      ACACACACAC ACACACACAC ACAC~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
S.salarSSO      ACACACACAC ACACACACAC ACACACACAC ACACACACAC ACAC-ACACA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       110        120        130        140        150
A.charSSOS      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~CAAGC
S.salarSSO      CACACACACA CCACACACAC ACCACACAAC ACACACACAC ACACACAAGC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       160        170        180        190        200
A.charSSOS      CACCACCATA GTATACAGAA TCT------- ---------- ----------
S.salarSSO      CACCACCATA GTATACAG.. .......... .......... ..........
```

SalI26SFU

```
             ....|....| ....|....| ....|....| ....|....| ....|....|
                    10         20         30         40         50
SalI26SFU    ACATACTCTG CCTTATTTCA TACGCTGGGT GTCTGTTGTG CTTTGTGTTG

             ....|....| ....|....| ....|....| ....|....| ....|....|
                    60         70         80         90        100
SalI26SFU    CTAATGTCCC TTTAACTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG

             ....|....| ....|....| ....|....| ....|....| ....|....|
                   110        120        130        140        150
SalI26SFU    TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG

             ....|....| ....|....| ....|....| ....|....| ....|....|
                   160        170        180        190        200
SalI26SFU    TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TATGTGTGTG

             ....|....| ....|....| ....|....| ....|....| ....|....|
                   210        220        230        240        250
SalI26SFU    TGTGTCCAGN TNAATGTCGG TGGAAGAGGA GCTGAAGAAA GGATCACCAA

             ..
SalI26SFU    TG
```

Omm1228

```
               ....|....| ....|....| ....|....| ....|....| ....|....|
                      110        120        130        140        150
A.charOmm1     ---------- ---------- --GAGTCCCT TCCTGTGTGT CGTTGTTTAA
O.mykissOm     GCTGTAGTGA CTCGAGGTGC AGGCTGCCCT TCCTGTGTGT CGTTGTTTAA

               ....|....| ....|....| ....|....| ....|....| ....|....|
                      160        170        180        190        200
A.charOmm1     TGTAATCGTT GTGGTTT-AC AGCCTTGATC GTGCTGGGGC TAACACACAC
O.mykissOm     TGTAATCGTT GTGGTTTTAC AGCCTTG-TA GTGCTGGGGC TAACACACAC

               ....|....| ....|....| ....|....| ....|....| ....|....|
                      210        220        230        240        250
A.charOmm1     ACACACACAC ACACACACAC ACACACACAC ACACACGAAA CCAACTGCTG
O.mykissOm     ACACACACAC ACACACACAC ACAC------ ---CA-GAAA CCAACTGCTG

               ....|....| ....|....| ....|....| ....|....| ....|....|
                      260        270        280        290        300
A.charOmm1     CTATTTCTGT GACGCTTGTT GTCATGTGCC TTGCTGGAGT GAGAGGTTGT
O.mykissOm     CTATTTCTGT GTCGCTTGTC GT---GTGCC TTGCTGGAGT GAGAAGTGAT

               ....|....| ....|....| ....|....| ....|....| ....|....|
                      310        320        330        340        350
A.charOmm1     GCATCTAACC AGTCTGAGGC GACACAGCCC TGTGACCTAA CC--TA--AG
O.mykissOm     GCATC-ATGC --TCTG---- ----CTTG-- TGTGATCTAA CCAGTCCGAG

               ....|....| ....|....| ....|....| ....|....| ....|....|
                      360        370        380        390        400
A.charOmm1     GCCGTCTCCT ACTGCCAAGT GACTCCTGAA TCAC------ ----------
O.mykissOm     GCCGTCTCCT ACTGCCAAGT GACTCCTGAG GGAAGGGAGC ATCAGTTACG
```

Sal05SFU

```
              ....|....| ....|....| ....|....| ....|....| ....|....|
                     10         20         30         40         50
Sal05SFU      AGATTGACTG CGTACCAATT CACTAACATT GAGAATCATG TCACTCATCT

              ....|....| ....|....| ....|....| ....|....| ....|....|
                     60         70         80         90        100
Sal05SFU      AACTGACTAC AGAGTATTGC CTAAAAGCAT GCACTTTATT ATTATTGTTG

              ....|....| ....|....| ....|....| ....|....| ....|....|
                    110        120        130        140        150
Sal05SFU      TTGTTGTATG AAGAGATTAG CTCAGATCAG TTCTGTGCTT ATACCATGGC

              ....|....| ....|....| ....|....| ....|....| ....|....|
                    160        170        180        190        200
Sal05SFU      ATTGTTGAAT ACTCTGTTCT GATTGGCTAT AAGGGCAGTC TGGTTTGTAT

              ....|....| ....|....| ....|....| ....|....| ....|....|
                    210        220        230        240        250
Sal05SFU      TATTTACCTA TAATCAGTCC CTCTAGGATT CCGCTGACAT TTTTTGTGAT

              ....|....| ....|....| ....|....| ....|....| ....|....|
                    260        270        280        290        300
Sal05SFU      AATCTTTTGG CAAAAATGTT AGATTTTTCT GCGGCAATTG ACATTTTGCA

              ....|....| ....|....| ....|....| ....|....| ....|....|
                    310        320        330        340        350
Sal05SFU      TGGCAAGGCA GAGGGGTGTT TGGGTTAGCT CTAACGAATT GTTAGAATGT

              ....|....| ....|....| ....|....| ....|....| ....|....|
                    360        370        380        390        400
Sal05SFU      TGCAGTCAGA ATGACTGCTC ATTAGCTTGA AGGGGGCTGC ATCGAGGCCC

