THE EVOLUTION OF SEX CHROMOSOMES AND SEX DETERMINATION MECHANISMS IN STICKLEBACK FISHES (GASTEROSTEIDAE)

James Ralph Urton

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2011

Program Authorized to Offer Degree:
Molecular and Cellular Biology

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

James Ralph Urton

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of the Supervisory Committee:

Catherine Peichel

Catherine Peichel

Reading Committee:

Susan Biggins

Susan Biggins

Susan Parkhurst

Susan Parkhurst

Catherine Peichel

Catherine Peichel

Date: July 21, 2011

In presenting this dissertation in partial fulfillment of the requirements for the doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, 1-800-521-0600, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature James Ralph Uhrig

Date 21 July 2011

University of Washington

Abstract

THE EVOLUTION OF SEX CHROMOSOMES AND SEX DETERMINATION MECHANISMS IN STICKLEBACK FISHES (GASTEROSTEIDAE)

James Ralph Urton

Chair of the Supervisory Committee:
Dr. Catherine L. Peichel
Associate Member, Human Biology Division; Fred Hutchinson Cancer Research Center
Affiliate Associate Professor, Department of Biology; University of Washington

In many vertebrate species, a bipotential gonad develops into either testes or ovaries based on the action of an initial sex determination signal. Sex determination signals vary widely among species and can be genetic or environmental signals. Closely related species can have different sex determination mechanisms. Among species with genetic sex determination mechanisms, such rapid turnover is easily seen in species with independently evolved sex chromosome systems. However, the mechanisms by which sex chromosome systems and sex determination mechanisms turnover are poorly understood.

Within the stickleback fish family (Gasterosteidae), at least five sex chromosome systems arose in the past 40 million years. However, we do not know the evolutionary relationships among these sex chromosome systems, nor do we know if the same sex determination gene is found in different stickleback sex chromosome systems. To help understand the evolutionary relationships among the stickleback sex chromosome systems, I undertook genetic and cytogenetic screens to map the ZZ-ZW sex chromosome system of the fourspine stickleback, *Apeltes quadracus*,

relative to the threespine stickleback, *Gasterosteus aculeatus*. I discovered that the *A. quadracus* ZZ-ZW sex chromosomes arose independently of the other stickleback sex chromosome systems. I also discovered one *A. quadracus* population with no visible sex chromosomes. To address whether sticklebacks share the same sex determination gene, we first wish to identify the sex determination gene in *G. aculeatus*, which has a XX-XY sex chromosome system. Thus, I designed and executed a high-throughput sequencing transcriptome screen and identified hundreds of genes that are differentially expressed between the sexes during the early stages of gonadal differentiation. These genes will shed light on how sexual differentiation pathways have evolved in the stickleback family and assist in the continued search for the *G. aculeatus* sex determination gene. In addition, my screen confirmed the lack of a global dosage compensation mechanism for X chromosome genes in this species. These results will spawn future studies to understand how sex chromosomes arose in this family, how the gene content of sex chromosomes can change over time, and how dosage tolerance evolves in a complex vertebrate genome.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

In *The Life of Timon of Athens*, William Shakespeare wrote, "'Tis not enough to help the feeble up, but to support him after." While I am far from feeble, I relied on support (and much more) from a number of people over the course of my graduate studies and research.

Mike Shapiro and Tom Near caught fish for genetic and cytogenetic experiments. Ken Olivera provided equipment and storage facilities for a fishing expedition to Massachusetts. I thank members of the Moens, Nelson, Olson, Tewari, Trask, and Tsukiyama laboratories at the Fred Hutchinson Cancer Research Center ("the Hutch") for equipment and supplies. Stephen Uy skillfully translated passages from Robert Matthey's works, including *Les chromosomes des vertébrés*.

For assistance with planning and executing my RNA-Seq experiments, I thank the staff of the array and sequencing facilities at the Hutch. I am especially grateful to Jeff Delrow and Andy Marty, who provided advice and reassurance. I could always count on Elizabeth Jensen, armed with an understated smile, to return genotyping data to me quickly in my search for *SEX*-linked markers in sticklebacks. Jenni Risler kindly provided training to use the Hutch's quantitative PCR facilities.

Marty McIntosh and his group are an invaluable resource at the Hutch for the analysis of next generation sequence data. Within that group, I am especially grateful to Martin Morgan and Jerry Davison, who executed the alignments I used to identify differentially expressed genes by RNA-Seq. Matt Fitzgibbon, a generous soul, mined my RNA-Seq data to begin the process of identifying transcripts from Y chromosome genes. I am grateful for Matt's insight and enthusiasm, which helped move this project past its early growing pains.

I thank all of the administrators and staff members in the Molecular and Cellular Biology graduate program. MaryEllin Robinson, Terry Duffey, Diane Darling, and Milli Morris always made themselves available to address any issue, crisis, or question. Michele Karantsavelos is the caretaker for all graduate students at the Hutch. I cannot imagine how hard this task must be, nor can I adequately express how grateful I am that she enjoys her job so much. Every graduate student

at the Hutch benefits from her warm smiles and her crisis-management skills, particularly her reassurances of "It's going to be okay!" and "Don't worry. We'll fix it."

The members of my supervisory committee were a source of tremendous wisdom, advice, and encouragement. Occasionally, they also kicked me in the butt for my own good. I am especially grateful to Sue Biggins, who agreed to be on my committee after I declined to join her laboratory. I cannot thank her enough for her steadfast support and frank advice over the years. I would also like to thank Sue and Susan Parkhurst for serving on my reading committee, and providing guidance on the scope and tone of this dissertation. Barb Trask's proofreading skills were an invaluable contribution to a manuscript that has been accepted for publication in *Cytogenetic and Genome Research*. I also thank Barb for kindly interrupting her retirement to attend my final examination. Graham Young served as my graduate school representative for five years, and was a source of both enthusiasm and encouragement at my meetings. When Graham discovered that he could not attend my final examination, he personally found a replacement in the form of his jovial colleague, Evan Gallagher. Sue, Susan, Barb, and Graham also served as references for me when I applied for postdoctoral positions, and I know their advocacy helped me find my new research home.

I wish to thank all the talented and dedicated scientists of the Peichel laboratory. Postdoctoral fellows Jun Kitano, Kerry Marchinko, Brian Fritz, and Matt Arnegard, and graduate student Jen Cech provided helpful ideas throughout my graduate studies. Graduate students Margaret Mills and Tiffany Malek were always available with suggestions, jokes, hugs, and entertaining stories as we pursued our own (very different) projects. Joe Ross initiated the laboratory's cytogenetic studies on stickleback sex chromosome evolution while he was a graduate student, and trained me on cytogenetic techniques. He overlooked my slow responses to his quick wit, and encouraged me to forge my own path in my graduate studies. Amanda Bruner and Shaun McCann kept our fish facilities running like a well-oiled machine. I also thank Amanda for giving me such a warm welcome to the Peichel laboratory, both as a rotation student and a permanent member in 2006. She is now a trusted friend, and I know I can always count on her to sample a platter of fried

midwestern cuisine with me. Shaun McCann is probably the most unique human being I will ever meet. His generosity, curiosity, mild obsessive-compulsive tendencies, and oddball sense of humor combine to make to make him a warm and helpful soul. Shaun is also a talented cytogeneticist, and his beautiful stickleback chromosome spreads adorn portions of two figures from Chapter 3 of this dissertation. I also thank Brooke Magnusson, Alex Moreno, and especially Pam Lindberg for serving as such dedicated and friendly administrative assistants for the Peichel laboratory over the years.

I had three wonderful "assistants" in the laboratory to help me move my RNA-Seq projects forward. Amanda Bruner spent the summer of 2008 extracting DNA and RNA and genotyping stickleback larvae for a pilot next generation sequencing screen. Amanda is also a learned scholar of fish gonad development, and I counted on her insight as I selected embryonic and larval tissues for RNA-Seq. The next summer, Reyes Balcells arrived from Barcelona, and spent ten weeks extracting DNA and genotyping thousands of embryonic and larval tissue samples. I was truly lucky to have Reyes in the laboratory that summer. She is a thorough and dedicated scientist, and I have no doubt she will excel in anything she puts her mind to. In the summer of 2010, high school student Jesse Clurman spent five weeks in the laboratory genotyping individuals for quantitative PCR experiments. His friendly demeanor and curiosity were a breath of fresh air as I wrapped up this project.

I could not have completed my Ph.D. without the support and encouragement of Abby Wark. Her warmth, honesty, sincerity, and quick wit were a source of tremendous strength, particularly during difficult periods where (to put it mildly) "nothing was working." I am grateful for her unique scientific insights and "go get 'em" attitude. Abby and her husband, Barry, also christened me an unofficial babysitter of their twins, Jack and Bella. I treasure these "parenting lessons" with the Wark family, and I have missed Abby a great deal since she moved to Boston.

Initially, I was intimidated by Anna Greenwood's immense wisdom. But over the years, I have learned to embrace and treasure her brilliant scientific insights. Anna is a talented and generous scientist. She freely gave both her time and skills to help me in almost every project I have undertaken in graduate school. Anna

especially devoted time and energy to proofread countless manuscripts and proposals, help me analyze RNA-Seq data, and tutor me on quantitative PCR and statistical tests. She also leads by example, showing firsthand how an open and collaborative approach to science yields unexpected and exciting insights for every experiment. Most of all, Anna encouraged me to have confidence in my own abilities, and to push myself beyond the perceived limits of those abilities.

I am truly blessed to have Katie Peichel as my advisor and mentor. From my first day in her laboratory, she challenged me to improve and sharpen my abilities as a scientist, and to advance my natural curiosity with rigorous and thoughtful experiments. But Katie's skills as an advisor encompass more than just bench work. She encouraged me to explore all of my interests, even if they veered from traditional career paths. Thus, with her exuberant support, I pursued new passions for public outreach and teaching. Her infectious enthusiasm and dedicated work ethic helped me through even the most difficult periods of graduate school. I cannot thank Katie enough for her steadfast support and frank advice over the past five years. I hope I have made her proud, and that my future career will serve as a testament to her dedication as a scientist and an advisor.

I could not have made it to (or through) graduate school without the love and support of my family. Words cannot express the gratitude I feel to my parents, who sacrificed so much to raise me, and always urged me to aim high. I am especially indebted to my mother, who taught me the benefits of self-improvement and supported me every step of the way in graduate school. Ellen, my older sister, inspired me to go to graduate school when she earned her own graduate degree. She has always been my faithful advocate and a source of strength.

Finally, I must thank my boyfriend, best friend, and partner Zach Carstensen. The road to and through graduate school has not been an easy one, but I cannot imagine taking this journey (or any journey) with anyone other than him. We cannot count on much in this world. But, if I have learned anything over our years together, it is that I can always count on his love, loyalty, patience, and capacity for forgiveness.

DEDICATION

To my teachers.

CHAPTER ONE

Introduction to the Evolution of Sex Determination Mechanisms, Sexual Differentiation Pathways, and Sex Chromosomes

SUMMARY

In vertebrates, a sex determination mechanism acts to differentiate an initially bipotential gonad primordium into either testes or ovaries. Testes and ovaries, in turn, govern the development of mature male and female forms. Though many downstream genes in sexual differentiation pathways are conserved, even among vertebrates and invertebrates, the upstream sex determination mechanism can vary even between closely related species. Thus, sex determination mechanisms can evolve rapidly, and can even transition between genetic and environmental mechanisms. Transitions between genetic sex determination mechanisms can also lead to sex chromosome diversity, both within and between species. This type of sex chromosome diversity is seen among the stickleback fishes (Gasterosteidae), where several sex chromosome systems have evolved independently within the last 40 million years. In this family, there are XX-XY and ZZ-ZW sex chromosome systems, as well as two sex chromosome systems derived from independent Y-autosome fusions. However, we do not know how many times sex chromosomes have arisen independently in this family. In addition, we do not know if different stickleback sex chromosome systems house the same sex determination gene. My research goals, which are described in detail in Chapters 2, 3, and 4, address these questions.

SEX DETERMINATION AND SEXUAL DIFFERENTIATION

Sex determination and sexual differentiation are the developmental processes by which females and males develop. Among vertebrates, a bipotential gonad develops into either testes or ovaries based on different genetic and environmental instructions (Nakamura, 1989). Signals from the gonad (testes or ovaries) govern the development of mature male or female forms. For the purposes of this dissertation, sex determination is defined as the mechanism that initiates the sex-specific development of an immature gonad into a testis or an ovary. Sexual differentiation is the process by which sex-specific differences in both the gonad and somatic tissues develop. The consequences of sex determination and sexual differentiation are far-reaching, not only governing gametogenesis but also the extensive sexual dimorphism in morphological, physiological, and behavioral traits observed in many species. However, these consequences begin with the sex determination mechanism, which acts to promote the gonad development and sexual differentiation process toward the development of one sex over the other.

I deliberately separate "sex determination" (a precise, narrow developmental event) from "sexual differentiation" (the broad downstream consequences of the sex determination mechanism) to ease in illustrating how evolution has shaped these processes.

The Principle of "Masters Change, Slaves Remain"

Sex determination mechanisms are not conserved among animals. Sex determination mechanisms can be a single gene (simple genetic sex determination), multiple genes (polygenic sex determination), an environmental signal, or a "polyfactorial" sex determination mechanism (a mixture of genetic and environmental cues) (Bull, 1983). Vertebrates provide a particularly compelling group to study these variable mechanisms. For example, in some vertebrate lineages, such as marsupial and placental mammals, a simple genetic sex determination mechanism has persisted for over 100 million years (Veyrunes *et al.*, 2008). Although birds have a different simple genetic sex determination mechanism than mammals, it is also quite stable (Handley *et al.*, 2004). However, both genetic and environmental sex

determination mechanisms (and everything in between) have been observed in many other vertebrate lineages such as fish, amphibians, and reptiles. Furthermore, different sex determination mechanisms can be found in closely related species in these groups. This is exemplified by fish in the genus *Oryzias*; even closely related species separated by only a few million years have different sex determination mechanisms (Matsuda, 2005; Takehana *et al.*, 2007; Tanaka *et al.*, 2007). These observations have led to the theory that sex determination mechanisms can turnover rapidly, leading to transitions between different mechanisms (whether they be genetic, environmental, or both).

In contrast to the variability of sex determination mechanisms, gonad fate and sexual differentiation appear to be governed by a conserved set of genes and gene families. For some gene families, such as the *Doublesex-MAB3* (DM) family (Raymond *et al.*, 1998; Volff *et al.*, 2003), this conservation even extends across invertebrate and vertebrate species. Other conserved genes and gene families include the SOX (*SRY* on the X) family in vertebrates (Koopman, 2005) and the *transformer* family in dipteran insects (Schutt and Nothiger, 2000; Graham *et al.*, 2003; Pomiankowski *et al.*, 2004). The molecular roles for these conserved genes vary from transcriptional regulators (DM and SOX genes) to hormone synthesis, such as the conserved vertebrate aromatase gene *Cyp19* (Callard *et al.*, 2001).

This downstream stability of sexual differentiation pathways stands in stark contrast to the rapid turnover of the signals initiating sexual determination (Marin and Baker, 1998; Schutt and Nothiger, 2000; Haag and Doty, 2005). This observation was once neatly summarized with the phrase, "Masters change, slaves remain" (Graham *et al.*, 2003). Behind this credo lie a number of outstanding questions regarding how evolutionary pressures could act so differently in sex determination and sexual differentiation (Wilkins, 1995). The goal of understanding how evolutionary pressures shape sex determination and differentiation pathways is made even more difficult by the effects of genetic sex determination mechanisms on genome architecture in the form of degenerate sex chromosomes.

ORIGIN AND EVOLUTION OF SEX CHROMOSOMES

Sex Chromosomes and Genetic Sex Determination

In species with a simple sex determination mechanism, a sex determination gene initiates the downstream sexual differentiation pathway to promote the development of one sex over the other. For example, in most mammalian species, the male-determining gene *SRY* promotes gonadal differentiation into testes (Berta *et al.*, 1990; Koopman *et al.*, 1990; Koopman *et al.*, 1991). Sex determination genes like *SRY* reside on one member of a pair of sex chromosomes. Sex chromosome pairs differ between the sexes in a given species. One sex carries one pair of two identical sex chromosomes, while the other sex carries a single copy of two different sex chromosomes. A sex chromosome system is XX-XY (male heterogametic) when females carry two identical sex chromosomes (X and X), and males have a single X chromosome and a Y chromosome. In contrast, for species with ZZ-ZW sex chromosomes (female heterogametic), males have two identical sex chromosomes (Z and Z), and females have a single Z chromosome and a W chromosome. The Y chromosome is male-limited, the W chromosome is female-limited, and each is present in a hemizygous state (Bull, 1983). At a minimum, the X and Y (or Z and W) differ by the presence or absence of a sex determination gene. For example, the sex determination gene in birds, *DMRT1*, is found on the Z chromosome, but is absent from the W (Smith *et al.*, 2009). Thus, in this case, the sex determination gene acts primarily through a dosage effect. Like in the mammalian case of *SRY*, in the medaka fish *Oryzias latipes*, the male sex determination gene *DMY* is only found on the Y chromosome (Matsuda *et al.*, 2002; Nanda *et al.*, 2002). The example of medaka is particularly instructive, because the X and Y chromosomes in this species have identical gene content over their entire length, save for a relatively small 258kbp region unique to the Y chromosome that contains only a single protein-coding gene, the sex determination gene *DMY* (Kondo *et al.*, 2006). Thus, the medaka Y chromosome could be thought of as a "minimalist" design, since the region of the Y chromosome that is unique to males contains the minimum number of genes required for a sex chromosome: a single sex determination gene.