              ....|....| ....|....| ....|....| ....|
                    410        420        430
Sal05SFU      AGGGTTTCTA GTGTTACTCA GGACTCATCA ATCAC
```

3.2.4 Designing hybridization probes from Arctic char-specific sequences

Hybridization probes of approximately 40 nucleotides in length were designed from the Arctic char-specific sequences of six sex-linked microsatellite markers and one sex-linked SCAR marker: SsaBHMS431, SsaBHMS490, SsaBHMS7.033, SSOSL32, Sall26SFU, Omm1228, and Sal05SFU, respectively. These were designed in a region between the two PCR primers, or overlapping one of the primers, but close to the marker of interest. A summary of the probe sequences and hybridization conditions is shown in Table 3.2.

Table 3.2 Hybridization probes designed from Arctic char-specific marker sequences. Optimal annealing temperatures (T_m) for each probe are indicated.

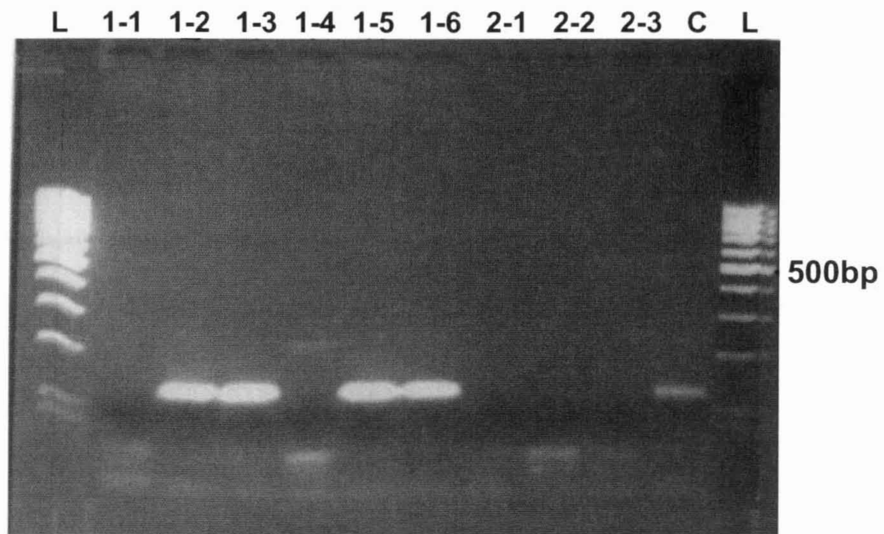
Marker Name	Hybridization Probe Sequences (5'-3')	T_m (°C)
SsaBHMS431	TACTGACTCAGTCATAAGTAGGTTTTAGAGAGAGCCACAC	65.5
SsaBHMS490	GGGTTGTGTGTTTGAGAGGGGAAATGTCAAAAGAGGGTGA	67.6
SsaBHMS7.033	CAGTGGAGCGCTTTCATCTCCACAGAGACATTCCGTCGCG	71.7
SSOSL32	TTATGTCAGACCAGGTGGGTATCTCTCACAGGTGAGC	67.8
Sall26SFU	GGTGTCTGTTGTGCTTTGTGTTGCTAATGTCCCTTTAACTG	66.5
Omm1228	GACGCTTGTTGTCATGTGCCTTGCTGGAGTGAGAGGTTGT	69.6
Sal05SFU	TTTACCTATAATCAGTCCCTCTAGGATTCCGCTGACATTT	64.5

3.3 Fosmid Library Screening

3.3.1 PCR amplification from mini-libraries

The primer pairs that were designed for the sex-linked markers were used to detect the presence of the corresponding sequences in the Arctic char fosmid library. DNA prepared from nine mini-libraries was used as a template in PCR reactions, and Arctic char DNA was used as a positive control template to verify the marker identity. The fosmid library was PCR-screened for SsaBHMS431, SsaBHMS490, Omm1228, Sall26, and Sal05SFU. For each of these markers, I was able to identify one or more mini-libraries that had clones that contained the marker (i.e., were PCR-positive). A sample image showing a mini-library PCR-screen for Sal05SFU is shown in Figure 3.3.

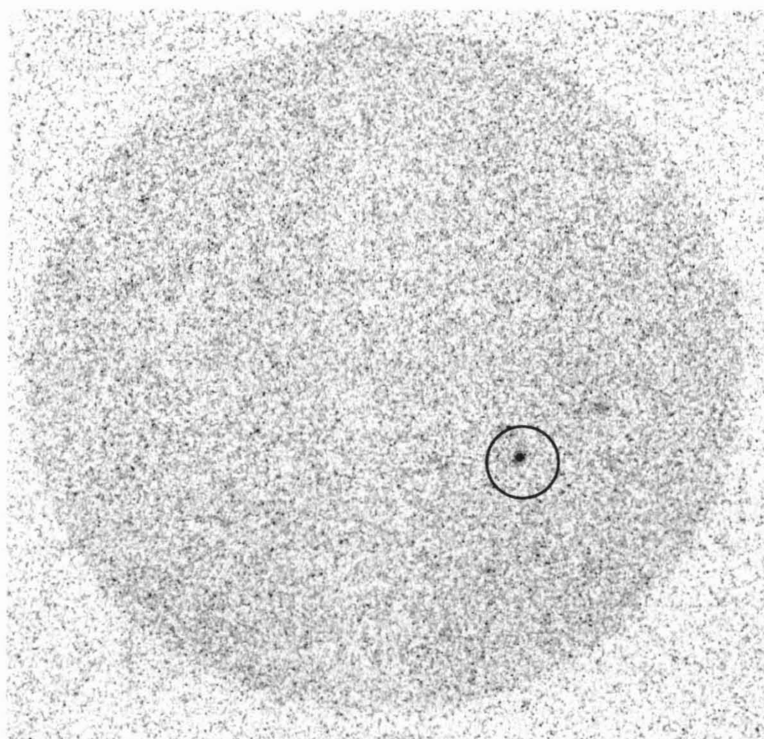
Figure 3.3 **1.5% agarose gel showing amplification products from a Sal05SFU mini-library PCR-screen.** Lanes labeled "L" contain a 100bp ladder, in which the brightest band is 500 bp. The lane labeled "C" is the positive control. Lanes are labeled at the top with the mini-library that the product was or was not amplified from. Sal05SFU is present in mini-libraries 1-2, 1-3, 1-5, and 1-6.



3.3.2 Colony lifts and hybridization using 40-mer probes

When a mini-library was known to contain a marker of interest, single clones containing the marker had to be isolated. For each marker that was confirmed present in a mini-library, samples of that fosmid mini-library were plated so that colonies could grow singly, at a density of approximately 100-200 colonies per plate. DNA was extracted from the fosmid colonies and fixed to membranes by colony lifts. Each membrane was hybridized overnight with the appropriate 40-mer probe, and exposed to either X-ray film or storage phosphor-screens. When hybridization signals were detected on the membranes, the fosmid colonies corresponding to the signals were located on the original plates, and isolated. Hybridization signals were detected for SsaBHMS431, SsaBHMS490, Sal05SFU, Omm1228, and Sall26. Thus, fosmid clones potentially positive for these markers were identified, and isolated. An example of a hybridization signal indicating a fosmid containing the marker of interest is shown in Figure 3.4.

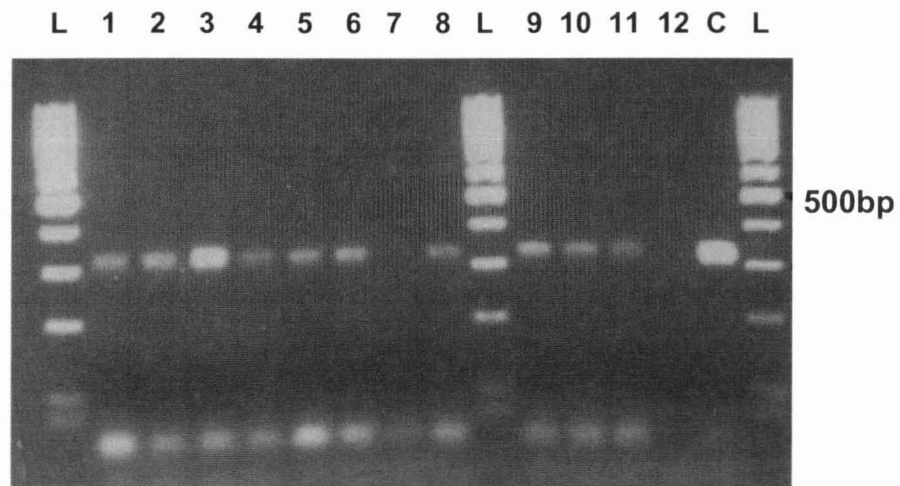
Figure 3.4 A phosphor image of a membrane hybridized with the Sal05SFU 40-mer probe. A hybridization signal is indicated by black circle, indicating a fosmid clone containing the marker Sal05SFU.



3.3.3 PCR confirmation of positive fosmid clones

Clones that were positive for markers of interest by hybridization needed to be confirmed as true positives. This was done by isolating the fosmid DNA from each clone, and using it as a template to amplify the marker using the appropriate primers (the primers originally designed and used to obtain Arctic char-specific sequences). If no amplification product was obtained from any of the clones at this point, the clones were deemed to be false-positives, and the mini-library was re-screened for that particular marker. Although clones were identified as positive by hybridization for SsaBHMS490, none of them was ever confirmed by PCR, and so clones containing this marker were never isolated. Clones were confirmed as PCR positive for Sal05SFU (Figure 3.5), Omm1228, and Sall26SFU.

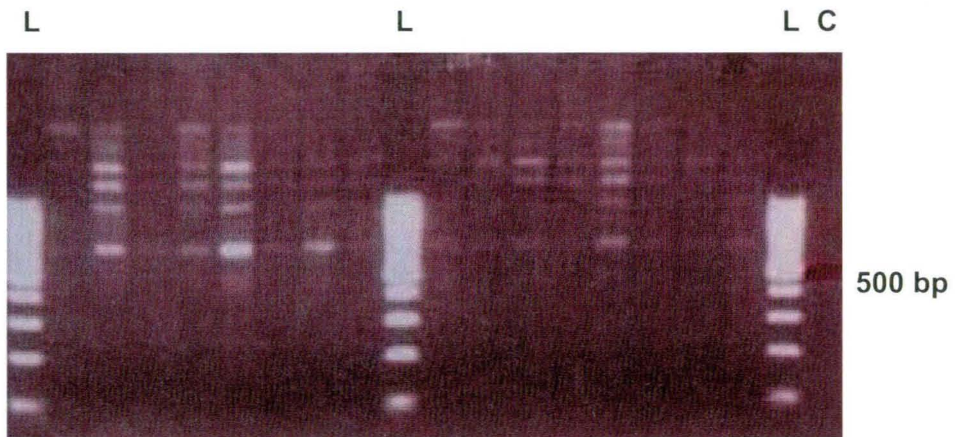
Figure 3.5 **1.5% agarose gel showing PCR products of Sal05SFU clones.** Clones that are PCR positive for Sal05SFU appear in lanes 1-6, and 8-11, corresponding to the clone numbers. Lanes labeled “L” contain a 100bp ladder, in which the brightest band is 500 bp. The lane containing the positive control is labeled “C”.



3.3.3.1 PCR confirmation of SsaBHMS431

Screening the fosmid library resulted in the identification many clones that potentially contained the BHMS431 microsatellite. When the clones were subjected to PCR-confirmation, the gel consistently displayed ladder-like products. When PCR was attempted on the clones using only the forward primer, there was no amplification. However, when PCR was attempted on the clones using only the reverse primer, the gel displayed the same ladder-like banding pattern of amplification as with both PCR primers (Figure 3.6), indicative of SsaBHMS431 placement in a highly repetitive region. The entire screening process was repeated several times for BHMS431, and although many clones were positive by hybridization, they always amplified ladder-like bands. Since PCR attempts on the hybridization-positive BHMS431 clones consistently gave this result, I was unable to PCR confirm them. As a result, fosmids containing BHMS431 were not isolated during this project, and I decided not to proceed any further with this particular marker.

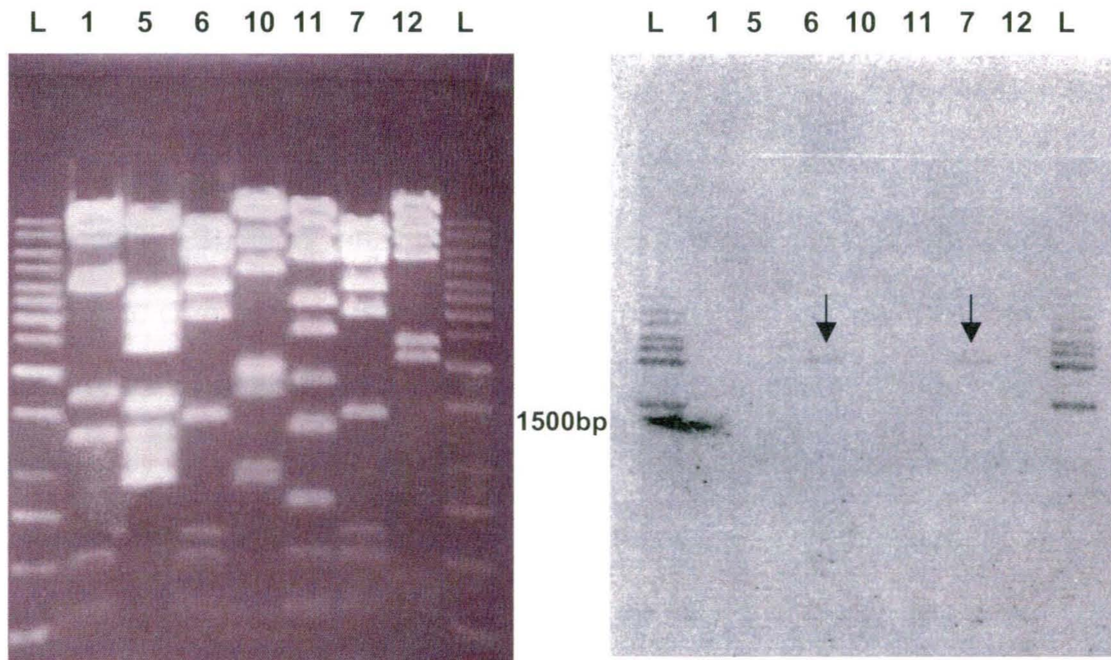
Figure 3.6 **PCR results of SsaBHMS431 amplified from 22 potentially positive fosmid clones using R primer only.** As expected, the positive control, labeled "C", shows no amplification. Lanes labeled "L" contain a 100bp ladder, in which the brightest band is 500 bp.



3.3.4 Characterization by *Hind*III digestion and Southern blotting