Suppression of Recombination and Sex Chromosome Degeneration

Many model sex chromosome systems do not merely differ by the presence or absence of the sex determination gene, and the sex chromosome pair can be distinguished using cytogenetic methods. Such heteromorphic sex chromosomes have evolved independently from autosomes in animals, plants, and fungi and share a number of characteristics (Fraser and Heitman, 2005). Typically, the X and Y chromosomes do not recombine over part (if not all) of their length (Bergero and Charlesworth, 2009). Since the Y chromosome is present in a hemizygous state, the non-recombining region of the Y is the male-specific region (and the corresponding region on a W chromosome is the female-specific region). The region of the Y chromosome that still recombines with the X during meiosis is the "pseudo-autosomal region". For the remainder of this section, when describing the features of XY and ZW sex chromosome pairs, I will describe only X and Y sex chromosomes. However, similar features are found on Z and W sex chromosome pairs.

Recombination suppression between X and Y sex chromosomes may initially encompass a small area and spread gradually through suppressive chromatin remodeling or chromosome rearrangements. Alternatively, large chromosome rearrangement events, such as an inversion, can instantaneously establish a male-specific region encompassing many genes (Lahn and Page, 1999). While the sex determination gene lies within the non-recombining region of the Y or X chromosome, the presence of the sex determination gene alone is not thought to be sufficient to select for a loss of recombination and a spread of recombination suppression. Instead, the main theory to explain the spread of recombination suppression proposes an important role for sexually-antagonistic loci in this process (Rice, 1987; Charlesworth *et al.*, 2005). By this theory, a mutation can arise with a sexually-antagonistic effect (for example, a mutation beneficial to males but detrimental to females). If this new male-beneficial/female-detrimental locus becomes closely linked to a male sex determination locus on a Y chromosome, selection will act to sequester the male-beneficial allele with the male-determination

locus through the suppression of recombination on the Y (Rice, 1987; Charlesworth *et al.*, 2005). Once the male-specific region of a Y chromosome has been established, other genes with sex-specific effects can transpose to the Y chromosome (Charlesworth *et al.*, 2005; Bachtrog, 2006).

In addition to accumulating genes with sex-specific effects, the male-specific region of the Y chromosome also begins to suffer some of the detrimental effects of the loss of recombination. These effects include the inability to repair mutations through recombination, so genes in the male-specific region of the Y gradually diverge in sequence from their X chromosome homologs (Bachtrog, 2006). In addition, large-scale rearrangements can occur, leading to deletions and inversions. Inversions themselves may help suppress recombination between the X and Y (Lahn and Page, 1999) and help expand the male-specific region if those inversions include sections of the pseudo-autosomal region (Charlesworth *et al.*, 2005). In addition, transposable elements often accumulate in the male-specific region of the Y chromosome (Liu *et al.*, 2004; Peichel *et al.*, 2004; Kondo *et al.*, 2006; Bachtrog *et al.*, 2008; Marais *et al.*, 2008). Possibly to silence these invasive mobile genetic elements, chromatin modifications can also transcriptionally silence sections of the male-specific region that have large numbers of invasive genetic elements (Bachtrog *et al.*, 2008). However, transposable elements can also be recruited to modify expression of a gene on the sex chromosomes, possibly affecting expression of the sex determination gene itself (Martin *et al.*, 2009; Herpin *et al.*, 2010).

As formerly-autosomal genes on the Y chromosome are lost due to mutation or deletion, a gene dosage imbalance between the sexes may develop for genes that have been lost on the Y and retained on the X. As Y degeneration spreads, the dosage imbalance may encompass most genes on the X chromosome. This dosage imbalance may persist with no selection for a mechanism to equalize X gene dosage between the sexes (dosage tolerance). However, the dosage imbalance may also select for the evolution of either a global or a gene-by-gene mechanism to equalize X gene dosage between the sexes (Charlesworth, 1978). In mammals, for example, this dosage compensation mechanism takes the form of global transcriptional inactivation of one X chromosome copy in females (Lucchesi *et al.*, 2005). In the

fruit fly *Drosophila melanogaster*, dosage compensation occurs via hypertranscription of the single X chromosome in males (Lucchesi *et al.*, 2005). However, there are several species with degenerate sex chromosomes that have no reported global dosage compensation mechanism, including the chicken, the trematode *Schistosoma mansoni*, the silkworm moth, and the threespine stickleback (*Gasterosteus aculeatus*) (Ellegren *et al.*, 2007; Leder *et al.*, 2010; Vicoso and Bachtrog, 2011; Walters and Hardcastle, 2011).

Degeneration of the male-specific region on the Y chromosome may begin rapidly following the cessation of recombination with the X chromosome. However, selection does likely preserve sections of the male-specific region. The sex determination gene, for example, must remain functional. In addition, sex chromosomes can fuse with autosomes, adding new sections to the male-specific region of the Y chromosome. These Y-autosome fusion events may even help bring new sexually-antagonistic loci onto the sex chromosomes (Charlesworth and Charlesworth, 1980).

Sex chromosomes are dynamic environments within the genome. The sex-specific regions of the Y or W can diverge rapidly between lineages. Thus, as populations diverge and speciation events occur, even closely related species can have sex chromosomes that, thought they were inherited from a common ancestor, differ in size, appearance, the size of the sex-specific region, and gene content (Paar *et al.*, 2011).

TRANSITIONS BETWEEN SEX DETERMINATION MECHANISMS

The Emergence of New Sex Determination Genes

New sex determination genes can evolve anywhere in the genome, regardless of the existence of a previous sex determination locus on a sex chromosome. The mechanisms by which new sex determination genes arise are not known. However, several theories point to sexually-antagonistic loci as a potential catalyst for the emergence of new sex determination genes. By these theories, an autosomal gene with fitness benefits favoring one sex over the other can select for a new sex determination gene arising closely linked to it (van Doorn

and Kirkpatrick, 2007). The new sex determination gene (and its closely-linked gene under sexually-antagonistic selection) would establish a new sex chromosome system, displacing the previous sex determination gene. The degree and direction of the initial sexually-antagonistic selection can even cause transitions between male and female heterogametic sex chromosome systems (van Doorn and Kirkpatrick, 2010).

We can only verify these theoretical models by identifying sex determination genes in a variety of closely-related species. However, few animal sex determination genes are known. Many of these genes are either transcription factors (such as the mammalian sex determination gene *SRY* and the bird sex determination gene *DMRT1*) or RNA splicing factors (such as *Sex-lethal* in *Drosophila melanogaster*) (Volff *et al.*, 2003; DiNapoli and Capel, 2008; Sekido and Lovell-Badge, 2009). Interestingly, the known vertebrate sex determination genes come from conserved genes and gene families of downstream sexual differentiation developmental pathways. These genes include the SOX family gene *SRY* on the Y chromosome of most mammalian species, and DM family genes such as *DMRT1* on the Z chromosome of birds, *DMY* on the Y chromosome of medaka, and *DM-W* on the W chromosome of the African clawed frog *Xenopus laevis* (Matsuda *et al.*, 2002; Nanda *et al.*, 2002; Volff *et al.*, 2003; Okada *et al.*, 2009; Smith *et al.*, 2009; Yoshimoto *et al.*, 2010).

Transitions Between Genetic and Environmental Mechanisms

The loss of a sex determination gene can lead to sex chromosome turnover, where the sex chromosomes themselves are no longer maintained by selection and ultimately lost. A new sex determination mechanism can assume its role at the top of the sex determination and sexual differentiation hierarchy. If this new mechanism is a sex determination gene, then a new sex chromosome system is established, as discussed previously (van Doorn and Kirkpatrick, 2007). But, this new mechanism need not be genetic. Transitions between different genetic and environmental mechanisms have been implied based on the distribution of sex determination mechanisms in different animal lineages. Among reptiles, for example, many

species have either sex chromosomes or an environmental sex determination mechanism such as temperature (Modi and Crews, 2005). In some lizard species, ZZ-ZW sex chromosomes and temperature-based sex determination coexist (Quinn *et al.*, 2007; Ezaz *et al.*, 2009). However, the selective pressures and mechanisms governing the transition between environmental and genetic sex determination mechanisms are poorly understood. In reptiles, recent theories have examined whether sexual selection and the presence of sexually dimorphic traits might predispose a species to developing a genetic or environmental sex determination mechanism (Valenzuela, 2009). It has been suggested that these selective pressures may trigger transitions between environmental and genetic sex determination (Valenzuela, 2009).

Several experimental approaches can help identify transitions between sex determination mechanisms. For example, if a species has a known sex determination gene, or genetic markers closely linked to the sex determination locus then, closely related species can be tested for the presence of those *SEX*-linked genetic markers or the sex determination gene (Kondo *et al.*, 2003; Matsuda *et al.*, 2003; Takehana *et al.*, 2007; Tanaka *et al.*, 2007). However, fewer methods exist to identify the transitions between genetic and environmental sex determination mechanisms. These transitions likely require knowledge of the ecology and behavior of a species to understand what types of environmental cues could supplant an established genetic sex determination mechanism.

The diversity of sex determination mechanisms and sex chromosome systems reported in the literature, including between closely related species, raises a number of questions concerning how they arise and evolve. Many of these questions are difficult to address in mammals due to the stability of the sex determination mechanism and the advanced stage of sex chromosome degeneration. To answer these questions, we must turn to model systems where closely related species have divergent sex determination and sex chromosome systems.

STICKLEBACK FISHES AS A MODEL SYSTEM FOR THE EVOLUTION OF SEX CHROMOSOMES

Sex Chromosome Diversity in the Sticklebacks

Sticklebacks are a family of small teleost fish, inhabiting freshwater and marine temperate habitats across the northern hemisphere (Wootton, 1976). All six species are classified in five genera within the family Gasterosteidae (Figure 1.1). For over a century, generations of ecologists and ethologists have studied stickleback species and populations, documenting numerous differences in behavioral and morphological traits both between and within species (Wootton, 1976; Bell and Foster, 1994). Historically, these ecological, morphological, and behavioral differences within and between stickleback species have been difficult to place in a larger evolutionary context due to disputed phylogenies within this family (Mattern, 2004). However, a recent phylogeny has clarified the evolutionary relationships among sticklebacks as this family has evolved over the past 40 million years (Kawahara *et al.*, 2009).

Stickleback evolution has not been confined to morphological and behavioral traits. Several studies have shown that sticklebacks differ in some aspects of genome organization. Karyotypes among stickleback species differ in both diploid number (2n) and chromosome morphology (Figure 1.1) (Chen and Reisman, 1970). This divergence is especially true for sex chromosomes in sticklebacks. At least five sex chromosome systems have been described in sticklebacks, and in four of these systems, genetic markers linked to a sex determination locus have been identified (*SEX*-linked markers; see Figure 1.1) (Chen and Reisman, 1970; Peichel *et al.*, 2004; Ocalewicz *et al.*, 2008; Kitano *et al.*, 2009; Ross *et al.*, 2009; Shapiro *et al.*, 2009). To better illustrate the evolutionary relationships among these sex chromosome systems, chromosome pairs in all stickleback species have been numbered relative to the genome of the threespine stickleback, *Gasterosteus aculeatus* (Figure 1.1; Figure 1.2) (Peichel *et al.*, 2001).

The three sex chromosome systems in the genus *Gasterosteus* likely arose from a common ancestral sex chromosome system (Figure 1.1). In most populations of *G. aculeatus* (Figure 1.2), chromosome 19 (Chr19) is an XX-XY sex

chromosome system (Peichel *et al.*, 2004; Ross and Peichel, 2008). This sex chromosome system is visible (heteromorphic), since the X and Y are morphologically distinguishable from one another (Ross and Peichel, 2008). At least 3 inversions and one large deletion have occurred on the male-specific region of the *G. aculeatus* Y chromosome since it stopped recombining with the X (Ross and Peichel, 2008). The male-specific region of the Y chromosome shows evidence for the invasion of transposable elements, and sequence divergence has proceeded to a point where distinct X and Y alleles for some genes can be detected (Withler *et al.*, 1986; Peichel *et al.*, 2004). However, *G. aculeatus* from the Sea of Japan have an $X_1X_1X_2X_2$-X_1X_2Y sex chromosome system, which likely arose from a fusion between the ancestral Y chromosome (Chr19) and a previously autosomal copy of Chr09 (Kitano *et al.*, 2009). A separate Y-autosome fusion between the ancestral Y and Chr12 likely gave rise to the $X_1X_1X_2X_2$-X_1X_2Y sex chromosome system found in the black-spotted stickleback, *G. wheatlandi* (Ross *et al.*, 2009).

In addition to the three sex chromosome systems in the genus *Gasterosteus*, two other species in the stickleback family have heteromorphic sex chromosomes. In ninespine sticklebacks, *Pungitius pungitius*, Chr12 is a XX-XY sex chromosome system (Ocalewicz *et al.*, 2008; Shapiro *et al.*, 2009). Cytogenetic evidence suggests that Chr12 has twice evolved as a sex chromosome system in this family; i.e. in the *P. pungitius* sex chromosome pair and the *G. wheatlandi* $X_1X_1X_2X_2$-X_1X_2Y sex chromosome system (Ross *et al.*, 2009). No heteromorphic sex chromosomes have been reported for the closest extant relative of *P. pungitius*, the brook stickleback (*Culaea inconstans*) (Figure 1.3). Fourspine sticklebacks (*Apeltes quadracus*) have a ZZ-ZW sex chromosome pair, but no genetic markers from this sex chromosome system have been reported (Figure 1.4) (Chen and Reisman, 1970; Ross *et al.*, 2009). Finally, no studies have yet searched for evidence of sex chromosomes in the fifteenspine stickleback, *Spinachia spinachia*.

Genetic and Genomic Studies of the Stickleback Sex Chromosome Systems

Though much data have been gathered regarding sex chromosome diversity in sticklebacks, very little is currently known about the gene content of these sex

chromosome systems. No sex determination gene has been found in sticklebacks. Thus, it is not known whether each sex chromosome system has a unique sex determination gene or if a sex determination locus has transposed to different sex chromosome pairs.

In addition to the sex chromosome diversity and rapid sex chromosome turnover in this family, a number of genetic and genomic tools have been developed for sticklebacks. These tools include the complete *G. aculeatus* female (XX) genome (BROAD S1 assembly, February 2006, available at http://www.ensembl.org/Gasterosteus_aculeatus/index.html), as well as the *G. aculeatus* bacterial artificial chromosome (BAC) libraries (Kingsley *et al.*, 2004). In addition, the complete sequence of the *G. aculeatus* Y chromosome should be available within the next year. Genome-wide microsatellite markers have been developed for both *G. aculeatus* and *P. pungitius* (Peichel *et al.*, 2001; Shapiro *et al.*, 2009). There are also protocols to introduce transgenic constructs into sticklebacks (Hoseman *et al.*, 2004; Chan *et al.*, 2010). Finally, a number of cytogenetic protocols, such as fluorescence *in situ* hybridization, have been adapted for sticklebacks (Ross and Peichel, 2008; Ross *et al.*, 2009).

GOALS OF THIS DISSERTATION

This dissertation summarizes several studies regarding the evolution of sex chromosome systems in the stickleback family, as well as the developmental and genetic regulation of sex determination and sexual differentiation in sticklebacks. These investigations were divided into two broad research questions, and the major findings of these studies are described in the next three chapters of this dissertation.

How Many Sex Chromosome Systems Have Evolved in the Stickleback Family?

The answer to this question is more than just a simple tally of different sex chromosome systems. Evidence so far suggests that at least two sex chromosome systems have arisen independently in this family: the Chr19-based sex chromosome systems of the *Gasterosteus* species, and the Chr12-based XX-XY sex chromosome

system of *P. pungitius* (Chen and Reisman, 1970; Peichel *et al.*, 2004; Ocalewicz *et al.*, 2008; Kitano *et al.*, 2009; Ross *et al.*, 2009; Shapiro *et al.*, 2009). However, there are many unanswered questions about sex chromosome systems in other stickleback lineages. These unresolved issues mask a true understanding of sex chromosome diversity in this family. For example, we do not know if the Chr12 or Chr19-based sex chromosome systems are found in other stickleback species. We also do not know if the *A. quadracus* ZZ-ZW sex chromosome system is derived from either the *P. pungitius* or *Gasterosteus* sex chromosome systems. Sex chromosomes are unknown, or have never been sought, in two stickleback species, and only a single survey has documented sex chromosome diversity between different populations of the same species (Kitano *et al.*, 2009).