When clones were confirmed to be positive for the marker of interest by PCR, they were characterized by *Hind*III digestion and Southern blotting. The digests were run on a 0.8% agarose gel, and subjected to Southern blotting. The membrane was hybridized with the same 40-mer probe that was initially used to identify the positive clones. This was done for clones containing the marker Sal05SFU, Sall26SFU and Omm1228. Figure 3.7 shows a *Hind*III digest and hybridization of the Southern membrane of seven fosmid clones identified as being PCR positive for the marker Sal05SFU. Hybridization signals are visible for Sal05SFU clones 6 and 7, which appear to be identical clones based on the *Hind*III banding pattern. By adding the band sizes together, the size of Sal05SFU-6 and Sal05SFU-7 fosmids was determined to be approximately 35,000 bp.

Figure 3.7 0.8% agarose gel of *Hind*III digested Sal05SFU PCR positive clones and phosphor image of hybridized Southern membrane. Clones are labeled by number at the top of each lane. Lanes labeled as "L" contain a 1 Kb ladder. Hybridization signals are indicated with black arrows in lanes labeled "6" and "7", corresponding to clones Sal05SFU-6 and Sal05SFU-7. Note that these two clones appear to be identical based on their *Hind*III digest patterns.

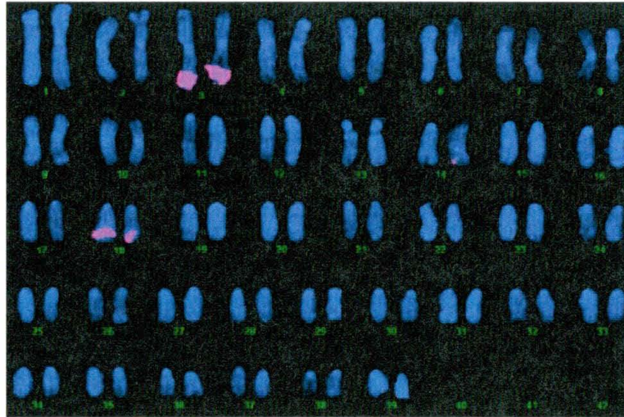


3.4 Fluorescent *in situ* Hybridization (FISH)

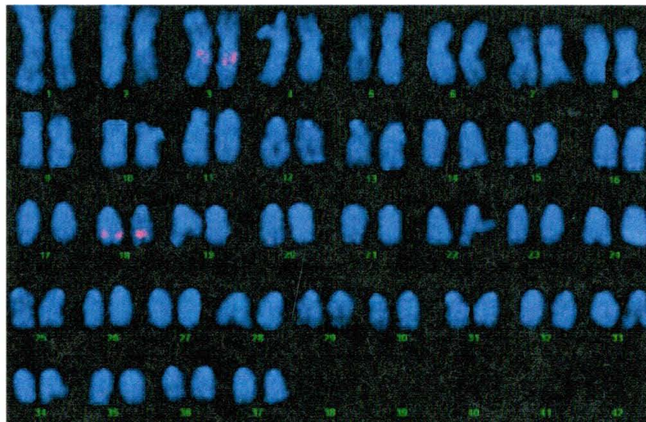
Positive fosmid clones were sent to Dr. Ruth Phillips' lab at Washington State University, Vancouver, WA. There, they were used to probe chromosomes prepared from a male Arctic char via FISH. Prior to the start of this project, fosmid clones positive for Sfo8LAV and Omy6DIAS had been sent to Dr. Phillips for FISH analysis. During the course of this project fosmid clones positive for Sal05SFU, Omm1228, and Sall26 were sent to be used as probes in FISH. Both Sfo8LAV and Omy6DIAS localized to two chromosome pairs; a large metacentric (chromosome 3), and a small acrocentric (chromosome 18). Sal05SFU only localized to chromosome 3, the same large metacentric as Omy6DIAS and Sfo8LAV. Omm1228 failed to localize to any particular chromosome pairs, but consistently hybridized in a scattered fashion across many chromosome pairs (data not shown), indicative of this clone containing a significant amount of repetitive DNA. FISH results for Sall26SFU, as well as a dual-FISH using Sfo8LAV and Sal05SFU have yet to be obtained. Figure 3.8 shows the FISH results for Omy6DIAS, Sfo8LAV and Sal05SFU.

Figure 3.8 **Results of Fluorescent *in situ* Hybridizations.** Both Omy6DIAS and Sfo8LAV hybridize to both chromosome pair 3 and 16, while Sal05SFU hybridized only to chromosome pair 3, identifying it as the Arctic char sex chromosome.

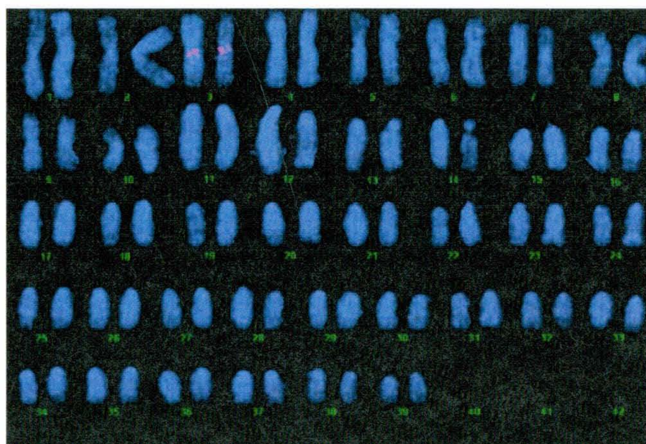
Omy6DIAS



Sfo8LAV



Sal05SFU



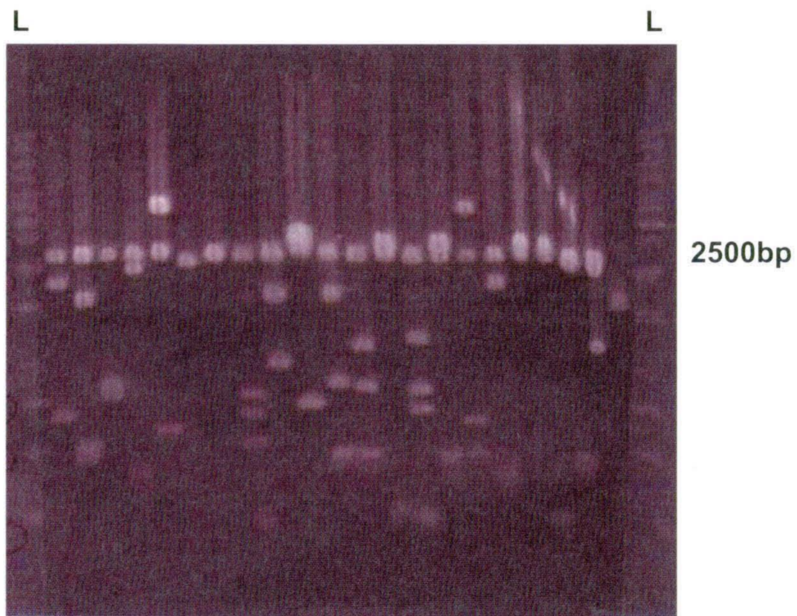
3.5 Sal05SFU-6 Sequencing

Fosmid clones were deemed true positives if they were both PCR positive and the probe hybridized to the *Hin*dIII digest/Southern blot membrane. Fosmids positive for Sal05SFU, Omm1228 and Sall26SFU were successfully isolated. Because Sal05SFU has 0% recombination with phenotypic sex and hybridized to a single chromosome pair, I decided to focus on the Sal05SFU-6 fosmid clone for detailed sequence analysis.

3.5.1 Shotgun library construction

To gain a better understanding of the genomic composition of the region surrounding *SEX*, a shotgun library of clone Sal05SFU-6 was generated. The Sal05SFU-6 shotgun library was constructed by first sonicating the fosmid DNA, end-repairing the sonicated DNA, running it on an agarose gel, and size-selecting the fragments between 2-5 Kb. These fragments were then cloned into pUC19 and transformed into *E. coli* cells. Hybrid recombinant, white colonies were selected and digested with *Pvu*II to separate the Sal05-6 fragment from the vector. The fragment sizes were visualized on a 0.8% agarose gel (Figure 3.9). From this, it was determined how much DNA to use per shotgun sequencing reaction based on the size of the cloned fragments.

Figure 3.9 **0.8% agarose gel showing *PVUII* digest of Sal05SFU-6 shotgun clones.** 22 shotgun clones from the Sal05SFU-6 shotgun library are shown. Lanes labeled "L" contain a 1Kb ladder. The bands corresponding to the linearized pUC19 vector can be seen at approximately 2500 bp.



3.5.2 Shotgun sequencing of Sal05SFU-6

The sequence coverage of Sal05SFU-6 was achieved by a combination of shotgun sequencing and directed sequence assembly. Sal05SFU-6 shotgun clones were sequenced directly from the pUC19 vector using the M13 forward (M13F) and reverse (M13R) primers. A total of 324 Sal05SFU-6 shotgun clones were sequenced, each with both the M13F and M13R sequencing primers. A total of 421 high-quality sequences were obtained as a result. These sequences initially collapsed into approximately 55 contiguous sequences, hereafter referred to as contigs, leaving approximately 249 singletons. Just prior to this stage, it was noted that shotgun sequencing was beginning to produce many duplicate reads (sequences that overlapped existing sequences in the project), and so another approach was adopted in order to extend the coverage of the contigs and thus obtain a more complete sequence of Sal05SFU-6.

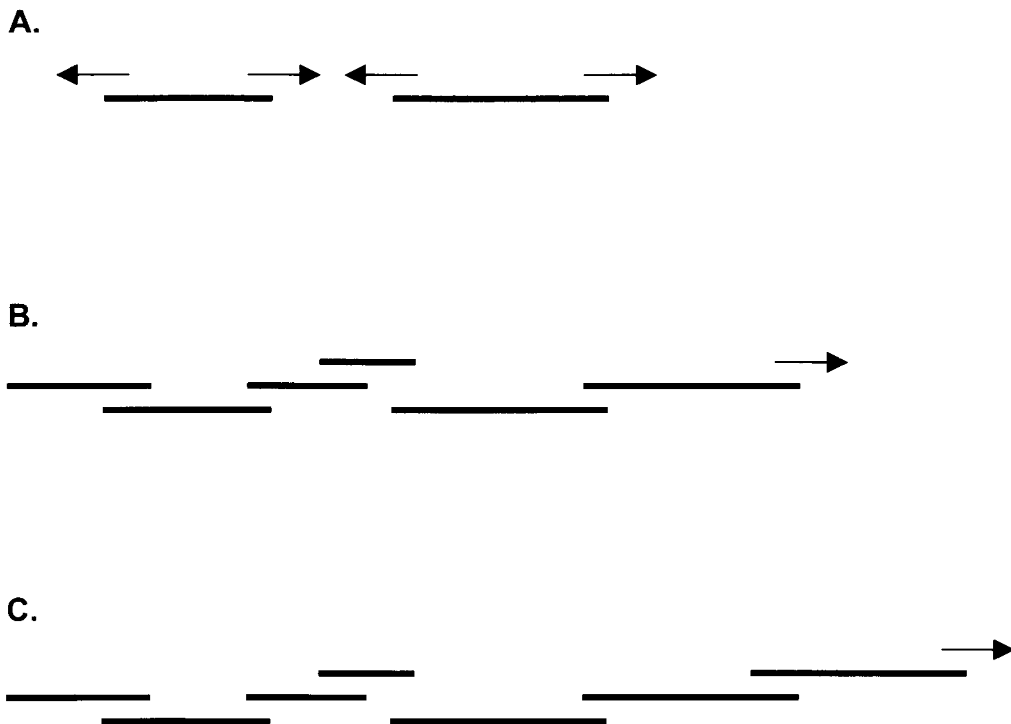
3.5.2.1 End-sequencing Sal05SFU-6

The Sal05SFU-6 fosmid clone was end-sequenced from both ends of the pCC1FOS™ vector using the pCC1™/pEpiFOS™ Forward and Reverse sequencing primers. The sequence obtained using the forward primer was approximately 450 nucleotides in length, while the sequence obtained using the reverse primer was approximately 325 nucleotides in length. These end-sequences allowed contigs corresponding to the ends of the genomic DNA insert to be identified, and also extended the sequence coverage of the clone.

3.6 Selected Primer Walking

In order to extend the contigs and fill in the gaps in the Sal05-6 fosmid sequence, a "selected primer walking" approach was taken. First, gaps that could be closed were identified by pairing up the ends of contigs based on the outer-most sequences. Primers were manually designed to sequence out from the ends of contigs to fill in sequence gaps between contigs, as well as to extend the ends of contigs that did not end-join with any other contig. An illustrative diagram of the selected primer walking process used in this project is shown in Figure 3.10.

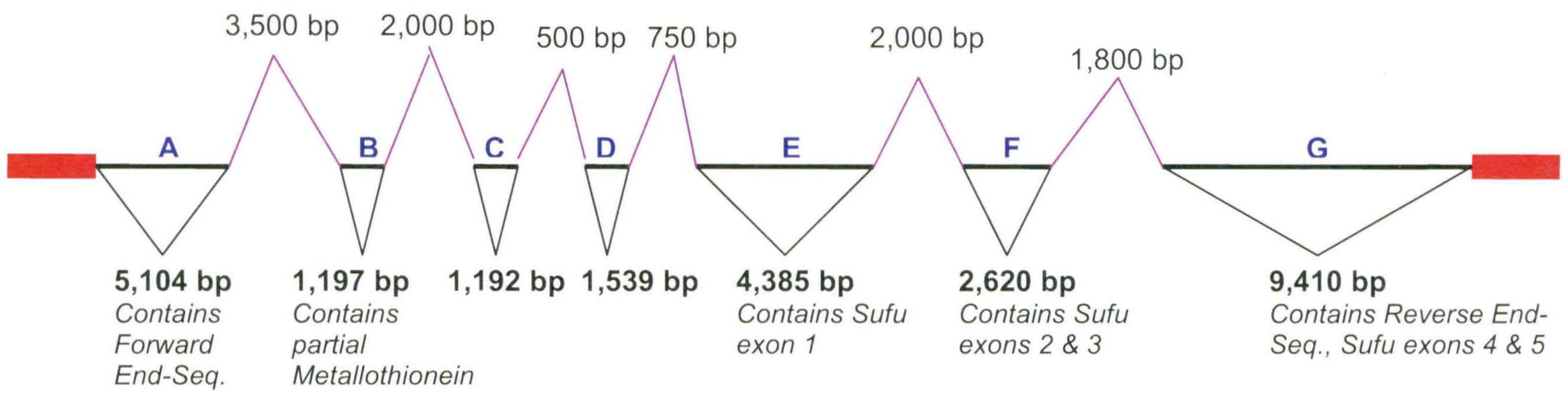