Thus, to uncover new answers to this broad question, I designed experiments to identify the sex determination mechanisms and sex chromosome systems that have evolved in the stickleback family (Gasterosteidae). In Chapter 2, I describe a set of experiments to search for genetic evidence of sex chromosome systems in two stickleback species, the brook stickleback (*C. inconstans*) and the fourspine stickleback (*A. quadracus*) (Figure 1.3; Figure 1.4). Based on the preliminary experimental results described in Chapter 2, I designed and executed an additional study to search for the sex chromosome system in *A. quadracus* (Figure 1.4). I report the design and major findings of this study in Chapter 3.

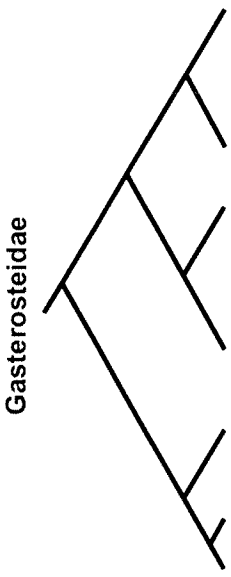
Do Sticklebacks Share the Same Sex Determination Gene?

By definition, sex chromosomes house sex determination genes. While we know that diverse sex chromosome systems have evolved in the stickleback family, we do not know if different stickleback species share the same sex determination gene. It is possible that each sex chromosome system in sticklebacks has its own sex determination gene. However, it is also possible that a single sex determination locus has transposed to different sex chromosome systems. To resolve this issue, we must identify the sex determination gene in one species with sex chromosomes, and see if this gene is present in other sex chromosome systems.

To address this question, I designed a screen to identify genes in the *G. aculeatus* sex determination pathway, particularly the master sex determination gene. These investigations were in the form of a comprehensive screen to identify genes that are differentially expressed between male and female *G. aculeatus* embryos and larvae. These differentially expressed genes include autosomal loci, as well as genes on the X and Y sex chromosomes of *G. aculeatus* (Figure 1.2). I report the major conclusions of this study in Chapter 4.

Chapter 5 summarizes the major findings of all of my studies. I also take the opportunity to suggest future experiments that will build on the results summarized in this dissertation. I sincerely hope my conclusions will spawn future studies of the evolution of sex chromosomes in this fascinating model system. Sticklebacks are a truly unique model through which we can investigate the early stages of sex chromosome degeneration and the enigmatic phenomenon of sex chromosome turnover.

Figure 1.1. A phylogeny of the stickleback fish family (Gasterosteidae).



Species	2n	System	Chr	References
Fourspine *Apeltes quadracus*	46	ZW	--	Chen and Reisman, 1970; Ross *et al.*, 2009; **this dissertation (chapters 2 and 3)**
Fifteenspine *Spinachia spinachia*	--	--	--	--
Brook *Culaea inconstans*	46	--	--	Ross *et al.*, 2009; **this dissertation (chapter 2)**
Ninespine *Pungitius pungitius*	42	XY	12	Ocalewicz *et al.*, 2008; Shapiro *et al.*, 2009; Ross *et al.*, 2009; Ocalewicz *et al.*, 2011
Black-spotted *Gasterosteus wheatlandi*	42	X_1X_2Y Male 2n = 41	12, 19	Ross *et al.*, 2009
Threespine (Sea of Japan) *Gasterosteus aculeatus*	42	X_1X_2Y Male 2n = 41	09, 19	Kitano *et al.*, 2009
Threespine *Gasterosteus aculeatus*	42	XY	19	Peichel *et al.*, 2004; Ross and Peichel, 2008; Ocalewicz *et al.*, 2011; **this dissertation (chapters 3 and 4)**

The stickleback phylogeny is based on Kawahara *et al.* (2009), and incorporates major findings regarding the evolution of sex chromosomes. Branch lengths are arbitrary. Listed are the common name and scientific name for each stickleback species, the diploid number, the heteromorphic sex chromosome system (if known), and the chromosome on which the sex determination gene is found (if known). Chromosomes in all stickleback species are numbered in reference to the genome of *G. aculeatus*.

Figure 1.2. A male threespine stickleback (*Gasterosteus aculeatus*).

From Lake Washington (Mercer Slough Nature Park, Bellevue, Washington, USA). Image provided courtesy of Shaugnessy McCann.

Figure 1.3. Brook sticklebacks (*Culaea inconstans*).

F1 progeny from a cross between a female from Fox Holes Lake (Northwest Territories, Canada) and a male from Pine Lake (Alberta, Canada). Image provided courtesy of Joseph Ross.

Figure 1.4. Fourspine sticklebacks (*Apeltes quadracus*).

Progeny of crosses between females and males from the West River (New Haven, Connecticut, USA).

CHAPTER TWO

The Search for Genetic Markers Linked to Sex Determination Loci in *Culaea inconstans* and *Apeltes quadracus*

SUMMARY

Previous studies have identified genetic markers linked to a sex determination locus in four sex chromosome systems among three different stickleback species. We searched for similar *SEX*-linked markers in two other stickleback species, *Culaea inconstans* and *Apeltes quadracus*, using two different techniques. Neither technique uncovered *SEX*-linked markers in either species, despite reports of a ZZ-ZW sex chromosome system in *A. quadracus*. However, we were able to show that the *A. quadracus* sex chromosome system evolved independently from the other known stickleback sex chromosome systems.

The experiments described in this chapter were executed by myself, Jessica Boland, and Catherine Peichel. Where applicable, I have used terms "I" and "we" to distinguish between experiments conducted solely by myself (such as amplified fragment length polymorphism genotyping) and experiments conducted by myself and my collaborators.

Portions of this chapter were published previously in:
Joseph A. Ross, James R. Urton, Jessica Boland, Michael D. Shapiro, and Catherine L. Peichel. 2009. "Turnover of Sex Chromosomes in the Stickleback Fishes (Gasterosteidae)" *PLoS Genetics* 5(2): e1000391.

INTRODUCTION

Genetic *SEX*-linked Markers in Sticklebacks

Four sex chromosome systems in the stickleback family were first identified by genetic markers linked to a sex determination locus (*SEX*-linked markers). Those four systems have been mapped relative to the genome of the threespine stickleback, *Gasterosteus aculeatus* (Figure 1.1). All of these sex chromosome systems were later confirmed by a cytogenetic technique, fluorescence *in situ* hybridization (FISH), using bacterial artificial chromosome (BAC) probes from *G. aculeatus* (Kingsley *et al.*, 2004; Peichel *et al.*, 2004; Ross and Peichel, 2008; Kitano *et al.*, 2009; Ross *et al.*, 2009; Shapiro *et al.*, 2009).

Lack of *SEX*-linked Markers in *C. inconstans*, *A. quadracus*, and *S. spinachia*

No publications report *SEX*-linked markers in the remaining species in the Gasterosteidae, *Apeltes quadracus*, *Culaea inconstans*, and *Spinachia spinachia*. There have been no investigations into the sex determination mechanism of *S. spinachia*, a stickleback species native to coastal regions of Europe. A single study reported no heteromorphic sex chromosomes in a *C. inconstans* population from New York, and a ZZ-ZW sex chromosome pair in *A. quadracus* (Chen and Reisman, 1970).

These reports raise a number of questions regarding the evolution of sex determination mechanisms. *C. inconstans* could have homomorphic sex chromosomes, or a sex determination mechanism that is not genetic. The *A. quadracus* ZZ-ZW sex chromosome pair could have arisen independently of the sex chromosome systems found in other stickleback species, or it could have evolved from a common ancestral sex chromosome system in this family. To address these questions, I sought to identify *SEX*-linked genetic markers in *C. inconstans* and *A. quadracus*. Parallel to this screen, my colleague Joseph Ross conducted a FISH screen in both *A. quadracus* and *C. inconstans* to investigate the possibility that sex chromosome systems in these species could be related to the sex chromosome systems in *Gasterosteus* species (Chr19) and *P. pungitius* (Chr12).

MATERIALS AND METHODS

Specimen Collection and Crosses

C. inconstans specimens were collected from Pine Lake (Wood Buffalo National Park, Alberta, Canada) and Fox Holes Lake (Northwest Territories, Canada) in June 2005 (permit WB05-1010). A single *C. inconstans* cross was generated using a female from Fox Holes Lake and a male from Pine Lake. *A. quadracus* specimens were collected from Pilgrim Lake (Cape Cod National Seashore, Massachusetts) in May 2004 (permit CACO-2005-SCI-0014), and a single cross was generated using a male and female. Progeny from all crosses were grown in 110L aquarium tanks (75cm length x 30cm depth x 46cm height). For both species, fish were kept in 0.35% seawater (3.5g/L Instant Ocean salt (Aquarium Systems, Mentor, Ohio, USA); 0.4mL/L $NaHCO_3$) at approximately 16°C in summer lighting conditions (16h light: 8h dark). Fish were fed brine shrimp nauplii twice daily. All animal procedures were approved by the Fred Hutchinson Cancer Research Center Institutional Animal Care and Use Committee (protocol 1575).

Sex Phenotyping

Fish were euthanized in 0.025% tricaine methanesulfonate (MS-222). The phenotypic sex of each individual was ascertained by direct examination of the gonads. The *C. inconstans* cross consisted of 16 females and 14 males. The *A. quadracus* cross consisted of 18 females and 10 males.

DNA Extraction

Genomic DNA from each individual was extracted as follows: a section of the caudal fin was removed and treated with 0.2mg Proteinase K (Invitrogen, Carlsbad, California, USA) for 14-16 hours at 55°C. Following an extraction with an equal volume of phenol:chloroform, genomic DNA was precipitated in ethanol and resuspended in 50µL Tris/EDTA (10mM Tris; 1mM EDTA) and stored at 4°C.

Microsatellite Genotyping

PCR-based genotyping of *G. aculeatus* and *P. pungitius* microsatellite markers was performed using genomic DNA from the parents and progeny of the *C. inconstans* and *A. quadracus* crosses. Genotyping was performed as previously described (Peichel *et al.*, 2001; Shapiro *et al.*, 2009) with the following exceptions. Reactions were run on an ABI 3100, and ABI GeneMapper 3.7 was used to analyze genotypes (Applied Biosystems, Carlsbad, CA, USA)

Amplified Fragment Length Polymorphism (AFLP) Genotyping

Amplified fragment length polymorphism (AFLP) screens (Vos *et al.*, 1995) were performed on parents and individuals from the *C. inconstans* and *A. quadracus* crosses. The AFLP Plant Mapping Protocol (Applied Biosystems) was used with the following alterations. Genomic DNA (1-2µg) was cut with 50 units of *Eco*RI (New England Biolabs, Ipswich, MA, USA) and 25 units of *Mse*I (New England Biolabs) with 10µg of BSA for 4 hours at 37°C. Reactions were then heated to 65°C for 20 minutes. Preselective amplifications were run on a Peltier Thermal Cycler-100 (PTC-100, MJ Research, Saint-Bruno-de-Montarville, Québec, Canada). For the AFLP selective amplifications, each of 8 *Eco*RI primer pairs (E-AAC, E-AAG, E-ACA, E-ACC, E-ACG, E-ACT, E-AGC, E-AGG) were paired with each of 6 *Mse*I primer pairs (M-CAA, M-CAC, M-CAG, M-CAT, M-CTA, M-CA), for a total of 48 primer pairs. *Eco*RI primers were labeled with 6-FAM. Selective amplifications were run on a PTC-100 (MJ Research). Reactions were run on an ABI 3100, and genotypes were analyzed with the "AFLP Default" method on ABI GeneMapper 3.7 (Applied Biosystems).

Linkage Analysis

Genetic linkage maps were generated using the default parameters in JoinMap3.0 (van Ooijen and Voorips, 2001). A Kruskal-Wallis test in MapQTL4.0 (van Ooijen *et al.*, 2002) was used to search for associations between marker genotype and sex phenotype.

RESULTS

SEX is Not Linked to Chr12 or Chr19 in *C. inconstans* or *A. quadracus* Crosses

We specifically tested whether genetic markers from Chr12 and Chr19 (which are *SEX*-linked in other species) are linked to a single *SEX*-determination locus governing female or male development in *C. inconstans* or *A. quadracus*. However, we found no associations between marker genotype and the phenotype of sex (Table 2.1). In the *C. inconstans* cross, a single Chr12 marker from *P. pungitius*, *Pun234*, was polymorphic, but not linked to *SEX*. In that same cross, four *G. aculeatus* and *P. pungitius* Chr19 makers were polymorphic, but not *SEX*-linked. Seven *G. aculeatus* and *P. pungitius* microsatellite markers from Chr12 were polymorphic in the *A. quadracus* cross, but none were *SEX*-linked. In addition, two Chr19 markers were not linked to sex in the *A. quadracus* cross.

No Microsatellite or AFLP Polymorphic Markers are *SEX*-Linked in *C. inconstans* or *A. quadracus* Crosses

To perform a genome-wide screen for markers that are *SEX*-linked in either *C. inconstans* or *A. quadracus*, we genotyped both crosses with all available *G. aculeatus* microsatellite markers (576 total) and all available *P. pungitius* microsatellite markers (162 total). However, many markers either did not yield a PCR product or were not polymorphic in either the *C. inconstans* cross or the *A. quadracus* cross (Table 2.2). In addition, among those markers that were polymorphic in these crosses, we found no evidence of any markers that were linked to the phenotype of sex in either the *C. inconstans* or *A. quadracus* cross (Table 2.2).

Those polymorphic microsatellite markers only represented 17 (of 21) *G. aculeatus* and 19 (of 30) *P. pungitius* linkage groups in the *C. inconstans* cross, and 15 *G. aculeatus* and 15 *P. pungitius* linkage groups in the *A. quadracus* cross. Since this survey of microsatellite markers did not cover all linkage groups in either species, I also used an amplified fragment length polymorphism (AFLP) approach to identify additional polymorphic markers in both species. I used 48 different AFLP primer combinations to identify 35 polymorphic loci in *C. inconstans* and 86

polymorphic loci in *A. quadracus*. However, none of these polymorphisms were linked to *SEX* in either species (Table 2.3).

DISCUSSION

C. inconstans Sex Determination Mechanism Evolved Independently of the *Gasterosteus* and *P. pungitius* Sex Chromosome Systems

My study did not uncover any markers that are linked to a locus controlling male or female sex determination in the *C. inconstans* cross. In parallel to this study, Joseph Ross conducted a survey of chromosome morphology in males and females from this same *C. inconstans* cross. Metaphase chromosome number between the sexes was identical, indicating the absence of the types of $X_1X_1X_2X_2$-X_1X_2Y sex chromosome systems found in *G. wheatlandi* and the *G. aculeatus* population from the Sea of Japan (Kitano *et al.*, 2009; Ross *et al.*, 2009). In addition, he uncovered no evidence for a heteromorphic sex chromosome pair in either males or females (Ross *et al.*, 2009). Thus, we find no evidence for a heteromorphic XX-XY or ZZ-ZW sex chromosome system in *C. inconstans*.

These data are consistent with Chen and Reisman (1970), who reported the absence of a heteromorphic sex chromosome pair in a different *C. inconstans* population (Cayuga Inlet, Ithaca, New York, USA). Thus, it is possible that *C. inconstans* could have a "cryptic" (homomorphic) sex chromosome system, in which the Y or W chromosome is morphologically identical to its X or Z counterpart in metaphase chromosome spreads. Homomorphic sex chromosome pairs may indicate that the non-recombining region on the Y or W is not large enough to alter chromosome morphology relative to the X or Z chromosome. Alternatively, the non-recombining region could be large, but sequence degeneration within the Y or W non-recombining region has not resulted in a divergence in metaphase chromosome morphology between the sex chromosomes. For example, if the non-recombining region on the Y or W is altered through the simultaneous deletion of native loci and the addition of transposable elements, then morphology of the Y or W may not change appreciably compared to the X or Z. There is evidence that this type of event may have occurred on the sex chromosomes of *G. aculeatus*. The X and Y

sex chromosomes of this species are of similar size in metaphase chromosome spreads, differing only in centromere position, despite the fact that at least one large deletion (encompassing 6Mb of X chromosome sequence) has occurred on the Y (Ross and Peichel, 2008). However, a limited analysis of approximately 0.5Mb of Y chromosome sequence has shown that transposable elements have invaded the non-recombining region of the Y chromosome (Peichel *et al.*, 2004). This addition of transposable elements to the *G. aculeatus* Y chromosome may explain why the X and Y chromosomes are of similar size in metaphase spreads. This example illustrates how *C. inconstans* could have a homomorphic sex chromosome system with a relatively large non-recombining region.

If *C. inconstans* has a genetic sex determination mechanism found on a homomorphic sex chromosome system, our current approach has not identified it. This could be due to a lack of genome-wide coverage of the markers. In our *C. inconstans* cross, we could not find polymorphic markers from 4 of 21 *G. aculeatus* linkage groups and 11 of 30 *P. pungitius* linkage groups. While none of an additional 35 polymorphic AFLP markers were linked to *SEX* in the *C. inconstans* cross, it is still likely that we lacked sufficient markers to survey all of the 23 chromosome pairs in *C. inconstans*. However, if *C. inconstans* has a sex chromosome system, its evolutionary origins are independent of the sex chromosome systems of *Gasterosteus* species and *P. pungitius*. Our results provide genetic evidence for this, since neither one Chr12 marker nor four Chr19 markers are linked to *SEX* in the *C. inconstans* cross. These data are consistent with FISH experiments demonstrating that *G. aculeatus* Chr12 and Chr19 BAC probes hybridize to identical locations on homomorphic chromosome pairs in both male and female *C. inconstans* chromosome spreads (Ross *et al.*, 2009). Thus, any sex chromosome system in *C. inconstans* likely arose independently of sex chromosome systems in *Gasterosteus* species and *P. pungitius*.

C. inconstans could also have a sex determination mechanism other than simple genetic sex determination (XX-XY or ZZ-ZW sex chromosomes). These mechanisms could be polygenic sex determination, environmental sex determination, or a polyfactorial signal of both genetic and environmental cues. *C.*

inconstans is a freshwater fish inhabiting diverse inland lake, stream, and river environments across the Interior Plans and Canadian Shield physiographic regions of North America (Wootton, 1976). Only a few of these populations have been examined for evidence of sex determination; it is also possible that different *C. inconstans* populations have evolved divergent sex determination mechanisms.

A remaining challenge is to identify the sex determination mechanism in this species. As environmental sex determination mechanisms in fish can range from temperature and photoperiod to social cues (Bull, 1983; Conover and Heins, 1987; Godwin *et al.*, 2003), it is impractical to propose screening even one *C. inconstans* population for possible environmental sex determination mechanisms. I believe that it is more beneficial to first rule out the possibility of homomorphic sex chromosomes and polygenic sex determination.

One approach to search for evidence of genetic sex determination sex is through the hormonal manipulation of the sex phenotype in individual fish, followed by an examination of the sex of their progeny. This approach employs experimental protocols established to control the phenotypic sex of individual fish, regardless of genotype (Iwamatsu, 1999; Hahlbeck *et al.*, 2004; Hamaguchi *et al.*, 2004). Recently-fertilized fish clutches are treated with hormones to override or supersede the natural sex determination mechanism. One group would be treated with androgens to cause the fish to develop as fertile phenotypic males, while the other group would be treated with estrogens to induce development as phenotypic females (Iwamatsu, 1999; Hamaguchi *et al.*, 2004). After these hormone-treated fish grow to reproductive maturity, we would mate phenotypic females to wild-type males and mate phenotypic males to wild-type females. This approach would identify a simple genetic sex determination system only if these matings generated some sons-only clutches (through the crossing of sex-reversed ZZ females to wild-type ZZ males) or daughters-only clutches (through the crossing of sex-reversed XX males to wild-type XX females) (Figure 2.1) (Iwamatsu, 1999; Hamaguchi *et al.*, 2004). Although these hormone-based sex-reversal and mating experiments would identify simple XX-XY or ZZ-ZW sex determination systems (Iwamatsu, 1999; Hamaguchi *et al.*, 2004), they would not reveal the degree of sex chromosome

degeneration, or whether additional genetic or environmental factors influence sex determination.

If hormonal studies suggested that there is a genetic sex determination mechanism, then genetic and genomic techniques could be used to identify the sex chromosome pair in these species. The sex chromosome systems of both *G. aculeatus* and *P. pungitius* were identified following the generation of genome-wide microsatellite markers for those species (Peichel *et al.*, 2004; Shapiro *et al.*, 2009). Thus, it is possible that the generation of sufficient genetic markers to cover all *C. inconstans* linkage groups would reveal the presence of sex chromosomes in this species. If the non-recombining region on the Y or W were sufficiently small, these genetic markers may still overlook a homomorphic sex chromosome system in this species. It is difficult to know what marker coverage is needed. In the medaka (*Oryzias latipes*), the male-specific region of the Y chromosome is less than 0.3Mb, yet sex-specific markers from this region have been reported (Matsuda *et al.*, 2002). However, multiple methods (including AFLP) used to screen over 16,000 polymorphic markers failed to identify any *SEX*-linked markers in the green-spotted pufferfish (*Tetraodon nigroviridis*) (Li *et al.*, 2002). Given that these negative results are difficult to interpret (i.e. is there no genetic sex determination mechanism, or limited coverage of the genome?), I believe such a search should only be undertaken once alternative methods (such as a hormone study) have revealed the existence of a sex chromosome system.

Cytogenetic techniques could also identify a sex chromosome pair. Comparative genomic hybridization (CGH) is an especially powerful technique which has been use to successfully identify sex chromosome pairs in reptiles (Ezaz *et al.*, 2005; Martinez *et al.*, 2008). In CGH, genomic DNA from males and females are each labeled with a different fluorescent probe, combined, and hybridized to male and female chromosome spreads. We could then scan female and male metaphase chromosome spreads for regions of differential fluorescent labeling, indicative of the non-recombining region of a hemizygous sex chromosome (Martinez *et al.*, 2008). However, it is possible that a small non-recombining region would not be detected by CGH. Of the techniques described here, the hormone-based sex reversal and

mating experiment has the greatest potential to reveal the existence of a genetic sex determination mechanism on a homomorphic sex chromosome pair in *C. inconstans*. Only if this technique discounted the possibility of sex chromosomes in this species should more complex polygenic, environmental, or polyfactorial sex determination mechanism be considered.

Evidence for the Independent Origin of the *A. quadracus* ZZ-ZW Sex Chromosome System

Chen and Reisman (1970) reported heteromorphic ZZ-ZW sex chromosomes in an *A. quadracus* population from Reid State Park in Maine. My colleague, Joseph Ross, found a heteromorphic ZZ-ZW sex chromosome pair in an *A. quadracus* population from Demarest Lloyd State Park and Pilgrim Lake in Massachusetts (Ross *et al.*, 2009). In our study, we have genetic evidence that the ZZ-ZW sex chromosome system in *A. quadracus* population from Pilgrim Lake is not related to the Chr19-based XX-XY system of *G. aculeatus* or the Chr12 XX-XY system of *P. pungitius*. Our evidence is in the form of Chr19 and Chr12 genetic markers from *G. aculeatus* and *P. pungitius*, none of which are *SEX*-linked in our *A. quadracus* cross from Massachusetts. In addition, Joseph Ross showed that *G. aculeatus* Chr12 and Chr19 BAC probes hybridize to homomorphic chromosome pairs in *A. quadracus* males and females (Ross *et al.*, 2009). In our study, no other polymorphic *G. aculeatus* markers, *P. pungitius* markers, or AFLP markers were linked to *SEX* in the *A. quadracus* cross.

As in *C. inconstans,* our screen most likely failed to identify *SEX*-linked genetic markers in *A. quadracus* due to the low coverage of polymorphic markers across the genome. Polymorphic markers represented only 15 of 21 *G. aculeatus* linkage groups and 15 of 30 *P. pungitius* linkage groups. The addition of 86 polymorphic AFLP markers probably still does not provide sufficient coverage of the 23 chromosome pairs in *A. quadracus* to identify the ZZ-ZW chromosome pair.

There are several approaches we could take to identify *SEX*-linked markers in *A. quadracus*. One approach would be to expand the coverage of polymorphic markers in our cross, coupled with a genetic survey of additional *A. quadracus*

crosses, to increase the chances of identifying markers from the ZZ-ZW pair. However, given that the extent of divergence between the Z and W has not been studied, we do not know the number of polymorphic markers and individuals needed to find markers form the non-recombining region of the sex chromosomes. Another method that would permit a thorough survey of each *A. quadracus* chromosome pair is FISH, which is possible due to the *G. aculeatus* BAC library (Kingsley *et al.*, 2004). FISH has already been used successfully to study the sex chromosomes in *G. aculeatus*, *G. wheatlandi*, and *P. pungitius* (Ross and Peichel, 2008; Kitano *et al.*, 2009; Ross *et al.*, 2009). This screen would permit us to test fluorescently-labeled genomic probes from each *G. aculeatus* chromosome pair on each *A. quadracus* chromosome pair and locate a set of probes that hybridize to the heteromorphic pair in female *A. quadracus* metaphase spreads. In Chapter 3, I report the results of such a screen on an *A. quadracus* population from Connecticut.

Table 2.1. *SEX*-linked *G. aculeatus* and *P. pungitius* microsatellite markers used for genotyping.

		G. aculeatus		*G. wheatlandi*		*P. pungitius*		*C. inconstans*		*A. quadracus*	
Marker	Chr	Position (Mb)	*SEX*-linked	Chr	*SEX*-linked	Chr	*SEX*-linked	PCR product	*SEX*-linked	PCR product	*SEX*-linked
Pun99	12	5.576	NT	12/19	yes	12	yes	FP	NT	yes	no
Stn327	12	5.800	no	12/19	yes	FP	NT	NP	NT	NP	NT
Pun7	12	8.475	no	NP	NT	12	yes	NP	NT	yes	no
Stn287	12	9.516	no	12/19	NT	12	yes	NP	NT	yes	no
Stn276	12	9.516	no	12/19	NT	12	yes	NP	NT	yes	no
Stn144	12	11.037	no	FP	NT	12	yes	NP	NT	yes	no
Stn142	12	12.635	no	12/19	yes	NP	NT	FP	NT	yes	no
Pun2	12	12.276	no	12/19	yes	12	yes	NP	NT	NP	NT
Pun234	12	15.613	NT	FP	NT	12	yes	yes	no	yes	no
Stn186	19	1.942	yes	NP	NT	19	no	yes	no	FP	NT
Pun117	19	6.325	yes	12/19	yes	19	no	FP	NT	yes	no
Stn235	19	7.396	yes	12/19	yes	NP	NT	FP	NT	FP	NT
Stn194	19	11.787	yes	12/19	yes	19	no	NP	NT	yes	no
Pun268	19	13.170	FP	FP	NT	19	no	yes	no	FP	NT
Stn284	19	13.658	yes	12/19	yes	NP	NT	NP	NT	FP	NT
Stn168	19	13.736	NP	NP	NT	19	no	yes	no	FP	NT
Cyp19b	19	16.671	yes	12/19	yes	19	no	yes	no	FP	NT

For each marker, the *G. aculeatus* chromosome (Chr) designation and position in the *G. aculeatus* genome assembly (Broad assembly S1, http://www.ensembl.org/Gasterosteus_aculeatus/index.html) is shown. If a marker was also mapped in *G. wheatlandi* or *P. pungitius*, the Chr designation in that species is also indicated (Ross *et al.*, 2009). FP ("Failed PCR") indicates that PCR product was not obtained for a marker in that species. NP ("Not polymorphic") indicates that the PCR product obtained was not polymorphic in the cross for that species. Markers are labeled NT ("Not tested") if sex linkage for a cross could not be tested.

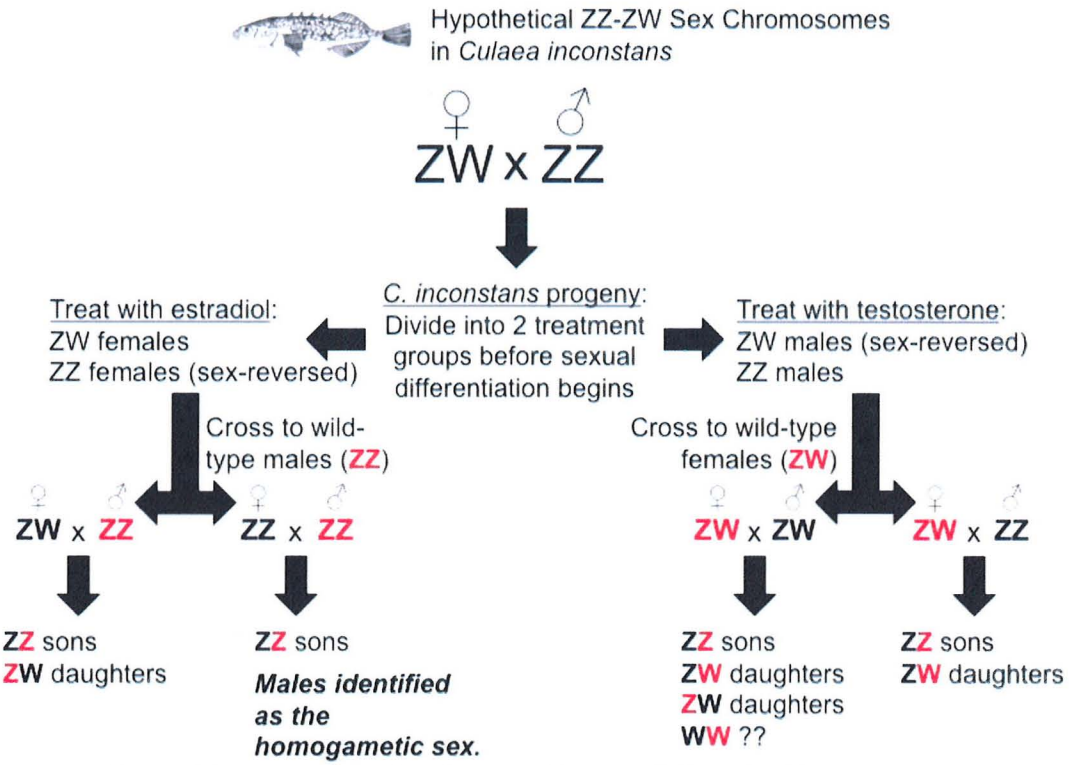
Table 2.2. Genome-wide microsatellite markers from *G. aculeatus* and *P. pungitius* used for genotyping *C. inconstans* and *A. quadracus* crosses.

Marker source	Total	*C. inconstans*			*A. quadracus*		
		Yielded PCR product	Polymorphic	*SEX*-linked	Yielded PCR product	Polymorphic	*SEX*-linked
G. aculeatus	576	225 (39.1%)	66 (11.5%)	0 (0%)	135 (23.4%)	47 (8.2%)	0 (0%)
P. pungitius	162	86 (53.1%)	43 (26.5%)	0 (0%)	51 (31.5%)	26 (16.1%)	0 (0%)

Table 2.3. AFLP-generated bands in *C. inconstans* and *A. quadracus*.

	C. inconstans	*A. quadracus*
Total number of AFLP bands generated	1,927	1,454
Polymorphic	35	86
SEX-linked	0	0

Figure 2.1. An overview of a hormone-based sex-reversal and mating experiment.



This approach could uncover a simple genetic sex determination or sex chromosome system (XX-XY or ZZ-ZW) in *C. inconstans*. The example assumes *C. inconstans* has a ZZ-ZW sex chromosome system. If *C. inconstans* had a XX-XY sex chromosome system, then the mating of a sex-reversed female to a wild-type female would yield all-female daughters, revealing females as the homogametic (XX) sex. If *C. inconstans* did not have a sex chromosome system, then no pairing of a hormone-treated fish to a wild-type fish should produce only progeny of a single sex.

CHAPTER THREE

Mechanisms for Chromosome Evolution and Sex Chromosome Diversity in *Apeltes quadracus*

SUMMARY

The fourspine stickleback, *Apeltes quadracus*, has a heteromorphic ZZ-ZW sex chromosome system. However, as shown in Chapter 2, I was unsuccessful in locating genetic markers linked to a sex determination locus in this species. I turned to fluorescence *in situ* hybridization (FISH) as an alternative mechanism to identify the *A. quadracus* ZZ-ZW pair. By FISH, I identified major differences in karyotype between *A. quadracus* and the threespine stickleback, *Gasterosteus aculeatus*. These differences likely reflect the types of chromosome rearrangements that have occurred in these species since they diverged from a common ancestor. I discovered that rearrangements such as centric fissions, Robertsonian fusions, and inversions occurred on approximately half of the chromosome pairs in the karyotypes of both species. In addition, I discovered that the *A. quadracus* population used for this screen lacks the heteromorphic ZZ-ZW sex chromosome pair found in other populations. This population may have a homomorphic ZZ-ZW sex chromosome pair related to the ZZ-ZW pair in other populations. Alternatively, a new sex chromosome pair or other sex determination mechanism may have evolved in this population. Thus, sex chromosomes in *A. quadracus*, and the stickleback family, may be more diverse than previously thought.

I performed and analyzed all experiments described in this chapter. Shaugnessy McCann assisted in fish collection and the preparation of chromosome spreads for FISH. Where applicable, I use the plural pronoun "we" to credit his contributions.

Portions of this chapter have been accepted for publication:
James R. Urton, Shaugnessy R. McCann, and Catherine L. Peichel. "Karyotype Differentiation Between Two Stickleback Species (Gasterosteidae)." *Cytogenetic and Genome Research.* In press.

INTRODUCTION

Sex Chromosome Diversity in the Sticklebacks

At least five sex chromosome systems have been identified in the stickleback family, and four have been mapped relative to the genome of the threespine stickleback, *Gasterosteus aculeatus* (Figure 1.1). Only one of these species, *Apeltes quadracus*, has ZZ-ZW sex chromosomes (Chen and Reisman, 1970; Ross *et al.*, 2009). In Chapter 2, I presented evidence that the ZZ-ZW sex chromosome pair of *A. quadracus* is not related to the other known stickleback sex chromosome systems (Ross *et al.*, 2009). However, our attempts to uncover genetic markers linked to a sex determination locus in *A. quadracus* were not successful.