Figure 3.10 An illustrative diagram of the approach used to extend sequence coverage of Sal05SFU-6. A. Two contigs with paired ends. Primers are designed to sequence the gap between the two contigs, as well as to sequence out from the opposite ends. B. The gap between the two contigs has been filled, allowing the contigs to collapse into one large contig, and both opposite ends have been extended. C. As contig sequences are extended, new primers are designed from the extension sequences to continue walking out. Contig sequences are illustrated by lines. Primers are illustrated by arrows.



3.6.1 Extension and orientation of Sal05SFU-6 contigs

By using the selected primer-walking approach to extend sequence coverage of fosmid clone Sal05SFU-6, it was possible to collapse the number of contigs down to seven large, super-contigs. To determine the order and orientation of the super-contigs, PCR amplifications were performed using primers to amplify out from the ends of each contig, in all possible combinations. By analyzing the sizes of the amplification products, it was possible to determine the orientation of the contigs with respect to one-another. The seven super-contigs range in size from approximately 1,200 bp to 9,400 bp. Together, the contig sequences from the Sal05SFU-6 fosmid cover approximately 25,500 bp out of 35,000 bp. One of the super-contigs is anchored to the fosmid vector sequence, which contains the Sal05SFU-6 end-sequence from the pCC1™/pEpiFOS™ Forward sequencing primer. The Sal05SFU reference sequence did not collapse into any of the super-contigs, and did not overlap any of the singleton sequences, and is thus positioned in one of the unsequenced gaps. Contigs are anchored to both vector ends. An illustrative diagram summarizing the results of the extension and orientation of the Sal05SFU-6 fosmid sequence is shown in Figure 3.11.

Figure 3.11 Illustrative diagram summarizing sequence coverage and contig orientation of Sal05SFU-6. Black lines illustrate the contigs of the Sal05SFU-6 sequence. Red lines illustrate the pUC19 vector ends. Contigs are labeled arbitrarily by letters A-G. Approximate sizes (in base-pairs) are indicated for all contigs, as well as the gaps between them. Contigs containing genes and end sequences are indicated. Diagram is not drawn to scale.



3.7 Sequence Analyses

3.7.1 Sequence similarity searches

In order to identify candidate sex-determining genes within the Sal05SFU-6 fosmid sequence, all super-contigs were subjected to BLASTn and BLASTx analyses against the non-redundant database, and against the protein database (PDB) (NCBI, 2006). The super-contigs were also subjected to BLAST analysis against the University of Victoria salmonid EST database (http://web.uvic.ca/cbr/grasp/). Three of the contigs (A, C and D) did not return any significant hits (e-value > 0.01), other than alignments with known transposable elements or other known repetitive sequences. Three of the super-contigs (E, F and G) returned alignments to *Suppressor of Fused* (*Sufu*). Contig B contains a sequence that resembles metallothionein.

3.7.2 Gene descriptions

The following is a brief description of each of the genes identified in the fosmid clone Sal05SFU-6.

3.7.2.1 *Suppressor of fused* (*Sufu*)

Suppressor of fused (*Sufu*) was originally identified in *Drosophila melanogaster* for its ability to suppress overactive *fused* mutations, which include phenotypic abnormalities such as cuticular abnormalities and tumorous ovaries in females. *Sufu* is also a Gli/Ci-interacting protein, which acts as a negative regulator of Hedgehog signaling in both *Drosophila* and vertebrates, and has also

been implicated as a tumor suppressor as its mutations have been found in medulloblastoma and prostate cancer (Barnfield et al., 2005). *Sufu* has also been shown to have an influence on the nuclear entry rate of Sxl, the master sex-determining factor in *Drosophila* (Horabin et al., 2003).

3.7.2.2 ***Metallothionein***

The metallothioneins (MTs) are a family of Cys-rich, low molecular weight (MW ranging from 3,500 to 14,000 Da) proteins. MTs have the capacity to bind both physiological (e.g., Zn, Cu, Se) and xenobiotic (e.g., Cd, Hg, Ag) heavy metals through the thiol group of its cysteine residues, which represent nearly 30% of its amino acidic residues. The functions of metallothionein remain unclear, but experimental data relate MTs with protection against metal toxicity, regulation of physiological metals (Zn and Cu) and protection against oxidative stress (NCBI, 2006).

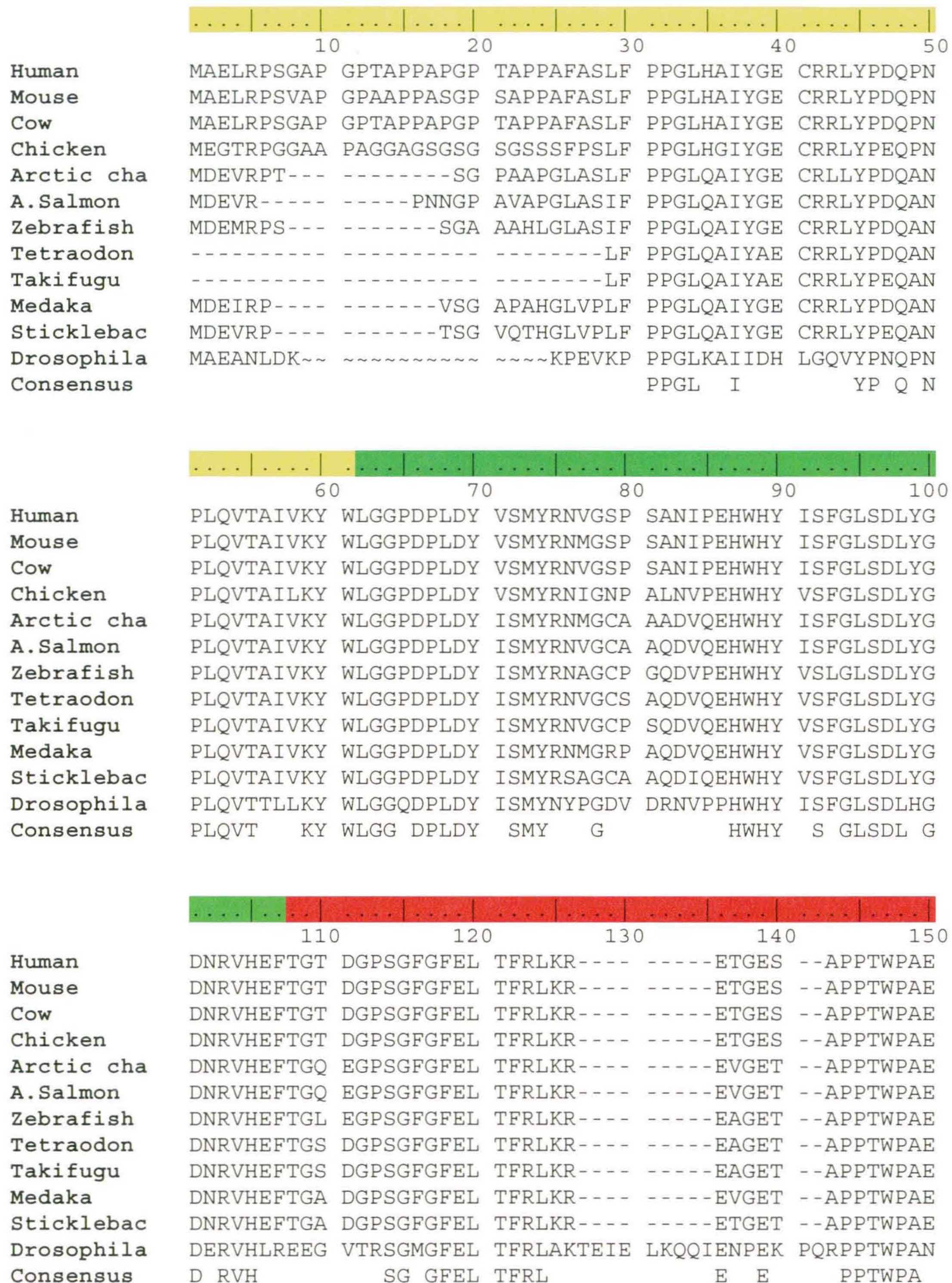
3.8 Characterization of Arctic char *Sufu*

3.8.1 *Sufu* amino acid alignment

The first five exons of the gene *Sufu* were identified within the Sal05SFU-6 contigs (Figure 3.11). The predicted amino acid sequences of these exons were aligned with the corresponding protein sequences from six fish, chicken, three mammals and *Drosophila* (Figure 3.12). The amino acids for the first five exons of *Sufu* in Atlantic salmon were obtained from the published EST database (http://web.uvic.ca/cbr/grasp; Rise et al., 2004), and the other species sequences were obtained through the NCBI protein database (PDB) (NCBI, 2006). Sufu

appears to be highly conserved in vertebrate species, and, as would be expected, Arctic char Sufu is more similar to that from other fish than mammals. Based on amino acid sequences in the first five exons, human and Arctic char Sufu are 87.1% similar, human and salmon Sufu are 86.6% similar, and human and medaka Sufu are 88.6% similar, while Arctic char and Atlantic salmon Sufu are 96.0% similar and Arctic char and medaka Sufu are 91.0% similar.

Figure 3.12 **Amino acid alignment of SUFU.** Alignment was produced using BioEdit Sequence Alignment Editor © (Version 7.0.5.3, 2005). Exon boundaries for the first five exons are indicated by different coloration above each exon sequence.



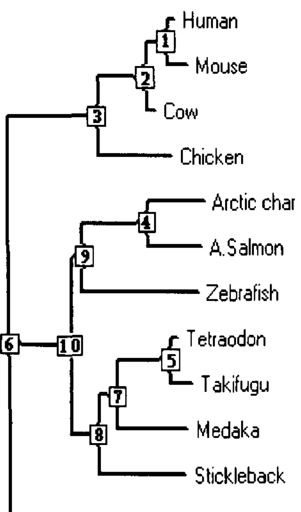
```
                  160        170        180        190        200
Human        LMQGLARYVF QSENTFCSGD HVSWHSPLD- NSESRIQHML LTEDPQMQPV
Mouse        LMQGLARYVF QSENTFCSGD HVSWHSPLD- NSESRIQHML LTEDPQMQPV
Cow          LMQGLARYVF QSENTFCSGD HVSWHSPLD- NSESRIQHML LTEDPQMQPV
Chicken      LMQGLARYVF QSENTFCSGD HVSWHSPLD- NSESRIQHML LTEDPQMQPV
Arctic cha   LMQGLARYVF QSENTFCSGD HISWHSPLD- NSESRIQHML LTEDPQMTPV
A.Salmon     LMQGLARYVF QSENTFCSGD HISWHSPLD- NSESRIRHML LTEDPQISPV
Zebrafish    LMQGLARYVF QSENTFCSGD HISWHSPLD- NSESRIQHML LTEDPQMHPV
Tetraodon    LMQGLARYVF QSENTFCSGD HVSWHSPLD- NSESRIQHML LTEDPQMQPI
Takifugu     LMQGLARYVF QSENTFCSGD HVSWHSPLD- NSESRIQHML LTEDPQMQPI
Medaka       LMQGLARYVF QSENTFCSGD HVSWHSPLD- NSESRIQHML LTEDPQMQPI
Sticklebac   LMQGLARYVF QSENTFCSGD HISWHSPLD- NSESRIQHML LTEDPQMQPV
Drosophila   LLQAIGRYCF QTGNGLCFGD NIPWRKSLDG STTSKLQNLL VAQDPQLGCI
Consensus    L Q   RY F Q  N  C GD    W   LD    S     L     DPQ
```

```
                  210        220        230        240        250
Human        QTPFGVVTFL QIVGVCTEEL HSAQQWNGQG ILELLRTVPI AGGPWLITDM
Mouse        RTPFGVVTFL QIVGVCTEEL HSAQQWNGQG ILELLRTVPI AGGPWLITDM
Cow          QTPFGVVTFL QIVGVCTEEL HAAQQWNGQG ILELLRTVPV AGGPWLITDM
Chicken      QTPFGVVTFL QIVGVCTEEL HAAQQWNGQG ILELLRTVPV AGGPWLITDM
Arctic cha   HTPFGLVNFL QIVGVCTEEL QAAQQWNGQG MLELLRGVHV A
A.Salmon     HTPFGLVNFL QIVGVCTEEL QAAQQWNGQG ILELLRGVHV AGGPWLITDM
Zebrafish    QTPFGHVSFL QIVGVCTEEL QAAQQWNGQG ILDLLRGVHI AGGPWLITDM
Tetraodon    QTPFGTVAFL QIVGVCTEEL QAAQQWNGQG ILELMRGVRV AGGPWLITDM
Takifugu     QTPFGTVAFL QIVGVCTEEL QAAQQWNGQG ILELMRGVRV AGGPWLITDM
Medaka       QTPFGTVSFL QIVGVCTEEL QAAQQWNGQG ILELMRGICI AGGPWLITDM
Sticklebac   QTPFGSVSFL QIVGVCTEEL QAAQQWNGQG ILELMRGVRV AGGSWLITDM
Drosophila   DTPTGTVDFC QIVGVFDDEL EQASRWNGRG VLNFLRQDMQ TGGDWLVTNM
Consensus     TP G V F  QIVGV   EL   A  WNG G  L   R
```

3.8.2 Phylogenetic analysis of Sufu

The alignment from amino acid 29 to 241 was used to construct a neighbor-joining tree (Figure 3.13). As expected, the sequences of Sufu from Arctic char and Atlantic salmon were most similar, and were thus grouped together on the tree. Using the calculated branch lengths between nodes and species (provided with the tree in Figure 3.13), the distance between node 9 and Arctic char is approximately 0.08361, while the distance between node 9 and Atlantic salmon is 0.08153. Based on these numbers, there did not appear to be an increased rate of change along the Arctic char lineage compared to Atlantic salmon or other fish, as might be expected if Sufu were the sex-determining factor in Arctic char but not in the other species. It should be noted that there is evidence for only one copy of *Sufu* in the fish genomes examined here.

Figure 3.13 Neighbor-joining tree based on the amino acid alignment of Sufu. The *Sufu* alignment between amino acids 29-241 was selected for the tree construction. The tree was assembled using BioEdit Sequence Alignment Editor© (Version 7.0.5.3, 2005). Branch lengths between nodes and species are indicated (units= average change per amino acid).

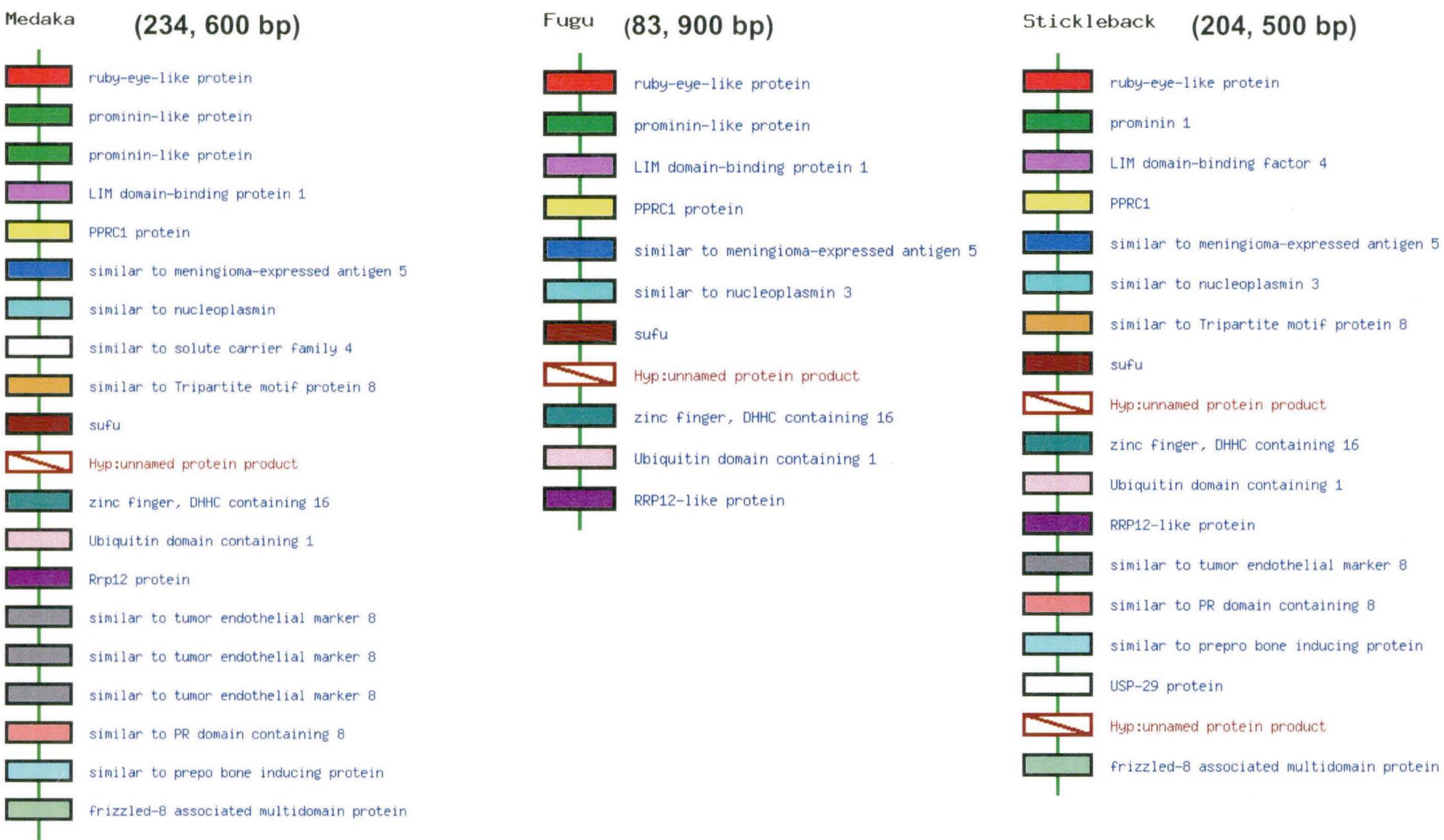


Between	And	Length
6	3	0.06198
3	2	0.03323
2	1	0.01363
1	Human	0.00446
1	Mouse	0.01394
2	Cow	0.00482
3	Chicken	0.04919
6	10	0.04289
10	9	0.00676
9	4	0.04542
4	Arctic char	0.03819
4	A.Salmon	0.03611
9	Zebrafish	0.07928
10	8	0.01892
8	7	0.01239
7	5	0.03630
5	Tetraodon	0.00429
5	Takifugu	0.01401
7	Medaka	0.04770
8	Stickleback	0.05851
6	Drosophila	1.09527

3.8.3 Comparative genomics

To identify additional candidate sex-determining genes for Arctic char, chromosomal regions adjacent to *Sufu* were compared across the genomes of medaka, Takifugu, threespine stickleback, tetraodon, and zebrafish (Figure 3.14). These species were chosen for this comparative analysis as it is known that *Sufu* is not the master sex-determining gene in medaka, and *DMY*, the master sex-determining gene in medaka is not the sex-determining gene in other fish species. Many proteins and predicted proteins are shared by at least four of the fish species, and in most cases, appear to be in the same order on linkage groups. Table 3.3 shows a summary of what is known about these genes and the proteins they encode.

Figure 3.14 **Comparative analysis of *Sufu* across other fish species.** Sizes indicate approximate distances from the start of the first exon (top) to the end of the last exon (bottom) on each group. Syntenic regions are shown for medaka (*Oryzias latipes*), fugu (*Takifugu rubripes*), threespine stickleback (*Gasterosteus aculeatus*), tetraodon (*Tetraodon nigroviridis*) and zebrafish (*Danio rerio*).



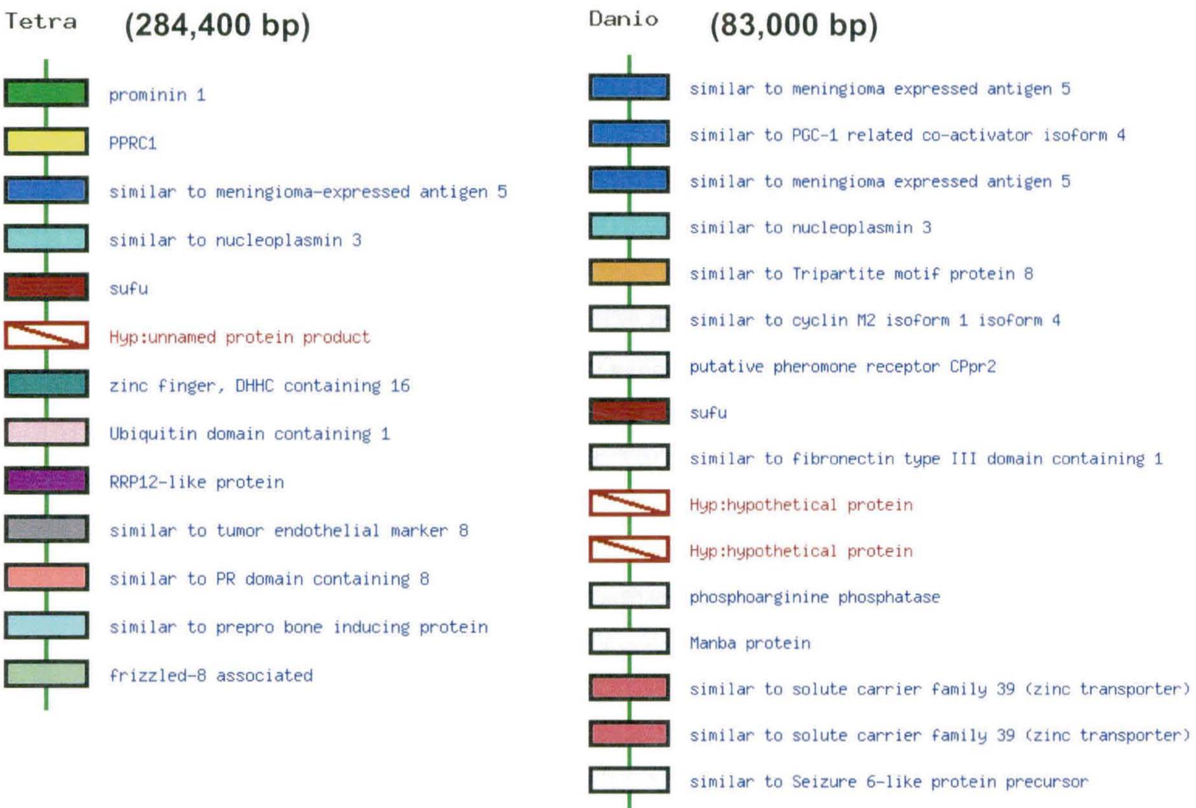

Tetra
(284,400 bp)
prominin 1
PPRC1
similar to meningioma-expressed antigen 5
similar to nucleoplasmin 3
sufu
Hyp:unnamed protein product
zinc finger, DHHC containing 16
Ubiquitin domain containing 1
RRP12-like protein
similar to tumor endothelial marker 8
similar to PR domain containing 8
similar to prepro bone inducing protein
frizzled-8 associated
Danio
(83,000 bp)
similar to meningioma expressed antigen 5
similar to PGC-1 related co-activator isoform 4
similar to meningioma expressed antigen 5
similar to nucleoplasmin 3
similar to Tripartite motif protein 8
similar to cyclin M2 isoform 1 isoform 4
putative pheromone receptor CPpr2
sufu
similar to fibronectin type III domain containing 1
Hyp:hypothetical protein
Hyp:hypothetical protein
phosphoarginine phosphatase
Manba protein
similar to solute carrier family 39 (zinc transporter)
similar to solute carrier family 39 (zinc transporter)
similar to Seizure 6-like protein precursor

Table 3.3 **Summary of roles/functions of genes/proteins identified through comparative genome analysis.** All of the information summarized here was obtained through NCBI (2006).

Protein/Gene Name	**General description and/or function(s)**
Ruby-eye protein	Organelle organization; mast cell granule secretion
Prominin	Stem cell surface marker; AC133 antigen
LIM domain-binding protein	Cofactor for LMO proteins, which have fundamental roles in development
PPRC1	Mitochondrial biogenesis; nucleic acid-binding protein
Meningioma-expressed antigen 5	Hydrolase activity (cytoplasm)
Nucleoplasmin	Widely expressed molecular chaperone (nucleus)
Solute carrier family 4	Anion exchanger (membrane transport); adaptor protein
Tripartate motif protein 8	Function unknown; zinc-ion binding protein
Zinc finger, DHHC domain containing 16	Putative transcription factor; probable palmitoyl transferase; may be involved in apoptosis
Ubiquitin, domain containing 1	Cell growth and/or maintenance
Rrp12	Ribosomal RNA processing and export
Tumor endothelial marker 8	Function unknown; plasma membrane protein
PR, domain containing 8	DNA and metal ion-binding protein; regulation of DNA-dependant transcription
Prepro bone inducing protein	Bone formation and remodeling
Frizzled-8 associated multidomain protein	Adaptor protein; molecular link between *Wnt*-signaling and actin cytoskeleton
USP29	Peptidase and hydrolase activity (Ubiquitin cycle)

Protein/Gene Name	**General description and/or function(s)**
Thrombospondin	Calcium-ion binding protein; cell adhesion
PGC-1 related co-activator	Roles in germ cell migration and pole cell development
Cyclin M2 isoform 2	Regulation of cell cycle (nucleus); cell communication and signal transduction
Pheromone receptor CPpr1	Contains two zinc-fingers (LIM2)
Fibronectin type III, domain containing 1	Activation of G-protein signaling (by similarity)
Phosphoarginine phosphatase	Calcium-dependent removal of serine and threonine-bound phosphate groups
Manba	Beta-mannosidase and hydrolase activity; cation-binding (metabolic processes)
Seizure related 6	Function unknown; type 1 transmembrane receptor
Solute carrier family 39	Function unknown; metal ion (zinc) transporter

CHAPTER 4 DISCUSSION

The objective of this study involved the integration of Arctic char genetic linkage maps associated with the sex-determining factor with Arctic char chromosomes. This enabled the identification of the sex chromosome pair in Arctic char, and further characterization of the sex-determining region in order to identify candidate sex-determining genes.

4.1 Fosmid Library Screening

4.1.1 Arctic char-specific primer and probe design

The basis of this project was to screen an Arctic char fosmid library for clones containing sex-linked markers. In order to do this, primers and hybridization probes unique to Arctic char had first to be designed. However, designing the primers to be unique to Arctic char was difficult due to the lack of a reference salmonid genome sequence. Despite this drawback, each of the designed primer sets was successful in amplifying the respective sex-linked marker from Arctic char genomic DNA. Furthermore, the hybridization probes designed directly from Arctic char marker sequences facilitated the isolation of fosmid clones containing the sex-linked markers Sal05SFU, Omm1228 and Sall26. This success demonstrates the usefulness of this method for exploring genomes for genes of interest without first having a reference sequence. However, to overcome the difficulties of extending fosmid contigs due to the high

density of repetitive sequences in the sex-determining region, an Arctic char bacterial artificial chromosome (BAC) genomic library should be constructed.

4.1.2 Arctic char Linkage Group 4 (AC-4)

It is notable that *SEX* was mapped to a linkage group end in both of the University of Guelph genetic maps with markers appearing on one side only, but mapped roughly to the middle of the SFU linkage map (markers appear on both sides of *SEX*). This difference may be due to a difference in strains that were used to generate the linkage maps. Unfortunately, this cannot be confirmed as, although it is know that the Davidson lab genetic linkage map was generated from a male of the Tree River strain, it is not known what strain(s) the two University of Guelph genetic linkage maps were created from. However, it is known that the male parent was not a Tree River strain, and was derived from the Fraser River (Labrador), Nauyuk Lake or a hybrid of these strains. Another possible explanation for the differences in location of *SEX* on the different linkage maps is that the markers appearing on either side of *SEX* in the Davidson lab map may not have been informative in the families that were used to make the University of Guelph maps.

There are some sex-linked markers for which primers and hybridization probes were successfully designed, but have not yet been used to probe the fosmid library. Future identification and isolation of clones containing the markers SsaBHMS490, SsaBHMS7.033 and SSOSL32, as well as successful FISH using

the Omm1228 and Sall26 clones identified in this study will aid in the orientation of AC-4 with respect to the sex chromosome.

4.2 Fluorescent *in situ* Hybridization of Fosmids Containing Sex-Linked Clones

4.2.1 Identification of the Arctic char sex chromosome

This study reports the first successful identification of the Arctic char sex chromosome. As seen from FISH analyses (Figure 3.8) fosmid clones containing the sex-linked markers Omy6DIAS and Sfo8LAV hybridize to both a large metacentric (chromosome 3), and a small acrocentric (chromosome 18), while fosmids containing the sex-linked Sal05SFU only localized to chromosome 3. These results strongly suggest that chromosome 3, rather than the acrocentric chromosome 18, is the Arctic char sex-determining chromosome.