FISH as a Tool for Studying Chromosome Rearrangements

To identify the *A. quadracus* sex chromosome pair relative to the *G. aculeatus* genome, I turned to the cytogenetic technique fluorescence *in situ* hybridization (FISH). FISH has been adapted as a tool for studying the sex chromosome systems of sticklebacks, making use of the bacterial artificial chromosome (BAC) libraries for *G. aculeatus* (Kingsley *et al.*, 2004). My colleague Joseph Ross has used FISH to map the divergence between the *G. aculeatus* X and Y sex chromosomes (Ross and Peichel, 2008), confirm two Y-autosome fusion events in sticklebacks (Kitano *et al.*, 2009; Ross *et al.*, 2009), and show that Chr12 and Chr19 BAC probes do not hybridize to heteromorphic chromosome pairs in *C. inconstans* or *A. quadracus* (Ross *et al.*, 2009).

FISH has also been widely used to study chromosome rearrangements in other species. Previous surveys of chromosome number and morphology in the stickleback family revealed diverse karyotypes among species (Chen and Reisman, 1970; Ocalewicz *et al.*, 2011), as well as within a single species (Kitano *et al.*, 2009). The phylogenetic relationships between species in the stickleback family have previously been disputed (Mattern, 2004). However, a recent comprehensive phylogeny has resolved the evolutionary relationship among the sticklebacks (Kawahara *et al.*, 2009). Mapping available karyotype data onto this phylogeny suggests a more rapid divergence in stickleback karyotype than previously

appreciated, since even closely related stickleback species have diverged in diploid chromosome number, karyotype, and sex chromosome system (Figure 1.1).

To explore the rapid karyotypic evolution of the stickleback family in more detail, I designed a comprehensive FISH survey to identify the major genomic rearrangements that have occurred between *G. aculeatus* and *A. quadracus*. These species have different diploid chromosome numbers, divergent karyotypes, and independently-evolved sex chromosome systems (Figure 1.1) (Chen and Reisman, 1970; Peichel *et al.*, 2004; Ross *et al.*, 2009). Thus, I designed my survey with three goals in mind: to identify the metaphase chromosome morphology of each *G. aculeatus* chromosome pair, to reveal the genomic rearrangements that generated the diverse karyotypes of these species, and to reveal the identity of the *A. quadracus* ZZ-ZW sex chromosome pair.

MATERIALS AND METHODS

Specimen Collection and Crosses

Male and female *A. quadracus* individuals were collected from the West River Memorial Park in New Haven, Connecticut in May 2009. Males and females bred naturally in the laboratory in July and August 2009. We grew progeny from these matings together in 474L aquarium tanks (183cm length x 46cm depth x 65cm height). In April and June 2010, we collected male *G. aculeatus* from two locations in Lake Washington: Union Bay in Seattle, Washington and the Mercer Slough Nature Park in Bellevue, Washington (Washington permit 10-049). We housed these fish in 110L aquarium tanks (75cm length x 30cm depth x 46cm height). For both species, we kept fish in summer lighting conditions (16h light: 8h dark) at approximately 16°C in 0.35% saltwater (3.5g/L Instant Ocean salt (Aquarium Systems, Mentor, Ohio, USA.); 0.4mL/L $NaHCO_3$). We fed the fish live brine shrimp nauplii twice daily. All procedures were approved by the Fred Hutchinson Cancer Research Center Institutional Animal Care and Use Committee (protocol 1575).

Selection of FISH Probes

I used *G. aculeatus* BAC clones from the CHORI-215 library (Kingsley *et al.*, 2004) as probes for FISH (Table 3.1). To identify BACs, I extracted 200kbp of genomic sequence from the end of each chromosome assembly (BROAD S1 assembly, February 2006, available at http://www.ensembl.org/Gasterosteus_aculeatus/index.html) and used this sequence in a BLAST search of the publicly available BAC end sequences (Kingsley and Peichel, 2007). If both the T7 and SP6 reads of the BAC clones aligned within the 200kbp region, I used the end sequences in a BLAT search against the *G. aculeatus* genome. If the paired end sequences of the clone aligned to the desired chromosome assembly in opposing orientation, with expected separation based on the average size of a CHORI-215 BAC insert (Kingsley *et al.*, 2004), I then used the clone as a FISH probe.

For 18 of 21 *G. aculeatus* chromosomes, the BAC clones aligned to a region within 2.6Mbp of the predicted ends of that chromosome assembly and hybridized to the ends of chromosomes in *G. aculeatus* male metaphase spreads. For three chromosomes, one BAC clone selected by this method did not hybridize to the end of a chromosome pair in *G. aculeatus* male metaphase spreads. In those three cases, I screened additional 200kbp regions of those chromosome assemblies until I identified BAC clones that met the above selection criteria and hybridized to the ends of a chromosome pair in male *G. aculeatus* metaphase spreads. For the chromosomes involved (Chr16, Chr19, and Chr20), errors in the *G. aculeatus* genome assembly, as previously reported for Chr19 (Ross and Peichel, 2008), or population-specific chromosome rearrangements could account for these observations.

Cytogenetic Analysis

We prepared metaphase spreads as described previously (Ross and Peichel, 2008), with the following modifications. We performed intraperitoneal injections of 5μL (*A. quadracus*) or 10μL (*G. aculeatus*) of 1% colchicine in phosphate-buffered saline into adult fish and incubated them for 12-14 hours in an aquarium. After

euthanizing the fish in 0.025% tricaine methanesulfonate (MS-222), we determined the sex of individual fish by examination of the gonads. We pooled spleen tissue from several individuals (Table 3.2) to make metaphase spreads of males and females from both species.

I conducted FISH as previously described (Ross and Peichel, 2008), with the following modifications. I extracted and purified BAC DNA using an AutoGen 740 automated system (AutoGen, Holliston, Massachusetts, USA) and quantified BAC DNA by gel electrophoresis. Using the Vysis nick translation kit (Abbott Labs, North Chicago, Illinois, USA), I labeled 1μg of each BAC clone with either ChromaTide Alexa Fluor 488-5-dUTP or 568-5-dUTP (Invitrogen, Carlsbad, California, USA). Two hundred nanograms of the labeled clones were ethanol precipitated with 10μg salmon sperm DNA and 10μg genomic DNA from the species used in the experiment. I allowed hybridization to proceed for 1-2 nights at 37°C. I viewed washed slides using the 100X objective on a Nikon Eclipse 80i microscope (Nikon, Shinjuku, Tokyo, Japan) with an automated filter turret using Chroma filters 31000v2 (DAPI), 41001 (FITC), and 41004 (Texas Red) (Chroma, Bellows Falls, Vermont, USA). I captured images using a Photometrics Coolsnap ES2 camera (Photometrics, Tucson, Arizona, USA) and used NIS Elements imaging software (BR 3.00, SP7, Hotfix8, Build 548, Nikon, Shinjuku, Tokyo, Japan) to pseudocolor those images grey (DAPI), green (Alexa 488), and purple (Alexa 568).

Chromosome Morphology

For *G. aculeatus* and *A. quadracus*, I used images from FISH experiments (Figure 3.1; Figure 3.2; data not shown) to assess the morphology of each chromosome. For each chromosome pair, I used Adobe Illustrator software (Adobe Systems, San Jose, California, USA) to make triplicate measurements of chromosome arms for both homologs in a FISH metaphase spread. I used these triplicate chromosome arm measurements to calculate the mean length of each chromosome arm for each homolog, and used the mean length of each chromosome arm to calculate the arm length ratio (long arm: short arm). I then used the arm length ratio for both homologs to calculate the mean arm length ratio for that

homologous pair. I repeated these measurements with a second FISH metaphase spread and averaged the chromosome arm length ratios for both chromosome spreads. I classified chromosomes according to this average arm length ratio (long arm: short arm). Chromosomes were classified as metacentric (1.0-1.7), submetacentric (1.7-3.0), acrocentric (3.0-7.0), or telocentric (7.0 or greater), following Levan *et al.* (1964) and Schrader *et al.* (1997) (Table 3.3). The only exception to this method is *G. aculeatus* Chr19, which is a heteromorphic XX-XY sex chromosome pair (Peichel *et al.*, 2004; Ross and Peichel, 2008). For Chr19, I measured the arm lengths of the single X chromosome from three male *G. aculeatus* FISH images to calculate the mean arm length ratio for this chromosome, and I did not measure the Y chromosome (Figure 3.2; data not shown). For all FISH experiments and karyotype analyses, I examined multiple metaphase spreads from multiple pools of individual fish (Table 3.2).

For chromosomes classified as "telocentric" in both species, I found variation in chromosome arm measurements both between homologs in the same metaphase spread and between metaphase spreads (data not shown). In addition, for Chr08 and Chr18 in *G. aculeatus* and Chr06 in *A. quadracus*, chromosome pairs measured as "acrocentric" in one FISH metaphase spread and "telocentric" in the other. This variation, both within "telocentric" chromosomes and between chromosome pairs in different metaphase spreads, is likely due to the small size of the short arms in these chromosomes, which makes accurate measurements dependent on the morphology and resolution of individual metaphase spreads. Given this variation in chromosome arm measurements within and between "acrocentric" and "telocentric" chromosomes, I have chosen to combine "acrocentric" and "telocentric" chromosomes as a single chromosome morphology in my discussion, particularly in reference to the transitions between different chromosome morphologies in *G. aculeatus* and *A. quadracus*.

Finally, I calculated the *nombre fondamental* (major chromosome arm number; NF) for both sexes of *G. aculeatus* and *A. quadracus* (Matthey, 1949; Klinkhardt, 1998). For calculating NF, I followed White (1978) and Klinkhardt (1998) and considered that metacentric and submetacentric chromosomes each have two

major chromosome arms, while acrocentric and telocentric chromosomes each have one major chromosome arm.

RESULTS

FISH Screen Identifies the Chromosome Morphology of Each *G. aculeatus* Chromosome Pair

The *G. aculeatus* male karyotype consists of 42 chromosomes: 8 autosomes are metacentric, 6 are submetacentric, 16 are acrocentric, 10 are telocentric, and there is one submetacentric X and one metacentric Y (Table 3.4) (Ross and Peichel, 2008). The relatively high NF (58) for *G. aculeatus* compared to its diploid number (2n = 42) reflects a relatively large number of metacentric and submetacentric chromosomes in the karyotype. I used *G. aculeatus* BAC clones from both ends of each chromosome assembly (Table 3.1) to identify the major morphological features of each *G. aculeatus* chromosome. The BAC clones hybridize by FISH to the ends of each chromosome (Figure 3.1; Figure 3.2). By combining FISH with measurements of chromosome arm length, I assigned each *G. aculeatus* chromosome pair to a morphological class (Table 3.3).

FISH Screen Identifies the Major Differences between the *G. aculeatus* and *A. quadracus* Karyotypes

I used the same BAC clones selected for the *G. aculeatus* male FISH analysis in a similar survey of the chromosomes in *A. quadracus* females and males in order to identify the major differences in karyotype between these species (Figure 3.1; Figure 3.2). The *A. quadracus* karyotype consists of 46 chromosomes, of which 6 chromosomes are metacentric, 4 are submetacentric, 32 are acrocentric, and 4 are telocentric (Figure 3.3; Table 3.4). The dominance of acrocentric and telocentric chromosomes in the *A. quadracus* karyotype (78%) is reflected in a lower ratio of NF to diploid chromosome number; NF:2n is 56:46 for *A. quadracus* while it is 58:42 for *G. aculeatus* (Table 3.4).

Two of the larger chromosome pairs in the *G. aculeatus* karyotype, Chr04 (submetacentric) and Chr07 (metacentric), each correspond to two pairs of relatively

small acrocentric chromosomes in *A. quadracus*, accounting for the higher diploid chromosome number in the latter species (Figure 3.1; Table 3.3). Eight chromosome pairs appear to have undergone inversion in the time since *G. aculeatus* and *A. quadracus* diverged; six are pericentric inversions, which encompass the centromere, and two are paracentric inversions, which do not encompass the centromere (Figure 3.1). Five of the six pericentric inversions have resulted in visibly different morphology for the orthologous *G. aculeatus* and *A. quadracus* chromosomes. In three of these cases, metacentric or submetacentric *G. aculeatus* chromosome pairs correspond to acrocentric chromosome pairs in *A. quadracus*; in the other two cases, telocentric *G. aculeatus* chromosome pairs correspond to metacentric chromosome pairs in *A. quadracus* (Figure 3.1; Table 3.3). Eleven *G. aculeatus* chromosomes have identical hybridization patterns and morphologies in both *G. aculeatus* and *A. quadracus* (Figure 3.2).

A. quadracus Males and Females from Connecticut Do Not Have Heteromorphic Sex Chromosomes

A heteromorphic ZZ-ZW sex chromosome pair has been described in *A. quadracus* populations from Maine (Chen and Reisman, 1970) and Massachusetts (Ross *et al.*, 2009). However, I found no evidence for a heteromorphic sex chromosome pair in *A. quadracus* male or female progeny from specimens collected in the West River in New Haven, Connecticut (Figure 3.3). In addition, hybridization patterns were identical in metaphase spreads from male and female *A. quadracus* for all FISH probes used in this study (data not shown).

DISCUSSION

Diversity of the *A. quadracus* ZZ-ZW Sex Chromosome Pair

I undertook a comparative FISH study to identify the major genomic rearrangements that have occurred in lineages separating two stickleback species, *A. quadracus* and *G. aculeatus*. These species are intriguing subjects for such a study, not only because of reported differences in karyotype, but also due to the divergence of sex chromosome systems (Chen and Reisman, 1970; Ross *et al.*,

2009). In *G. aculeatus*, Chr19 is an XX-XY sex chromosome system (Peichel *et al.*, 2004; Ross and Peichel, 2008). Two previous studies have shown that *A. quadracus* populations from Maine and Massachusetts have a ZZ-ZW sex chromosome pair (Chen and Reisman, 1970; Ross *et al.*, 2009). As shown in Chapter 2, genetic markers from Chr19 are not linked to the sex determination locus in an *A. quadracus* family from Massachusetts (Ross *et al.*, 2009), and we found no genetic markers linked to a sex determination locus in *A. quadracus*.

In *A. quadracus* from Connecticut, I find that neither females nor males have a heteromorphic chromosome pair indicative of either a ZZ-ZW or XX-XY sex chromosome system. Two possible explanations could account for the differences in sex chromosomes between the Connecticut population of *A. quadracus* and the Maine and Massachusetts populations. First, *A. quadracus* from Connecticut could have a cryptic (homomorphic) ZZ-ZW sex chromosome pair that is homologous to the ZZ-ZW pair in the Massachusetts and Maine populations. The W chromosome in the Maine and Massachusetts populations is large and acrocentric, while the Z chromosome is also acrocentric, but has a smaller long arm than the W chromosome (Chen and Reisman, 1970; Ross *et al.*, 2009). The larger size of the W chromosome relative to the Z chromosome in those populations could be due to an increase in the amount of heterochromatin on the W chromosome relative to the Z chromosome following the cessation of recombination with the Z chromosome. If negligible amounts of heterochromatin were found on the *A. quadracus* W chromosome in the Connecticut population, then the Z and W chromosomes in this population would not be morphologically distinguishable. Such a population-specific difference in the amount of heterochromatin or repetitive DNA sequences present on the sex chromosomes has been reported in the guppy *Poecilia reticulata* (Hornaday *et al.*, 1994), the platyfish *Xiphophorus maculatus* (Nanda *et al.*, 2000), and the glass knifefish *Eigenmannia virescens* (de Almeida-Toledo *et al.*, 2001). Differences in the amount of heterochromatin on the sex chromosomes have also been reported between closely related species of salmonids (Moran *et al.*, 1996) and South American catfishes (Andreata *et al.*, 1992).

Second, *A. quadracus* from Connecticut could have a different sex determination mechanism than the Massachusetts and Maine ZZ-ZW sex chromosome system. This mechanism could be genetic, environmental, or a combination of genetic and environmental signals. Extreme variation in the mechanism of sex determination within a single species is not without precedent. In the Atlantic silverside, *Menidia menidia,* both environmental and genetic sex determination mechanisms are present and vary by latitude (Conover and Heins, 1987). Although we currently have little data on the ecology or population structure of *A. quadracus*, it would be interesting to determine whether variation in the sex determination mechanisms is associated with any ecological or geographical factors in this species.