Support for the conclusion that the Arctic char sex-chromosome is a metacentric chromosome comes from a study by May et al. (1989), in which it was concluded that *SEX* in Arctic char is across the centromere from three biochemical allozyme markers. Logically, for this to be true, *SEX* must be on a chromosome that has arms on both sides of a centromere, and therefore it must be on a metacentric chromosome. The conclusion from the study by May et al. (1989) is consistent with the results produced in this project (i.e., that *SEX* is on one of the largest metacentric chromosomes), solidifying the identification of the Arctic char sex chromosome.

4.2.2 The Arctic char sex-determining region

It is worth discussing the point that two of the sex-linked markers, Sal05SFU and Omy6DIAS, hybridize to opposite ends of the same chromosome arm (Figure 3.8), yet they show 0% recombination between one another in approximately 200 meioses (Figure 3.1). It is possible that a very high density of repetitive DNA lies in the region between those two markers. This accumulation of repetitive DNA is to be expected on sex-chromosomes. It has been suggested that the presence of repetitive elements in the sex-determining region may hinder recombination with the homologous chromosome, which in turn may help to maintain the male-specific location of the sex-determining gene (Charlesworth, 1991). Suppression of recombination around the sex-determination locus is a hallmark of sex chromosomes (Peichel et al., 2004), and is consistent with the identification of the Arctic char sex chromosome in this study.

4.2.3 Possible locations of *SEX* on the Arctic char sex chromosome

The interpretation of the data presented here suggests three possibilities for the placement of *SEX* on the sex chromosome of Arctic char. Each of these scenarios deserves its own consideration and discussion.

1) *SEX lies near the centromere, on the other side of Sal05SFU*

It has been suggested that sex-determining loci are near centromeres so as to reduce recombination, and thus to maintain the locus as "male" specific (Bull, 1983). In addition, it has been observed that male salmonids have lower rates of recombination than females in chromosomal regions near the

centromeres (Sakamoto et al., 2000), supporting the idea that *SEX* would be "protected" by its placement adjacent to the centromere. Based on the observation that the sex-linked Sal05SFU and Sfo8LAV clones hybridized very closely to the centromere, and that those markers are very tightly linked with *SEX*, it is entirely possible that Arctic char *SEX* lies very close to the centromere.

2) *SEX lies near the telomere, on the other side of Omy6DIAS/Sfo8LAV*

The *Omy6DIAS* and *Sfo8LAV* marker hybridize close to the telomere of the third chromosome arm, and it is possible that *SEX* lies even closer to the telomere than they do. However, it is known that males have higher rates of recombination for genes near telomeres (Sakamoto et al., 2000). Also, if a sex-specific region is at the telomere of the Y chromosome, it could be more easily translocated to other chromosomes. This has been observed in insects that are at an early stage of sex-chromosome differentiation, where the *SEX* locus appears to have jumped to different chromosome pairs (Stein et al., 2002). It then seems more favorable for *SEX* to be located away from the less stable telomeric region. However, it was recently shown that Atlantic salmon *SEX* is placed at the telomeric end of the sex chromosome (Artieri, 2005), and so this may also be the case in Arctic char.

3) *SEX lies between Sal05SFU and Omy6DIAS/Sfo8LAV*

The FISH results of Omy6DIAS/Sfo8LAV and Salo5SFU also suggest that *SEX* lies between these markers. These markers all show a 0% recombination with phenotypic sex, yet they hybridize quite far apart in FISH analysis (Figure 3.8). As mentioned previously, it is highly likely that there is a region of repetitive sequence between those markers. It is known that regions of highly repetitive DNA are relatively gene-poor (Wallrath, 1998), and additional studies have shown that genes within such regions may lack active transcription (Schulz and Tyler, 2005). It seems more likely, then, that *SEX* lies within a non-repetitive, euchromatic region. In medaka, the Y chromosome-specific region appears to be very small (only a few hundred kilobases in length) (Kondo et al., 2003), so *SEX* in Arctic char could in fact lie between dense regions of repetitive DNA, in a small region of Y-specific sequence.

As nothing is known about the primary genetic mechanism of sex-determination in salmonids, it is also entirely possible that *SEX* exists within the highly repetitive region, perhaps within a block of heterochromatin. Under this scenario, it is possible that the primary genetic sex-determining signal is not present in the Arctic char fosmid library used in this study, and techniques to locate *SEX* such as chromosome "walking" will be ineffective. Alternate methods to locate *SEX* may include screening salmonid EST (expressed sequence tag) libraries for candidate sex-determining genes based on motifs, such as the DM domain of the sex-determining gene in medaka (Matsuda et al., 2002). Although

a large number of salmonid cDNA libraries have been developed using mRNA isolated from various tissues and developmental stages (Rise et al., 2004), it is possible that such libraries do not contain any ESTs that represent the primary sex-determining factor of Arctic char. Sex-determining genes in other non-fish vertebrate species are known to display narrow temporal and tissue-specific expression windows during male embryogenesis (Payen et al., 1996; Meyers-Wallen, 2003). However, it has been shown that the primary sex-determining factor in medaka, DMY, is expressed in various tissues during the entire life span of a male medaka (Ohmuro-Matsuyama et al., 2003). Although DMY is the only primary sex-determining signal identified in teleost fish to date, it does suggest that fish possess sex-determining genes that may be displayed in a non-temporal and tissue specific pattern.

One piece of evidence required to reveal a more precise location of *SEX* would be a dual FISH hybridization using clones containing the markers Sall26 and Omm1228. Since *SEX* appears to lie between these two markers on LG-4 (Figure 3.1-C), a FISH analysis using both of these probes would narrow the location of *SEX* down to the region between the regions where those two hybridize.

4.3 Sequences Analyses

4.3.1 Genes identified in Sal05SFU-6

Sequence similarity searches for the seven Sal05SFU-6 contigs revealed significant similarity to two genes; *Supressor of fused* (*Sufu*) and *Metallothionein* (*Met*). The most substantial alignment was to *Sufu*, as the first five full exons of this gene were identified in the Sal05SFU-6 contigs (Figure 3.12).

4.3.1.1 Arctic char *Supressor of Fused (Sufu)*

Sufu exon identification assisted and confirmed the Sal05SFU-6 contig order and orientation (Figure. 3.11). As seen in Figure 3.13, the alignments of the amino acids sequences corresponding to the first five exons of *Sufu* reveal that Sufu is highly conserved across vertebrate species. Although a preliminary phylogenetic analysis of Sufu in this study suggests there has not been an increased rate of change along the Arctic char lineage compared to Atlantic salmon or other fish, *Sufu* can still be considered a candidate sex-determining gene until evidence suggests otherwise. *Sufu* can be considered a candidate gene for sex-determination in Arctic char, simply based on it being positioned in the sex-determining region. As mentioned previously, *Sufu* is linked to sex-determination in *Drosophila* by influencing the nuclear entry rate of Sxl, the master sex-determining gene (Horabin et al., 2003). Again, this makes it a logical candidate for the sex-determining gene in Arctic char. Clearly, additional research investigating the link between *Sufu* and sex-determination needs to be

conducted, especially since the current link (mentioned here) exists in *Drosophila,* albeit a non-vertebrate system.

Although expression analyses were not conducted on *Sufu* or any of the other candidate sex-determining genes in this study, a search against the Atlantic salmon EST database (http://web.uvic.ca/cbr/grasp; Rise et al., 2004) using the *Sufu* exons revealed that it was identified in a mixed cDNA library made from brain, kidney and spleen tissues. So from this, it is known that in Atlantic salmon, *Sufu* is at least expressed in one, two, or all of those tissues. However, assuming adult tissues were used to make the cDNAs, this does not reveal anything about the expression during early development, so it does not mean that it is not solely expressed in gonads at that point. It also does not mean that the expression pattern of *Sufu* in Arctic char is the same as in Atlantic salmon.

4.3.1.2 *Metallothionein*

It is notable that a portion of *Metallothionein* (*Met*) was identified in Sal05SFU-6, as *MetBSFU* was also mapped as a sex-linked marker on the Arctic char linkage maps (Figure 3.1-B). This offers confirmation that the characterized Sal05SFU-6 sequence is in the vicinity of Arctic char *SEX*. However, the alignment was very small (approximately 30 amino acids) and displayed only 40% similarity in that region. It should be noted that there was no sequence similarity to Arctic char *Met*, whose sequence has been determined (Gerpe et al., 1998). Thus, additional sequence extension on the contig that aligned with *Met*

needs to be performed in order to give a more definitive BLAST result to this gene.

4.3.1.3 Repetitive and transposable elements

Most of the BLASTx results returned alignments to repetitive sequence or transposable elements. Although uninformative in terms of genes, as mentioned previously, highly repetitive sequences are an indicator of a major sex-determining region. In addition, the accumulation of transposable elements is believed to be to be linked to the evolution of heteromorphic sex chromosomes (Harvey et al., 2002), and so it is also not surprising to find them in the Arctic char sex-determining region. Thus, the abundance of repetitive and transposable elements identified in Sal05SFU-6 in fact suggests that the sequence-characterized fosmid indeed lies close to or within the sex-determining region.

From this study, the sex-determining region of Arctic char appears to display many properties of a nascent Y chromosome. That is, by using sex-linked markers, the region has been shown to have suppression of recombination, accumulation of repetitive sequences, and heterozygosity in males (Harvey et al., 2002). To definitively confirm the identity of the Arctic char sex-determining region, it should next be shown that there is a substantial nucleotide divergence from its homologous region on the X chromosome. It would then be favorable to extend the X and Y contigs to identify the gene or genes responsible for sex-determination in Arctic char.