Even though *A. quadracus* individuals from Connecticut lack a heteromorphic ZZ-ZW sex chromosome system, my comparative FISH study could help to identify the ZZ-ZW sex chromosome system in *A. quadracus*. The ZZ-ZW pair in both Massachusetts and Maine populations of *A. quadracus* is one of the larger chromosome pairs in the karyotype (Chen and Reisman, 1970; Ross *et al.*, 2009). In addition, the karyotypes of the Connecticut and Massachusetts *A. quadracus* are similar, differing only in the presence of the large W chromosome in *A. quadracus* females from the Massachusetts population (Ross *et al.*, 2009). Thus, the heteromorphic ZZ-ZW pair in the Massachusetts *A. quadracus* karyotype could correspond to one of the larger chromosome pairs in the Connecticut *A. quadracus* karyotype. Two of these larger chromosomes, Chr19 and Chr12, have previously been ruled out as the Massachusetts *A. quadracus* ZZ-ZW pair (Chapter 2) (Ross *et al.*, 2009). However, the Connecticut *A. quadracus* karyotype contains several other large chromosomes, including Chr01, Chr02, Chr09, Chr13, and Chr20. These chromosomes should be targets for further studies of the ZZ-ZW sex chromosome pair in *A. quadracus*.

Population differences in sex chromosomes and sex determination mechanisms are also found in other species of fish. Studies have documented sex chromosome diversity within species of poeciliid fish (Nanda *et al.*, 2000; Volff *et al.*, 2003; Schultheis *et al.*, 2009) and within several species of neotropical fish,

including *E. virescens* (de Almeida-Toledo *et al.*, 2002; Henning *et al.*, 2011), *Erythrinus erythrinus* (Bertollo *et al.*, 2004; Cioffi *et al.*, 2010), and *Hoplias malabaricus* (Bertollo *et al.*, 2000; Cioffi and Bertollo, 2010). Sex chromosome systems also differ among populations of other stickleback species. *G. aculeatus* from the Sea of Japan have a $X_1X_1X_2X_2$-X_1X_2Y sex chromosome system, which likely arose from the ancestral XX-XY sex chromosome system found in other *G. aculeatus* populations (Kitano *et al.*, 2009). Future studies should reveal whether most *A. quadracus* populations have ZZ-ZW sex chromosomes, or if this species has a diverse array of sex determination mechanisms.

Mechanisms Responsible for the Differences in Karyotype between *G. aculeatus* and *A. quadracus*

Although *G. aculeatus* and *A. quadracus* diverged approximately 40 million years ago, my comparative FISH study has uncovered genomic rearrangements encompassing nearly half of the 21 *G. aculeatus* chromosome pairs. These rearrangements account for the two major karyotypic differences between these species: *A. quadracus* has a larger diploid number, coupled with a lower number of metacentric and submetacentric chromosomes, when compared to *G. aculeatus* (Table 3.4).

The larger diploid number (46) in *A. quadracus* is explained by my finding that two *G. aculeatus* chromosome pairs (Chr04 and Chr07) correspond to four *A. quadracus* chromosome pairs. Examination of other stickleback species does not reveal whether the *G. aculeatus* or *A. quadracus* diploid number is more similar to the ancestral state in this family (Figure 1.1). The closest extant relative of *G. aculeatus* also has a diploid chromosome number of 42, as does a more distant relative, the ninespine stickleback, *P. pungitius* (Ocalewicz *et al.*, 2008; Ross *et al.*, 2009; Shapiro *et al.*, 2009; Ocalewicz *et al.*, 2011). However, many differences in chromosome morphology exist between the *Gasterosteus* species and *P. pungitius* (Ocalewicz *et al.*, 2008; Ross *et al.*, 2009; Ocalewicz *et al.*, 2011). A recent report shows that *P. pungitius* has a higher number of metacentric and submetacentric chromosomes than *G. aculeatus* (Ocalewicz *et al.*, 2011). However, since that study

did not compare specific chromosome pairs between species, it is not known if the chromosome pairs that differ in morphology between *G. aculeatus* and *P. pungitius* are identical to the chromosome pairs that differ between *G. aculeatus* and *A. quadracus* in my study. The closest extant relative of *P. pungitius*, the brook stickleback (*Culaea inconstans*), has the same diploid number as *A. quadracus*, but these two species also differ in karyotype (Chen and Reisman, 1970; Ross *et al.*, 2009). No published reports document the diploid chromosome number and chromosome morphology of the fifteenspine stickleback (*Spinachia spinachia*), which is the closest extant relative of *A. quadracus* (Kawahara *et al.*, 2009). Thus, we lack comprehensive karyotype data from a sufficient number of stickleback species and populations to know whether the karyotype of *G. aculeatus* or *A. quadracus* is the more ancestral for this family.

Though the ancestral stickleback karyotype is unknown, my FISH screen has clarified some of the mechanisms that have shaped the *A. quadracus* and *G. aculeatus* karyotypes since they diverged. Robertsonian fusions are the joining of two acrocentric chromosomes at their centromeres (generating a single large metacentric or submetacentric chromosome), while centric fission (or dissociation) is the splitting of a single metacentric or submetacentric chromosome into two acrocentric chromosomes (White, 1978; Klinkhardt, 1998). Robertsonian fusions and centric fissions change the diploid chromosome number (2n) without affecting the chromosome arm number (NF) (Klinkhardt, 1998). While metacentric and submetacentric chromosomes make up less than one quarter of the *A. quadracus* karyotype, they comprise more than one third of the *G. aculeatus* karyotype (Table 3.4). One factor that contributes to this difference is the separation of *G. aculeatus* Chr04 and Chr07 in the *A. quadracus* karyotype (Figure 3.1; Table 3.3). The differential state of both Chr04 and Chr07 in *G. aculeatus* and *A. quadracus*, and the fact that they are both acrocentric in *A. quadracus*, suggest that Robertsonian fusions and/or centric fissions have played roles in the evolution of stickleback karyotypes. These fusions may leave signatures in the form of interstitial telomeric sequences (ITSs) in the species with the derived chromosome state. A recent study of a *G. aculeatus* population from Poland failed to detect any ITSs by FISH;

however, it is possible that ITSs were eliminated, or remain in a copy number that is below the detection threshold for FISH (Ocalewicz *et al.*, 2011). Thus, the ancestral state of Chr04 and Chr07 in the stickleback family is still unknown, and proper determination of the ancestral state will clarify whether fusions or fissions are responsible for these differences in karyotype.

I also uncovered evidence for inversions involving at least eight chromosomes, based on differences between relative centromere position and/or BAC probe hybridization positions between *G. aculeatus* and *A. quadracus*. Six of these inversion events encompass the centromere, and five of those inversions alter the gross chromosome morphology between *G. aculeatus* and *A. quadracus*. In particular, three of these five inversions are on chromosomes that are metacentric or submetacentric in *G. aculeatus* and acrocentric or telocentric in *A. quadracus*, which is one factor that accounts for the lower NF in this species compared to *G. aculeatus* (Table 3.4). Future studies of chromosome morphology and relative probe hybridization positions in other stickleback species could further illuminate the role of pericentric inversions in the evolution of karyotypes across this family.

Possible Role of Chromosome Rearrangements in Adaptation and Speciation

My observations indicate that the major karyotypic differences between *G. aculeatus* and *A. quadracus* can be explained by chromosome inversion and either Robertsonian fusions or centric fissions. Such chromosome rearrangements might have been adaptive at some point during the evolution of these species. It has been suggested that chromosome rearrangements might play a role in both adaptation and speciation across diverse lineages (White, 1978; King, 1993; Ayala and Coluzzi, 2005; Hoffmann and Rieseberg, 2008).

My study shows that variation in Robertsonian fusion and centric fission is found between closely related species, which has also been reported for a number of mammalian species, including rodents, primates, deer, sheep, pigs, and horses (White, 1978; King, 1993; Searle, 1993). In particular, extensive variation in Robertsonian fusion and centric fission has been documented among different populations of the house mouse (Nachman and Searle, 1995; Britton-Davidian *et al.*,

2000) and the common shrew (Searle, 1993). Although variation in Robertsonian fusions among populations and species suggest that they might play a role in speciation (White, 1978; King, 1993), the mechanisms by which chromosome fusion and fission contribute to the formation of new species are still not well understood.

I also found evidence for chromosomal inversions between stickleback species. Many studies document the presence of inversion polymorphisms within and between species, and inversions have also been proposed to contribute to adaptation and speciation (White, 1978; King, 1993; Ayala and Coluzzi, 2005; Hoffmann and Rieseberg, 2008). A recent study in the yellow monkeyflower (*Mimulus guttatus*) provides some of the first direct experimental evidence that a chromosomal inversion contributes to local adaptation and reproductive isolation (Lowry and Willis, 2010). Similar studies in sticklebacks could resolve whether the inversions we have identified in the *A. quadracus* and *G. aculeatus* karyotypes were adaptive or contributed to the divergence of species within this family.

Table 3.1. BAC clones used as FISH probes in this study.

Chr	Length (bp)	BAC	Color	Location (bp)
01	28,185,914	CH215-01F04	purple	808,402-984,867
		CH215-36H04	green	26,520,970-26,685,556
02	23,295,652	CH215-14A06	purple	1,256,893-1,362,297
		CH215-51B03	green	22,208,609-22,374,157
03	16,798,506	CH215-22E01	purple	904,453-1,069,729
		CH215-18O23	green	14,811,263-14,979,917
04	32,632,948	CH215-37J16	purple	1,324,641-1,484,145
		CH215-03K06	green	31,667,549-31,830,249
05	12,251,397	CH215-10F03	purple	906,711-1,099,727
		CH215-15A11	green	11,086,429-11,245,223
06	17,083,675	CH215-39G07	purple	1,009,878-1,175,833
		CH215-30J20	green	16,083,765-16,163,690
07	27,937,443	CH215-24C20	purple	1,118,953-1,279,901
		CH215-50F19	green	26,759,364-26,929,269
08	19,368,704	CH215-18K21	purple	1,728,015-1,897,447
		CH215-50P04	green	18,234,424-18,411,738
09	20,249,479	CH215-09O14	purple	1,205,121-1,367,430
		CH215-35K20	green	18,760,858-18,914,793
10	15,657,440	CH215-32N20	purple	939,436-1,124,528
		CH215-13K03	green	14,608,763-14,740,402
11	16,706,052	CH215-22E24	purple	1,328,616-1,481,096
		CH215-32F20	green	15,594,645-15,722,431
12	18,401,067	CH215-35E14	purple	951,837-1,050,566
		CH215-19P17	green	17,405,092-17,588,100
13	20,083,130	CH215-29N20	purple	1,278,281-1,129,621
		CH215-41J19	green	18,303,573-18,452,590
14	15,246,461	CH215-50I22	purple	1,523,325-1,641,391
		CH215-04P19	green	14,054,345-14,228,449
15	16,198,764	CH215-42G09	purple	1,136,990-1,288,015
		CH215-47G02	green	14,910,768-15,091,535
16	18,115,788	CH215-25D03	purple	5,417,553-5,595,117
		CH215-59O23	green	15,244,265-15,412,230
17	14,603,141	CH215-03L14	purple	1,687,087-1,798,208
		CH215-25N10	green	13,556,132-13,679,389
18	16,282,716	CH215-60K06	purple	2,420,846-2,594,158
		CH215-58L04	green	14,897,105-15,061,421
19	20,240,660	CH215-23N18	purple	849,540-1,016,623
		CH215-16P13	green	4,756,773-4,923,988
20	19,732,071	CH215-32D17	purple	16,730,161-16,890,598
		CH215-09B01	green	18,782,835-18,919,085
21	11,717,487	CH215-11F14	purple	982,376-1,039,352
		CH215-34H24	green	10,642,915-10,791,824

For each *G. aculeatus* chromosome, the total length in the *G. aculeatus* genome assembly (version BROAD S1, February 2006, http://www.ensembl.org/Gasterosteus_aculeatus/index.html) is given. The pseudocolor for each BAC clone is indicated, as are the positions of the ends of each BAC clone in the respective chromosome assembly.

Table 3.2. Number of individual samples and pools analyzed by FISH in *A. quadracus* and *G. aculeatus* and karyotyped in *A. quadracus*.

	FISH analyses			Karyotypic analyses		
	Total pools analyzed	Individuals/ pool	Metaphases analyzed/chr	Total pools analyzed	Individuals/ pool	Metaphases analyzed
A. quadracus Female	8	12 (9-17)	5 (3-10)	3	11 (7-15)	6
A. quadracus Male	7	8 (7-11)	3 (3-5)	2	8 (6-11)	5
G. aculeatus Male	5	21 (15-36)	3 (3-4)	Not applicable	Not applicable	Not applicable

We pooled individual male and female *A. quadracus* and male *G. aculeatus* individuals to make metaphase spreads for FISH analysis (Figure 3.1; Figure 3.2; Table 3.3); we pooled individual male and female *A. quadracus* individuals to make karyograms (Figure 3.3; Table 3.4). Here I show the number of pools used for FISH and karyotype experiments, the mean number (range in parenthesis) of individuals per pool, and the mean number (range in parenthesis) of metaphases analyzed for each chromosome (FISH screen) or total number of metaphases analyzed for each experiment (karyotypes).

Table 3.3. Chromosome morphology and evidence for chromosome rearrangements differentiating the chromosomes of *G. aculeatus* and *A. quadracus*.

G. aculeatus chromosome	Mean arm ratio (L:S)	Chromosome morphology	*A. quadracus* chromosome	Mean arm ratio (L:S)	Chromosome morphology	Proposed rearrangement
Chr01	1.90	submetacentric	Chr01	3.33	acrocentric	Pericentric inversion
Chr02	5.28	acrocentric	Chr02	4.30	acrocentric	
Chr03	4.67	acrocentric	Chr03	5.10	acrocentric	Pericentric inversion
Chr04	2.15	submetacentric	Chr04a	3.39	acrocentric	Fission/fusion
			Chr04b	5.04	acrocentric	
Chr05	2.13	submetacentric	Chr05	1.98	submetacentric	Paracentric inversion
Chr06	4.38	acrocentric	Chr06	5.92	acrocentric	Paracentric inversion
Chr07	1.23	metacentric	Chr07a	4.88	acrocentric	Fission/fusion
			Chr07b	6.52	acrocentric	
Chr08	9.17	telocentric	Chr08	1.32	metacentric	Pericentric inversion
Chr09	6.04	acrocentric	Chr09	3.41	acrocentric	
Chr10	13.68	telocentric	Chr10	3.48	acrocentric	
Chr11	5.37	acrocentric	Chr11	11.65	telocentric	
Chr12	5.51	acrocentric	Chr12	3.77	acrocentric	
Chr13	5.74	acrocentric	Chr13	3.08	acrocentric	
Chr14	1.49	metacentric	Chr14	1.36	metacentric	
Chr15	9.97	telocentric	Chr15	3.06	acrocentric	
Chr16	9.73	telocentric	Chr16	8.12	telocentric	
Chr17	1.42	metacentric	Chr17	3.23	acrocentric	Pericentric inversion
Chr18	6.19	acrocentric	Chr18	3.52	acrocentric	
Chr19 (X)	2.67	submetacentric	Chr19	2.44	submetacentric	
Chr20	9.88	telocentric	Chr20	1.36	metacentric	Pericentric inversion
Chr21	1.39	metacentric	Chr21	3.36	acrocentric	Pericentric inversion

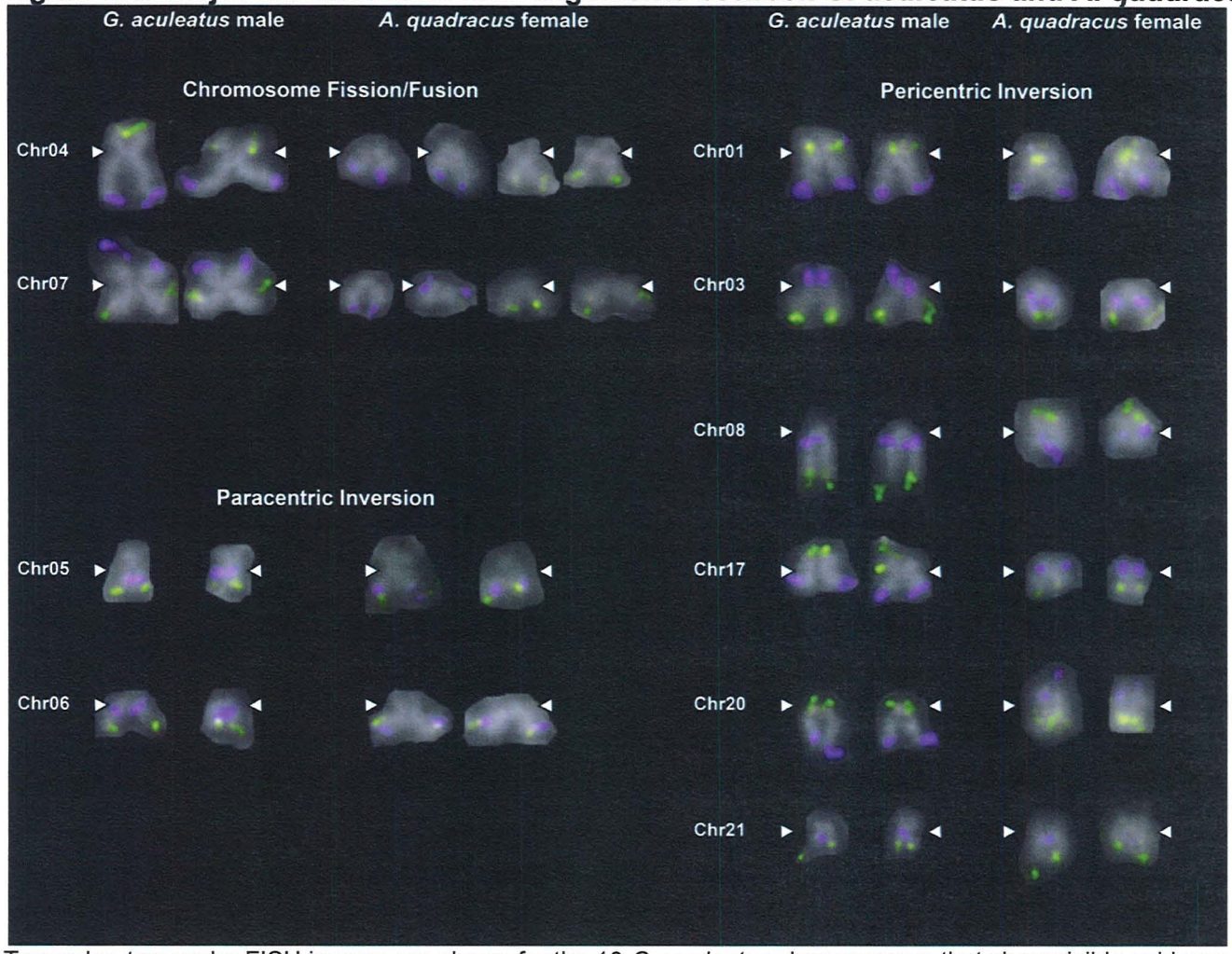
Based on the chromosome morphology, relative centromere position, and BAC probe hybridization positions for each chromosome in our FISH screen, I list the types of chromosome rearrangements that could explain the observed differences between the *G. aculeatus* and *A. quadracus* karyotypes.