4.3.2 Candidate Arctic char sex-determining genes

Since there has been no master sex-determining gene identified in Arctic char thus far, any gene identified within the sex-determining region that is not well characterized for this species is a potential candidate (Figure 3.14, Table 3.3). Those genes that possess DNA-binding sites, or resemble transcription factors (e.g., *PR*, domain containing 8 and *Sufu*) are particularly attractive as candidates. Two currently known sex-determining genes, *SRY* and *DMY* both encode transcription factors, and are both in vertebrate systems. It then follows that a system such as alternative splicing is less likely to be the main determining factor of sex in salmonids than it is in *Drosophila*.

As indicated earlier, to test whether the candidate genes identified in this study are truly candidate master sex-determining genes, a logical next step would be expression analyses of the genes on Arctic char tissues. If the genes show expression in only the gonad tissues (or are not widely distributed in other tissues), particularly early in development, then they would be considered true candidate genes. However, as discussed earlier, genes that do not display such a pattern of expression should not immediately be ruled out as candidate sex-determining genes.

In the future, identification of the master sex-determination gene in Arctic char will not necessarily mean that the sex-determining gene has been discovered for all salmonids. Although the master sex-determining gene *SRY* is shared across mammals, it is unlikely that a single sex-determining gene will be

shared across all species of salmonids. Many fish sex-determining systems appear to have evolved recently, including salmonid sex chromosomes (Ezaz et al., 2006), and it is likely that they are still in the process of rapid differentiation. If this is the case, then it is reasonable to suggest that salmonids do not share one single master sex-determining gene. *DMY* is not the master sex-determining gene in some of medaka's most closely related species (Kondo et al., 2003), and therefore, it would not be a surprise if even the most closely related species of salmonids do not share the same sex-determining gene. The elucidation of the gene(s) responsible for sex-determination in Arctic char and other salmonids remains a challenge, but the results produced in this thesis have brought us one step closer to this goal.

REFERENCES

Agrawal, A.F. 2006. Evolution of sex: Why do organisms shuffle their genotypes? Current Biology 16: R696-R704.

Allendorf, F.W. and Thorgaard, G.H. 1984. *Tetraploidy and the evolution of salmonid fishes*, pp. 1-53 in Evolutionary Genetics of Fishes, edited by BJ Turner. Plenum Press, New York.

Artieri, C. 2005. “Characterization of the Atlantic salmon sex-determining chromosome”. M.Sc. dissertation. Simon Fraser University, Burnaby, Canada.

Barnfield, P.C., Zhang, X., Thanabalasingham, V., Yoshida, M. and Hui, C.C. 2005. Negative regulation of Gli1 and Gli2 activator function by Suppressor of fused through multiple mechanisms. Differentiation 73 (8); 397-405.

Baroiller, J.-F., Guigen, Y. and Fostier, A. 1999. Endocrine and environmental aspects of sex differentiation in fish. Cell. Mol. Life Sci. 55; 910-931.

Bell, G. 1982. The masterpiece of Nature: The evolution and genetics of sexuality. University of California press, Berkely. 635 p.

Beye, M., Hasselmann, M., Fondrk, M.K., Page, R.E., and Omholt, S.W. 2003. The gene csd is the primary signal for sexual development in the honeybee and encodes an SR-type protein. Cell 114: 419-429.

Brennen, J. and Capel, B. 2004. One tissue, two fates: molecular genetic events that underlie testis versus ovary development. Nature Reviews Genetics 5: 509-521.

Brunelli, J.P. and Thorgaard, G.H. 2004. A new Y-chromosome specific marker for Pacific salmon. Trans. Am. Fish. Soc. 133 (5); 1247-1253.

Bull, J.J. 1983. Evolution of sex determining mechanisms. Benjamin/Cummings publishing, Don Mills. 316 p.

Bulmer, M. 1987. Sex determination in fish. Nature 326: 440-441.

Charlesworth, B. 1991. The evolution of sex chromosomes. Science 251; 1030-1033.

Charlesworth, B. 2004. Sex determination: Primitive Y chromosomes in fish. Current Biology 14; R745-R747.

Charlesworth, B. 2006. The evolutionary biology of sex. Current Biology 16: R693-R695.

Chevassus, B. 1989. "Modification du phenotype sexuel et du mode de reproduction chez les poisons salmonidés: Inversion sexuelle, gynogénèse, hybridation interspécifique et polypoïdisation." Ph.D. dissertation, Université de Paris XI, Orsay, France.

Christoffels, A., Koh, E.G., Chia J.M., Brenner, S., Aparicio, S. and Venkatesh, B. 2004. Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. Mol. Biol. Evol. 21; 1146-1151.

Conover, D.O and Heins, S.W. 1987. Adaptive variation in environmental and genetic sex determination in fish. Nature 326: 496-498.

Cossins, A.R. and Crawford, D.L. 2005. Fish as models for environmental genomics. Nature Reviews Genetics. 6: 324-333.

Cunado, N., Barrios, J., San Miguel, E., Amaro, R., Fernandez, C., Hermidia, M. and Santos, J.L. 2002. Synaptonemal complex analysis in oocytes and spermatocytes of threespine stickleback, *Gasterosteus aculeatus* (Teleosti, Gasterosteidae). Genetica 114; 53-56.

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*). Canadian Journal of Fisheries and Aquatic Sciences 48; 1606-1612.

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.

Devlin, R.H., Stone, G.W. and Smailus, D.E. 1998. Extensive direct tandem organization of a long repeat DNA sequence on the Y chromosome of Chinook salmon (*Oncorhynchus tshawytscha*). J. Mol. Evol. 46; 277-287.

Devlin, R.H., Biagai, C.A. and Smailus, D.E. 2001. Genetic mapping of Y-chromosomal DNA markers in Pacific salmon. Genetica 111; 43-58.

Devlin, R.H. and Nagahama, Y. 2002. Sex determination and sex differentiation in fish: an overview of genetic, physiological and environmental influences. Aquaculture 208; 191-364.

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

Ezaz, T., Stiglec, R., Veyrunes, F. and Marshall Graves, J.A. 2006. Relationships between vertebrate ZW and XY sex chromosome systems. Current Biology 16; R736-R743.

Flybase: A Database of *Drosophila* Genes and Genomes, 2006. <www.flybase.org>

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. U.S.A. 91-1628-1631.

Froschauer, A., Körting, C., Katagiri, T., Aoki, T., Asakawa, S., Shimizu, N., Schartl, M. and Volff, J.-N. 2002. Construction and initial analysis of bacterial artificial chromosome (BAC) contigs from the sex-determining region of the platyfish *Xiphophorus maculatus*. Gene 295; 247-254.

Gerpe, M., Kling, P. and Olsson, P.E. 1998. Metallothionein Gene expression in Arctic Char (*Salvelinus alpinus*) following metal and PCB exposure. Marine Environmental Research 46; 551-554(4).

Gharbi, K. 2001. "Construction d'une carte génétique pertielle du genome tétraploide de la truite commune (*Salmo trutta*)". Ph.D. dissertation, Institut National Agronomique Paris-Grignon, Paris, France.

Gharbi, K., Gautier, A., Danzmann, R.G., Gharbi, S., Sakamoto, T., Hoyheim, B., Taggart, J.B., Cairney, M., Powell, R., Krieg, F., Okamoto, N., Ferguson, M.M., Holm, L.-E. and Guyomard, R. 2006. A linkage map for brown trout (*Salmo trutta*): chromosome homeologies and comparative genome organization with other salmonid fish. Genetics 172; 2405-2419.

Gilbey, J., Verspoor, E., McLay, A. and Houlihan, D. 2004. A microsatellite linkage map for Atlantic salmon (*Salmo salar*). Anim. Genet. 35; 98-105.

Guyomard, R., Mauger, S., Tabet-Canale, K., Martineau, S., Genet, C., Krieg, F. and Quillet, E. 2006. A Type I and Type II microsatellite linkage map of Rainbow trout (*Oncorhynchus mykiss*) with presumptive coverage of all chromosome arms. BMC Genomics 7; 302.

Hartley, S.E. 1989. Chromosomes and constitutive heterochromatin distribution in Arctic char, *Salvelinus alpinus* (*L.*) (Pisces:Salmonidae). Genetica 79; 161-166.

Harvey, S.C., Masabanda, J., Carrasco, L.A.P., Bromage, N.R., Penman, D.J. and Griffin, D.K. 2002. Molecular-cytogenetic analysis reveals sequence differences between the sex chromosomes of *Oreochromis niloticus*: evidence

for an early stage of sex-chromosome differentiation. Cytogenet. Genome. Res. 79; 76-80.

Hoegg S., Brinkman, H., Taylor, J.S. and Meyer, A. 2004. Phylogenetic timing of fish-specific genome duplication correlates with the diversification of teleost fish. J. Mol. Evol. 59; 190-203.

Horabin, J.I., Walthall, S., Vied, C. and Moses, M. 2003. A positive role for Patched in Hedgehog signaling revealed by the intracellular trafficking of Sex-lethal, the *Drosophila* sex determination master switch. Development 130; 6101-6109.

Howard, R.S. and Lively, C.M. 1998. The maintenance of sex by parasitism and mutation accumulation under epistatic fitness functions. Evolution 52; 604-610.

Hunter, G.A., Donaldson, E.M., Stoss, J. and Baker II. 1983. Production of monosex female groups of chinook salmon (*Oncorhynchus tshawytscha*) by the fertilization of normal ova with sperm from sex-reversed females. Aquaculture 33; 355-364.

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 of male heterogemety. Trans. Am. Fish. Soc. 111; 367-372.

Johnstone, R., Simpsons, T.H., Youngsen, A.F. and Whitehead, C. 1979. Sex reversal in salmonid culture. PartII. The progeny of sex reversed rainbow trout. Aquaculture 18; 13-19.

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

Kallman, K.D. 1984. A new look at sex determination in poeciliid fishes. In: Turner, B.J. (Ed.), Evolutionary Genetics in Fishes, Plenum, New York, pp. 95-171.

Keller, G.H. and Manak, M.M. 1989. DNA Probes. 1st Ed. Stockton Press, New York.

Kettlewell, J.R., Raymond, C.S. and Zarkower, D. 2000. Temperature-dependant expression of turtle *Dmrt1* prior to sexual differentiation. Genesis 26; 174-178.

Kondo, M., Nanda, I., Hornung, U., Asakawa, S., Shimizu, N., Mitani, H., Schmid, M., Shima, A. and Schartl, M. 2003. Absence of the candidate male sex-determining gene *dmrt1b(Y)* of medaka from other fish species. Current Biology 13; 416-420.

Kondo, M., Nanda, I., Hornung, U., Schmid, M. and Schartl, M. 2004. Evolutionary origin of the Medaka Y chromosome. Current Biology 14; 1664-1669.

Kondrashov, A.S. 1993. Classification of hypotheses on the advantage of amphimixus. J. Hered. 84: 372-387.

Koopman, P. 2005. Sex determination: a tale of two *Sox* genes. Trends in Genetics 21; 367-370.

Leberg, P. 1998. Influence of complex sex determination on demographic stochasticity and population viability. Conservation Biology 12: 456-459.