Table 3.4. Major features of the karyotypes of *G. aculeatus* and *A. quadracus*.

Species	Sex	2n	Metacentric	Submetacentric	Acrocentric	Telocentric	NF
G. aculeatus	Female	42	8	6+2X	16	10	58
	Male	42	8+Y	6+X	16	10	58
A. quadracus	Female	46	6	4	32	4	56
	Male	46	6	4	32	4	56

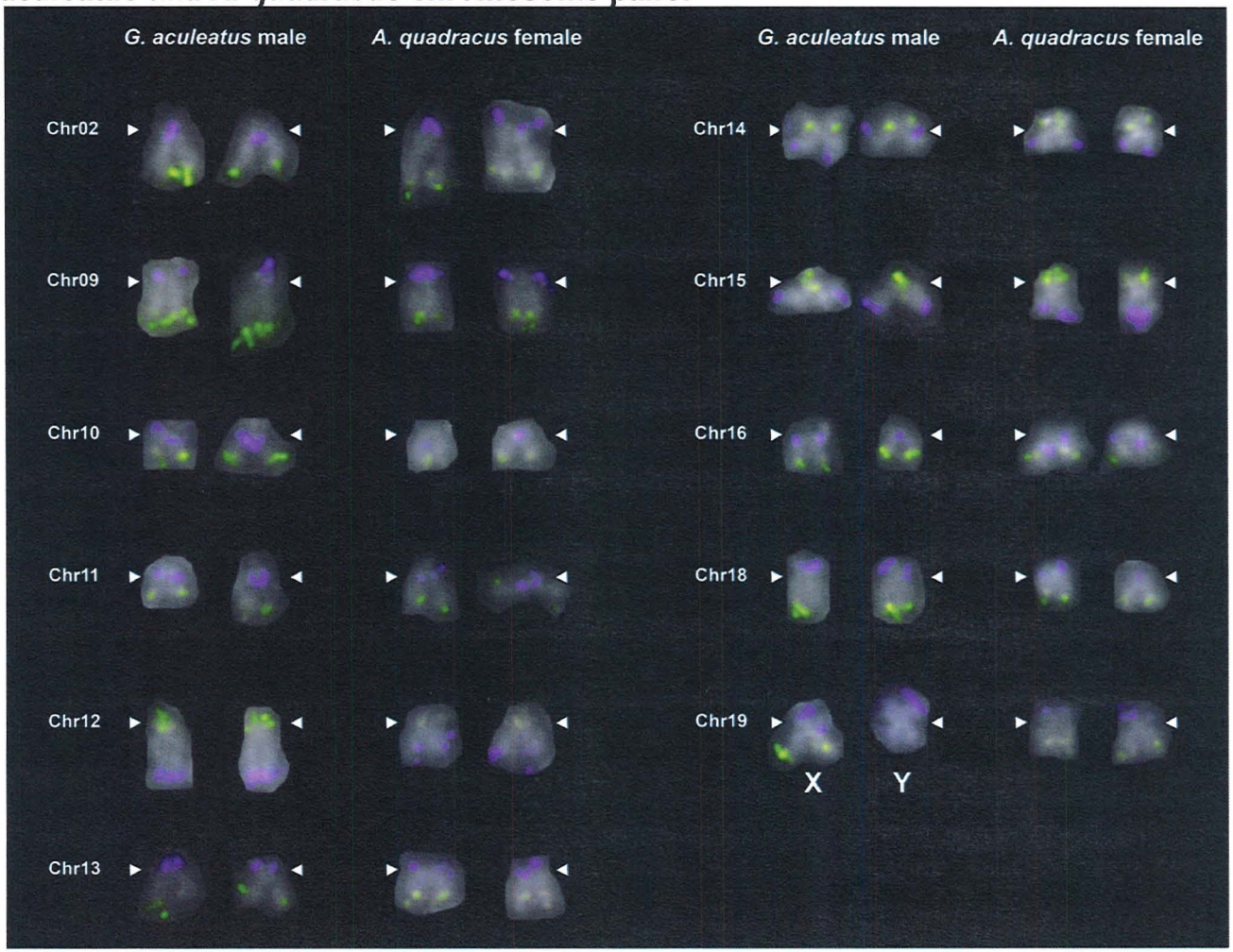
2n is the diploid chromosome number. NF ('*nombre fondamental*') is the number of major chromosome arms in a karyotype; it does not count the short arms of acrocentric and telocentric chromosomes (Matthey, 1949; White, 1978; Klinkhardt, 1998). We inferred karyotype features for *G. aculeatus* females based on the absence of the metacentric Y and the presence of a second submetacentric X chromosome (Ross and Peichel, 2008).

Figure 3.1. Major chromosomal rearrangements between *G. aculeatus* and *A. quadracus*.



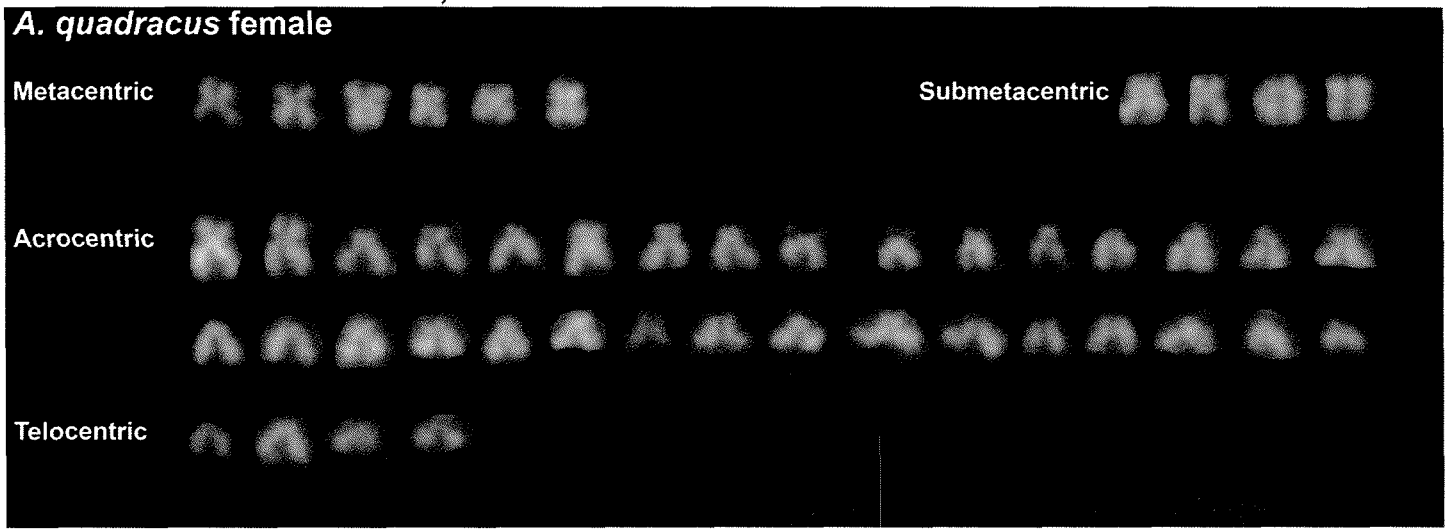
Two-color, two-probe FISH images are shown for the 10 *G. aculeatus* chromosomes that show visible evidence for chromosome rearrangement. Pairs of homologs from individual metaphase spreads of *G. aculeatus* males and *A. quadracus* females are shown. Chromosomes are arranged by the type of chromosome rearrangement: chromosome fission/fusion (2 chromosome pairs), paracentric inversion (2 chromosome pairs), and pericentric inversion (6 chromosome pairs). White arrowheads indicate the position of the centromere. Figure 3.2 shows the FISH images of the *G. aculeatus* and *A. quadracus* chromosomes that show identical morphology and hybridization patterns between species.

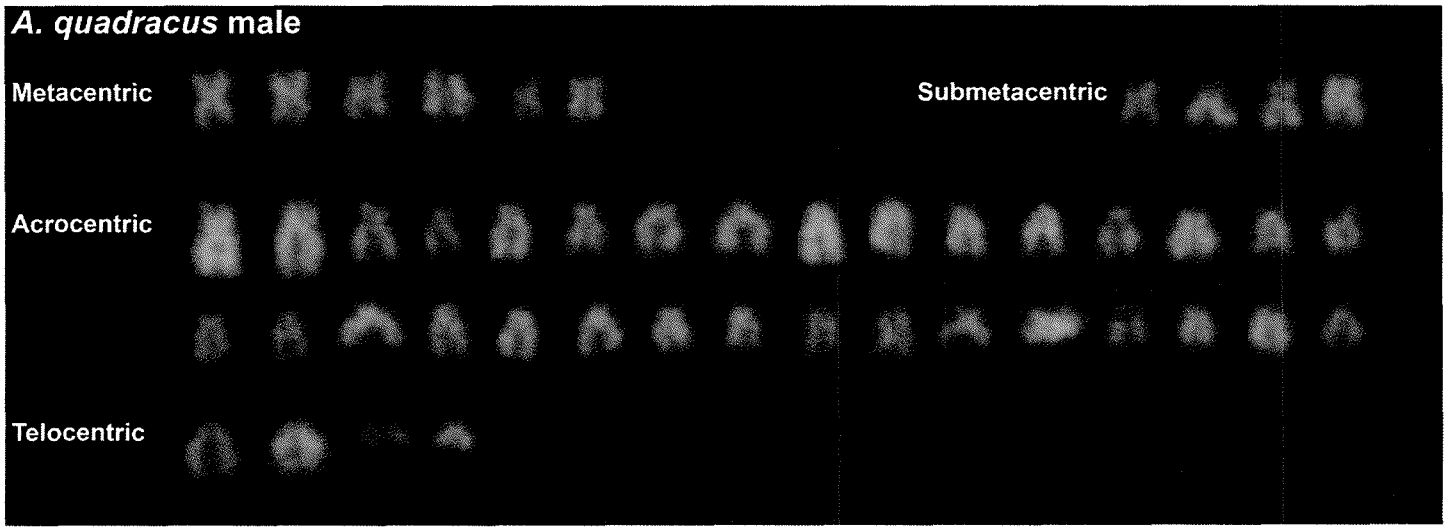
Figure 3.2. No major differences in chromosome morphology or probe hybridization exist between 11 *G. aculeatus* and *A. quadracus* chromosome pairs.



Two-color, two-probe FISH images are shown for both members of each chromosome pair from metaphase spreads of *G. aculeatus* males and *A. quadracus* females. White arrowheads indicate the position of the centromere. Figure 3.1 shows the FISH images of *G. aculeatus* and *A. quadracus* chromosome pairs that show evidence of rearrangements between these species. One FISH probe for Chr19 does not hybridize to the metacentric Y chromosome, most likely due to a large deletion on the Y chromosome in *G. aculeatus* (Ross and Peichel, 2008).

Figure 3.3. Karyograms of an *A. quadracus* female (upper panel) and an *A. quadracus* male (lower panel) from the West River in New Haven, Connecticut.





CHAPTER FOUR

Genes Expressed During Early Gonadal Differentiation in the Threespine Stickleback, *Gasterosteus aculeatus*

SUMMARY

Despite the diversity of sex chromosome systems in the stickleback family, we do not know how many sex determination genes have evolved in sticklebacks. In threespine sticklebacks (*Gasterosteus aculeatus*), the sex determination gene resides on the large Y chromosome in males. To identify this gene, I designed and executed a high-throughput next generation cDNA sequencing screen to identify genes that are differentially expressed between male and female threespine sticklebacks (*Gasterosteus aculeatus*) just as sexual differentiation begins. Male-biased genes included genes from the Y chromosome; identifying such genes will aid in the search for the *G. aculeatus* sex determination gene. I used real-time quantitative polymerase chain reaction (qPCR) to confirm the differential expression of 2 male-biased genes and 6 female-biased genes. These genes may represent some of the first genes to initiate male and female-specific developmental pathways. I also confirmed previous reports that there is no global dosage compensation mechanism in *G. aculeatus* to equalize dosage of most X chromosome genes between females and males.

In this large and complex screen, I relied on the aid and advice of many collaborators. Martin Morgan and Jerry Davison performed the genome alignments that permitted me to identify differentially expressed autosomal and X chromosome genes. Matthew Fitzgibbon devised and executed the protocol to identify expressed Y chromosome genes. Reyes Balcells genotyped over 1,000 *G. aculeatus* embryos and larvae to identify males and females. Finally, Anna Greenwood trained me in qPCR protocols and taught me the statistical concepts required to identify differentially expressed genes by RNA-Seq and confirm their expression by qPCR.

INTRODUCTION

Stickleback Sex Determination Genes

Sex chromosome systems have evolved at least three times in the stickleback family (Peichel *et al.*, 2004; Ross *et al.*, 2009; Shapiro *et al.*, 2009). However, we do not know if sex determination genes arose independently to establish each sex chromosome system, or if the same ancestral sex determination gene has transposed to different sex chromosome systems. To eliminate one of these hypotheses, the most straightforward approach is to first identify the sex determination gene in one stickleback species, and then see if it is present in the others.

The threespine stickleback, *Gasterosteus aculeatus* is the most logical stickleback species in which to look for a sex determination gene. In this species, chromosome 19 is a heteromorphic XX-XY sex chromosome pair (Peichel *et al.*, 2004; Ross and Peichel, 2008). Furthermore, a number of genetic and genomic tools have been developed for this species (Kingsley *et al.*, 2004; Kingsley and Peichel, 2007). A female (XX) *G. aculeatus* genome has been sequenced, and the sequence of the Y chromosome should be completed within the next year. In addition, protocols to make transgenic *G. aculeatus* have been developed (Hoseman *et al.*, 2004; Chan *et al.*, 2010), which will be required to test candidate sex determination genes. Thus, the number of tools developed for *G. aculeatus* make this species the most practical choice in which to search for a sex determination gene.

Sex Determination in *G. aculeatus*

In the 1950s, stickleback researcher Har Swarup generated triploid *G. aculeatus*. He used cold-shock treatments of unfertilized *G. aculeatus* eggs to force reabsorption of the second polar body and then fertilized those diploid eggs with sperm from males (Swarup, 1959a). Swarup was unaware that *G. aculeatus* had a sex chromosome system. However, based on his methods, we can assume that his triploid *G. aculeatus* had a sex chromosome compliment of either XXX or XXY, based on whether or not the diploid eggs were fertilized with X or Y-bearing sperm.

Though Swarup did not examine sex determination and sexual differentiation in the triploid *G. aculeatus*, he did obtain triploid *G. aculeatus* of both sexes (Swarup, 1959b). In addition, he did not mention any intersex or ambiguous sexual phenotypes (Swarup, 1959a; Swarup, 1959b), which could be expected if the *G. aculeatus* sex determination gene resided on the X chromosome and had a dosage effect on sex determination and sexual differentiation. Thus, based on Har Swarup's experiments with triploid *G. aculeatus*, we believe the sex determination gene in this species is a genetically dominant factor on the Y chromosome.