Lively, C.M. and Loyd, D.G. 1990. The cost of biparental sex under individual selection. Am. Nat. 135: 489-500.

Marshall, G. 2002. The rise and fall of SRY. Trends in Genetics 18; 259-264.

Matsuda, M., Nagahama, Y., Shinomiya, A., Sato, T., Matsuda, C., Kobayashi, T., Morrey, C.E., Shibata, N., Asakawa, S., Shimizu, N., Hori, H., Hamaguchi, S. and Sakaizumi, M. 2002. *DMY* is a Y-specific DM-domain gene required for male development in the medaka fish. Nature 417; 559-563.

Matsuda, M. 2005. Sex determination in the teleost medaka, *Oryzias latipes*. Annu. Rev. Genet. 39; 293-307.

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

McCarthy, J. 2004. An Arctic char linkage map. Bisc 498 Undergraduate Research.

McElreavey, K., Vilain, E., Abbas, N., Herskowitz, I. and Fellous, M. 1993. A regulatory cascade hypothesis for mammalian sex determination: SRY represses a negative regulator of male development. Proc. Natl. Acad. Sci. U.S.A. 90; 3368-3372.

McGowan, C., Goel, A., Davidson, E.A., Mosher, A., Ng, S., and Davidson, W. 2004. Enhancing the diversity of Arctic char (*Salvelinus alpinus*) broodstock; A final report prepared for Icy Waters Ltd. Describing the genetic analysis carried out by the Davidson lab at Simon Fraser University.

Metcalf, V.J. and Gemmel, N.J. 2006. Sexual genotype markers absent from small numbers of male New Zealand *Oncorhynchus tshawytscha*. Journal of Fish Biology 68: 136-143.

Meyers-Wallen, V.N. 2003. Sry and Sox9 expression during canine gonadal sex determination assayed by quantitative reverse transcription-polymerase chain reaction. Mol Reprod. Dev. 65; 373-381.

Moen, T., Hoyheim, B., Munck, H. and Gomez-Raya, L. 2004. Alinkage map of Atlantic Salmon (*Salmo salar*) reveals an uncommonly large difference in recombination rate between the sexes. Anim. Genet. 35; 81-92.

Mosher, A. 2005. Plant and Animal Genome XIII Conference Abstract; P595

Nakayama, I., Baigi, C.A., Koide, N. and Devlin, R.H. 1999. Identification of a sex-linked GH pseudogene in one of two species of Japanese salmon (*Oncorhynchus masou* and *O. rhodurus*). Aquaculture 173; 65-72.

Nanda, I., Shan, Z., Schartl, M., Burt, D.W., Koehler, M., Nothwang, H., Grutzner, F., Paton, I.R., Windsor, D., Dunn, I., Engel, W., Staeheli, P., Mizuno, S., Haaf, T. and Schmid, M. 1999. 300 million years of conserved synteny between chicken Z and human chromosome 9. Nat. Genet. 21; 258-259.

Nanda, I., Hornung, U., Kondo, M., Schmid, M. and Schartl, M. 2003. Common spontaneous sex-reversed XX males of the medaka *Oryzias latipes*. Genetics 163; 245-251.

Nanda, I., Kondo, M., Hornung, U., Asakawa, S., Winkler, C., Shimizu, A., Shan, Z., Haaf, T., Shimizu, N., Shima, A., Schmid, M. and Schartl, M. 2002. A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. Proc. Natl. Acad. Sci. USA 99; 11778-11783.

NCBI, 2006. Nucleotide-nucleotide BLASTx: <http://www.ncbi.nlm.gov/BLAST/>

Nelson, J.S. 1984. Fishes of the World, 3rd Edition. Wiley and Sons, New York.

Ohmuro-Matsuyama, Y., Matsuda, M., Kobayashi, T., Ikeuchi, T. and Nagahama, Y. 2003. Expression of DMY and DMRT1 in various tissues of the medaka (*Oryzias latipes*). Zoolog. Sci. 20; 1395-1398. Erratum in Zoolog. Sci. 20; 1491.

Ohno, S., Wolf, U. and Atkin, N.B. 1968. Evolution from fish to mammals by gene duplication. Heriditas. 59; 169-187.

Ohtsuka, M., Kikuchi, N., Nogami, M., Inoko, H., Ozato, K. and Kimura, M. 2002. Rapid screening of a novel arrayed medaka (*Oryzias latipes*) cosmid library. Marine Biotechnology 4; 173-178.

Onozato, H. 1989. Androgenesis in masu salmon. Physiol. Ecol. Japan. Spec. 1; 543.

Otto, S.P. and Gerstein, A.C. 2006. Why have sex? The population genetics of sex and recombination. Biochemical Society Transactions 34: 519-522.

Otto, S.P. and Nuismer, S.L. 2004. Species interactions and the evolution of sex. Science 304; 1018-1020.

Ottolenghi, C. and McElreavey, K. 2000. Deletions of 9p and the quest for a conserved mechanism of sex determination. Mol. Genet. Metab. 71; 397-404.

Ou, G. 2005. Generation of SCAR markers from AFLP marker CAC/ACT-465. Bisc 498 Undergraduate Research.

Payan, E., Pailhoux, E., Abou Merhi R., Gianquinto, L., Kirszenbaum, M., Locatelli, A. and Cotinot, C. 1996. Characterization of ovine SRY transcript and developmental expression of genes involved in sexual differentiation. Int. J. Dev. Biol. 40; 567-575.

Phillips, R. and Ihssen, P.E. 1985. Chromosome banding in salmonid fish: nucleolar organizer regions in Salmo and Salvelinus. Can. J. Genet. Cytol. 27; 433-440.

Phillips, R.B. 2001. Application of fluorescence *in situ* hybridization (FISH) to fish genetics and genome mapping. Marine Biotechnology 3; 5145-5152.

Phillips, R. and Ráb, P. 2001. Chromosome evolution in the Salmonidae (Pisces): an update. Biol. Rev. 76; 1-25.

Phillips, R.B., Konkol, N.R., Reed, K.M. and Stein, J.D. 2001. Chromosome painting supports lack of homology among sex chromosomes in *Oncorhynchus*, *Salmo*, and *Salvelinus* (Salmonidae). Genetica 111; 119-123.

Phillips, R.B., Matsuoka, M.P. and Reed, K.M. 2002. Characterization of charr chromosomes using fluorescence in situ hybridization. Env. Biol. Fishes 64; 223-228.

Phillips, R.B., DeKoning, J., Morasch, M.R., Park, L.K. and Devlin, R.H. 2007. Identification of the sex chromosome pair in chum salmon (*Oncorhynchus keta*) and pink salmon (*Oncorhynchus gorbuscha*)

Piferrer, F., Baker, I.J. and Donaldson, E.M. 1993. Effects of natural, synthetic, aromatizable, and nonaromatizable androgens in inducing male sex-differentiation in genotypic female Chinook salmon (*Oncorhynchus tshawytscha*). General and Comparative Endocrinology 91: 59-65.

Pomiankowski, A., Nöthiger, R. and Wilkins, A. 2004. Evolution of the *Drosophila* sex determination pathway. Genetics 166; 1761-1773.

Prodöhl, 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.

Raymond, C.S., Murphy, M.W., O'Sullivan, M.G., Bardwell, V.J. and Zarkower, D. 2000. *Dmrt1*, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. Genes. Dev. 14; 2587-2595.

Raymond, C.S., Shamu, C.E., Shen, M.M., Seifert, K.J., Hirsch, B., Hodgkin, J. and Zarkower, D. 1998. Evidence for evolutionary conservation of sex-determining genes. Nature 391; 691-695.

Reed, K.M. and Phillips, R.B. 1995. Molecular cytogenetic analysis of the double CMA3 chromosome in lake trout, *Salvelinus namaycush*. Cytogenet. Cell. Genet. 70; 104-107.

Rise, M.L., von Schaulburg, K.R., Brown, G.D., Mawer, M.A., Devlin, R.H., Kuipers, N., Busby, M., Beetz-Sargent, M., Alberto, R., Gibbs, A.R., Hunt, P., Shukin, R., Zeznik, J.A., Nelson, C., Jones, S.R., Smailus, D.E., Jones, S.J., Schein, J.E., Marra, M.A., Butterfield, Y.S., Stott, J.M., Ng, S.H., Davidson, W.S and Koop, B.F. 2004. Development and application of a salmonid EST database and cDNA microarray: data mining and interspecific hybridization characteristics. Genome Res. 14; 478-490.

Sakamoto, T., Danzmann, R.G., Gharbi, K., Howard, P., Ozaki, A., Khoo, S.K., Woram, R.A., Okamoto, N., Ferguson, M.M., Holm, L.E., Guyomard, R. and Hoyheim, B. 2000. A microsatellite linkage map of rainbow trout (Oncorhynchus mykiss) characterized by large sex-specific differences in recombination rates. Genetics 155; 1331-1345.

Sambrook, J., Fritsh, E.F. and Maniatis, T. 1989. Molecular cloning. 2nd Ed. Cold Spring Harbour Lab Press, U.S.A. 1.38-1.39.

Scherer, G. 2002. *The molecular genetic jigsaw puzzle of vertebrate sex determination and its missing pieces*, pp. 225-239 in The Genetics and Biology of Sex Determination, edited by D. Chadwick and J. Goode. John Wiley & Sons Ltd, West Sussex. 266 p.

Schulz, L. and Tyler, J. 2005. Heterochromatin focuses on senescence. Mol. Cell. 17; 168-170.

Schütt, C. and Nöthiger, R. 2000. Structure, function and evolution of sex-determining systems in Dipteran insects. Development 127; 667-677.

Stein, J.D., Phillips, R.B. and Devlin, R.H. 2001. Identification of the Y chromosome in Chinook salmon (*Oncorhynchus tshawytsha*). Cytogenet. Cell Genet. 92; 108-110.

Stein, J.D., Reed, K.M. and Phillips, R.B. 2002. Isolation and characterization of a sex-linked microsatellite locus from the lake trout sex chromosome. Cytogenet. Cell. Genet. 92; 108-110.

Stothard, P and Pilgrim, D. 2003. Sex-determination gene and pathway evolution in nematodes. Bioessays 25; 221-231.

Swain, A. 2002. Vertebrate sex determination: A new player in the field. Current Biology 12; R602-R603.

Taggart, J.B., Prödohl, 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.

Thorgaard, G.H. 1977. Heteromorphic Sex Chromosomes in Male Rainbow Trout. Science 196; 900-902.

Thorgaard, G.H. 1983. Chromosomal differences among rainbow trout populations. Copeia 3; 650-662.

Veith, A.-M., Froschauer, A., Körting, C., Nanda, I., Hanel, R., Schmid, M., Schartl, M., and Volff, J.-N. 2003. Cloning of the *dmrt1* gene of *Xiphophorus maculatus*: *dmY/dmr1Y* is not the master sex-determining gene in the platyfish. Gene 317: 59-66.

Vilain E. 2002. *Anomolies of human sexual development: clinical aspects and genetic analysis*, pp. 43-53 in The Genetics and Biology of Sex Determination, edited by D. Chadwick and J. Goode. John Wiley & Sons Ltd, West Sussex. 266 p.

Volff, J.-N. 2005. Genome evolution and biodiversity in teleost fish. Heredity 94; 280-294.

Wallrath, L.L. 1998. Unfolding the mysteries of heterochromatin. Curr. Opin. Genet. Dev. 8; 147-153.

Werren, J.H. and Beukeboom, L.W. 1998. Sex determination, sex ratios, and genetic conflict. Annu. Rev. Ecol. Syst. 29: 233-261.

Woram , R.A., Gharbi, K., Sakamoto, T., Hoyheim, B., Holm, L.E., Naish, K., McGowan, C., Ferguson, M.M., Phillips, R.B., Stein, J., Guyomard, R., Cairney,

M., Taggart, J.B., Powell, R., Davidson, W. and Danzmann, R.G. 2003. Comparative genome analysis of the primary sex-determining locus in salmonid fishes. Genome Res. 13; 272-80.

Woram. R.A., McGowan, C., Stout, J.A., Gharbi, K., Ferguson, M.M., Hoyheim., B., Davidson, E.A., Davidson, W.S., Rexroad, C., and Danzmann, R.G. 2004. A genetic linkage map for Arctic char (*Salvelinus alpinus*): evidence for higher recombination rates and segregation distortion in hybrid versus pure strain mapping parents. Genome 47: 304-315.

Wright Jr., 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.

Young, W.P., Wheeler, P.A., Coryell, V.H., Keim, P. and Thorgaard, G.H. 1998. A detailed linkage map of rainbow trout produced using doubled haploids. Genetics 148; 839-850.

Zhang, Q., Nakayama, I., Fujiwara, A., Kobayashi, T., Oohara, I., Masaoka, T., Kitamura, S. and Devlin, R.H. 2001. Sex identification by male-specific growth hormone pseudogene (GH-psi) in Oncorhynchus masou complex and related hybrid. Genetica 111; 111-118.

Zhang, J. 2004. Evolution of *DMY*, a newly emergent male sex-determination gene of medaka fish. Genetics 166; 1887-1895.