The *G. aculeatus* Y chromosome is nearly identical in size to its X chromosome counterpart, based on metaphase chromosome morphology (Ross and Peichel, 2008). However, we do not know the physical size of the Y chromosome, since it has not been completely sequenced. The X chromosome is approximately 20Mbp in length, and the X and Y do not recombine over approximately 16Mbp of X chromosome length (Peichel *et al.*, 2004; Ross and Peichel, 2008). The relatively small (4Mbp) pseudo-autosomal region accounts for all recombination between the X and the Y (Peichel *et al.*, 2004; Ross and Peichel, 2008). The sex determination gene likely resides in the male-specific region of the Y chromosome. We know that at least three inversions have occurred in this region, as well as one deletion event (Ross and Peichel, 2008). This deletion event (6Mbp) was large enough to eliminate hundreds of genes, including a copy of the aromatase gene *Cyp19a1* (Peichel, unpublished). Small regions of the male-specific region on the Y chromosome were previously sequenced and found to contain Y loci that have diverged in sequence from their homologues on the X, as well as multiple transposable elements (Peichel *et al.*, 2004). We do not know the frequency with which genes on the Y chromosome have degenerated to the point of becoming pseudogenes. Some of these diverged genes on the Y include "housekeeping" genes, such as the isocitrate dehydrogenase gene *Idh*, which may still have a functional Y allele (Withler *et al.*, 1986; Peichel *et al.*, 2004). Other genes include proteins implicated in gonad development and hormone synthesis, such as *17ßHSD* (17ß hydroxysteroid dehydrogenase) and *WT1* (Wilms' tumor protein 1). However, limited analyses of these genes did not uncover evidence to suggest that they are the *G. aculeatus* sex determination gene (Peichel

and Mills, unpublished). The large size of the non-recombining male-specific region, coupled with the lack of a complete Y chromosome sequence, made the search for the sex determination gene on the Y chromosome itself impractical.

Rather than locating the *G. aculeatus* sex determination gene on the Y chromosome, my approach has been to identify genes expressed when and where the sex determination gene should be expressed. The sex determination gene is most likely expressed prior to the first sign of morphological differentiation between the sexes. The earliest reported sign of morphological differentiation in *G. aculeatus* is an increase in primordial germ cell number in females, but not males (Lewis *et al.*, 2008). This is seen by late stage 27 or early stage 28, based on Swarup's developmental stages for *G. aculeatus* (Swarup, 1958; Lewis *et al.*, 2008). Thus, the sex determination gene is likely expressed prior to this stage and acts on downstream target genes to cause this morphological change between the sexes.

In order to identify the sex determination gene in sticklebacks, I believe it is important to characterize gene expression patterns between females and males prior to the morphological differentiation of the sexes at late stage 27 or early stage 28. This type of broad-scope approach will allow us to identify putative sex determination genes; i.e. Y chromosome genes that are expressed in male *G. aculeatus* (including the sex determination gene) prior to the first signs of morphological divergence between males and females. But, this approach will also let us address another fundamental question in the evolution of sexual differentiation pathways: whether or not genes in sexual differentiation pathways are conserved between *G. aculeatus* and other vertebrate species. These conserved genes could include genes and gene families implicated in gonad differentiation, hormone synthesis, and the development of secondary sexual characteristics in other organisms (Morrish and Sinclair, 2002; Haag and Doty, 2005).

I have chosen to use high-throughput "next generation" sequencing of transcripts (RNA-Seq) to characterize female and male *G. aculeatus* gene expression patterns (Wang *et al.*, 2009; Costa *et al.*, 2010). RNA-Seq will yield short sequences ("reads") from both female and male tissue from any developmental stage of my choosing. Since morphological differentiation between the sexes begins

by at least late stage 27 or early stage 28, I chose to characterize gene expression patterns between female and male *G. aculeatus* at the three latest stages of pre-hatching embryonic development (stages 22, 23, and 24), as well as the earliest stage of post-hatching larval development (stage 26) (Swarup, 1958). Stage 25 is the act of hatching and is not included in this screen due to the brief time it encompasses in the total lifespan of *G. aculeatus*; other stages of life are defined by a specific series of morphological, temporal, and behavior changes in individuals (Swarup, 1958).

By using RNA-Seq to characterize gene expression patterns between female and male *G. aculeatus* individuals, I had two broad goals aimed at understanding the evolution of sexual differentiation pathways in this species. First, I wanted to identify all genes that are differentially expressed between females and males during those late embryonic (stages 22-24) and early larval (stage 26) developmental periods. Second, I wanted to identify genes from the Y chromosome that are expressed in males during stages 22, 23, 24, and 26. These Y chromosome genes are candidate sex determination genes, and their identification will bring us closer to identifying the genetic mechanism that *G. aculeatus* uses to initiate sex-specific developmental pathways.

MATERIALS AND METHODS

Specimen Collection and Crosses

All procedures were approved by the Fred Hutchinson Cancer Research Center Institutional Animal Care and Use Committee (protocol 1575). In March 2009 and May and June 2010, adult male and female *G. aculeatus* were collected from two locations in Lake Washington: Union Bay in Seattle, Washington and the Mercer Slough Nature Park in Bellevue, Washington (Washington permits 09-038, 10-049). These fish were housed in 110L aquarium tanks (75cm length x 30cm depth x 46cm height) in summer lighting conditions (16h light: 8h dark) at approximately 16°C in 0.35% saltwater (3.5g/L Instant Ocean salt (Aquarium Systems, Mentor, Ohio, USA); 0.4mL/L $NaHCO_3$). Fish were fed live brine shrimp nauplii twice daily.

Crosses were generated in the laboratory by *in vitro* fertilization. Embryos and larvae were grown in 100mm x 15mm petri dishes in the same water and climate conditions as adult fish. *G. aculeatus* larvae did not need to be fed, as they were euthanized (below) before the yolk sac had been absorbed.

Tissue Collection

Embryos and larvae were euthanized in 0.025% tricaine methanesulfonate (MS-222). Embryos and larvae were examined to determine developmental stage (Swarup, 1958). Embryos were dechorionated using two pairs of fine-tipped forceps. Both stage and total body length (from the tip of the head to the tip of the tail) were recorded for all embryos and larvae. The posterior section of the body where the primordial gonad is found (Figure 4.1) was removed, placed into a 96-well plate, flash frozen over a bath of dry ice and 95% ethanol, and stored at -80°C for RNA extraction. The rest of the body was stored at room temperature in 95% ethanol for genomic DNA extraction. Tissues from all individuals were stored separately until the sex of each fish was determined by *Idh* locus genotyping (Peichel *et al.*, 2004).

Genomic DNA Extraction and *Idh* Locus Genotyping

Genomic DNA was extracted and purified from tissue stored in 95% ethanol at room temperature using procedures already described in this dissertation (see "Materials and Methods" section of Chapter 2). PCR-based genotyping of the *Idh* locus was performed as previously described to determine the sex of each individual embryo or larvae (Peichel *et al.*, 2004).

RNA Extraction

Tissue samples from several individuals were pooled by stage and sex for total RNA extraction for both RNA-Seq and qPCR (Table 4.1). Tissue samples were homogenized using a Tissuemiser Homogenizer (Fisher Scientific International, Hampton, New Hampshire, USA) in the TRIzol reagent (Invitrogen, Carlsbad, California, USA). One milliliter of TRIzol was used for every 50-100mg of tissue.

Total RNA was purified by chloroform extraction (1:5 chloroform:TRIzol) and isopropanol precipitation (1:2 isopropanol:TRIzol), before being resuspended in 20µL deionized water. Total RNA samples used for RNA-Seq (Table 4.1) were taken through an additional ethanol precipitation and then resuspended in 20µL deionized water. A Bioanalyzer 2100 (with the RNA 6000 Nano kit) was used to confirm RNA integrity and quantify RNA (Agilent Technologies, Santa Clara, California, USA).

Generation of cDNA Fragment Library

Four separate cDNA fragment libraries (two male and two female) were constructed (Figure 4.2) for two independent RNA-Seq high-throughput sequencing experiments. For each library, 1µg of total RNA from each of four developmental stages (22, 23, 24, and 26) was combined; the RNA sample from each stage was made by pooling multiple individuals (Table 4.1). Each cDNA fragment library was made from an independent set of RNA samples. The mRNA-Seq sample preparation kit (Illumina, San Diego, California, USA) was used according to the manufacturer's instructions to generate cDNA fragment libraries with a target length of 300bp (including adapters). Following chemical fragmentation of RNA samples, cDNA synthesis was done using SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, California, USA). PCR purification steps in the mRNA-Seq sample preparation protocol were carried out using a QIAquick PCR purification kit or a MinElute PCR purification kit (Qiagen, Hilden, North Rhine-Westphalia, Germany), and a QIAquick gel extraction kit was used for the single gel purification step (Qiagen, Hilden, North Rhine-Westphalia, Germany). Each cDNA fragment library was quantified using the High Sensitivity DNA kit on a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, California, USA).

RNA-Seq

Single-end high-throughput sequencing of each cDNA library was carried out on a Genome Analyzer II (Illumina, San Diego, California, USA) for 72 cycles, yielding reads of 80-100bp. Technical replicates of each library were sequenced on

the same flow cell, while biological replicates (i.e. the two male libraries or the two female libraries) were sequenced on different flow cells (Table 4.2).

Analyzing RNA-Seq Datasets to Identify Differentially Expressed Genes

The total number of sequence fragments ("reads") obtained by RNA-Seq for each cDNA fragment library (including technical replicates) are shown in Table 4.2. The Burrows-Wheeler Aligner (BWA) tool (Li and Durbin, 2009) was used to align reads to the *G. aculeatus* female (XX) genome (BROAD S1 assembly, February 2006, available at http://www.ensembl.org/Gasterosteus_aculeatus/index.html) (Table 4.2; Figure 4.3).

The BWA alignment data were used to identify autosomal and X chromosome genes that are differentially expressed between the sexes in each independent RNA-Seq experiment (Table 4.2). Data from all samples were normalized to account for the different numbers of total reads obtained (reads per kilobase of exon model, per million mapped reads; RPKM) (Mortazavi *et al.*, 2008). Then, for each predicted gene in the *G. aculeatus* female genome, the female:male expression ratio was calculated based on the adjusted RPKM value for each sex. Once log-transformed, the Z-score and significance (p) values for each gene were calculated to determine which genes showed a significant expression difference ($p<0.05$) between the sexes (Cheadle *et al.*, 2003).

Analyzing RNA-Seq Datasets to Identify Expressed Y Chromosome Genes

Reads were aligned to the *G. aculeatus* female (XX) genome (BROAD S1 assembly, February 2006, available at http://www.ensembl.org/Gasterosteus_aculeatus/index.html) and 27 bacterial artificial chromosome (BAC) sequences from the ongoing *G. aculeatus* Y chromosome sequencing (Table 4.3) project using Bowtie and Tophat (Trapnell *et al.*, 2009), while transcripts were assembled *de novo* using Cufflinks (Trapnell *et al.*, 2010). Velvet was then used to assemble reads that aligned to Y chromosome BACs or that did not align to anything into longer contiguous sequences (Zerbino and Birney, 2008). Bowtie was used to re-align these contiguous sequences to the

G. aculeatus female genome and the 27 Y chromosome BACs. In addition, the Cufflinks-assembled *de novo* transcripts and the Velvet-assembled contiguous sequences were used in a BLAST against the NCBI non-redundant protein database to identify putative homologues of these genes (Figure 4.3). These data were used to identify potential Y chromosome genes that are expressed in male *G. aculeatus.*

RNA Samples for qPCR

Real-time quantitative PCR (qPCR) was used to validate RNA-Seq expression data. The RNA templates chosen were 31 biological replicates from the following *G. aculeatus* developmental stages (Table 4.1): 8 samples of "pre-hatching" females (stages 22-24), 7 samples of "pre-hatching" males (stages 22-24), 8 samples of "post-hatching" females (stage 26), and 8 samples of "post-hatching" males (stage 26). These samples will be known as "test templates" in future sections. None of these samples had been used previously in the construction of cDNA fragment libraries for RNA-Seq.

In addition to these 31 test templates, a mixture of RNA samples was prepared for the standard curve qPCR reactions. The standard curve sample consisted of equal amounts of RNA from males and females from the four developmental stages (22, 23, 24, and 26) used in construction of the cDNA fragment libraries for RNA-Seq. The standard curve sample consisted of RNA from individuals that had also been used in the construction of cDNA fragment libraries for RNA-Seq.

cDNA Synthesis for qPCR

The RNA test templates and standard curve sample were treated with amplification-grade DNaseI (Invitrogen, Carlsbad, California, USA) according to the manufacturer's instructions. First-strand cDNA synthesis was carried out using the SuperScript III First-strand Synthesis System for RT-PCR (Invitrogen, Carlsbad, California, USA) according to the manufacturer's instructions.

A control cDNA synthesis reaction was also prepared, using RNA mixed in equal portions from one stage 26 male individual and one stage 26 female

individual. For this cDNA synthesis, the sample was not treated with reverse transcriptase, and a corresponding volume of deionized water was added instead. Aside from that deviation, no other steps in the DNaseI treatment and first-strand cDNA synthesis were altered. This control sample was used to assess whether the DNaseI treatment was effective at clearing genomic DNA from the sample prior to first-strand cDNA synthesis.

Quantitative PCR

Real-time quantitative PCR (qPCR) with SYBR Green Dye (Applied Biosystems, Carlsbad, California, USA) was used to validate expression differences between males and females for 10 genes (Table 4.4). Reactions were executed on a 7900HT Fast Real-Time PCR System (Applied Biosystems, Carlsbad, California, USA). For the 31 test samples, the 10µL reactions consisted of 5ng cDNA (RNA equivalent), 1X SYBR Green PCR Master Mix (Applied Biosystems, Carlsbad, California, USA), and 5pmol each of the forward and reverse oligonucleotides (Table 4.4). Standard curve reactions were carried out under the same reaction conditions, but in a dilution series of 50ng, 25ng, 5ng, 0.5ng, and 0.05ng of cDNA template (RNA equivalent). Two control reactions were also used: one with deionized water substituted for template, and the other is the cDNA synthesis reaction in which no reverse transcriptase was used (see preceding "cDNA Synthesis for qPCR" section). Triplicate reactions were performed for every test template, negative control, and standard curve sample for all oligonucleotide pairs.

All qPCR reactions were carried out under the following conditions: 50°C for 2 minutes, 95°C for 10 minutes, and 40 cycles of a two-temperature fluctuation (95°C for 15 seconds, 60°C for 1 minute). An additional ramp cycle of 95°C for 15 seconds, 60°C for 15 seconds, and back to 95°C for 15 seconds was used to calculate melting curves for each oligonucleotide pair.

Cycle threshold (CT) values were exported from Sequence Detection Systems 2.3 software for the 7900HT Fast Real-Time PCR System (Applied Biosystems, Carlsbad, California, USA), and data were analyzed in Excel 2008 for Mac, version 12.2.8 (Microsoft, Redmond, Washington, USA). Since triplicate

reactions were done for all templates, the mean CT value was calculated for each template. The mean CT values for the six standard curve samples were used to calculate the real-time PCR efficiency (E) of each oligonucleotide pair,

$$E = 10^{(-1/Slope)}$$

where "Slope" is the slope of the best-fit line where mean CT for the standard curve templates are graphed as a function of cDNA input into each reaction (Pfaffl, 2001).

E values were then used to calculate the relative expression of each gene for each test template, normalized to the expression of the control gene *EEF1B2*. *EEF1B2* was chosen due to its stable expression between sexes by RNA-Seq (data not shown) and qPCR (Figure 4.4). In addition, other elongation factor proteins have previously been used as stable reference genes for qPCR in *G. aculeatus* (Hibbeler *et al.*, 2008). Gene expression values, normalized by *EEF1B2*, were calculated for each test template using the formula

$$\text{Normalized expression} = E_{EEF1B2}^{CT(EEF1B2)} / E_{gene}^{CT(gene)} \times 100$$

where E_{gene} and E_{EEF1B2} are the respective efficiency values for the gene in question and the reference gene *EEF1B2* (both of which were calculated above using the standard curve samples), and CT(gene) and CT(*EEF1B2*) are the respective mean cycle threshold values of the test sample for gene in question and *EEF1B2* (Pfaffl, 2001).

The relative expression values of each gene for each test template were then exported into PASWStatistics 18.0 (SPSS/IBM, Armonk, New York, USA). Independent samples t-tests were used to identify genes that showed a significant difference in expression ($p<0.05$) by sex or age. A two-way analysis of variance (ANOVA), followed by least-significant difference *post hoc* tests, was used to identify significant differences in gene expression by both sex and age.

RESULTS

Differentially Expressed Genes in *G. aculeatus*

I performed high-throughput next-generation sequencing of two separate cDNA fragment libraries for both female and male *G. aculeatus* (RNA-Seq). We independently aligned all reads from both sets of biological replicates (two sets of

reads from males, and two sets of reads from females) to the *G. aculeatus* female (XX) genome using BWA (Table 4.2; Figure 4.3). I compared sets of differentially expressed genes between both biological replicates to determine which genes were male-biased ($p<0.05$) in both replicates and which genes were female-biased ($p<0.05$) in both replicates. In this comparison, I discarded any gene less than 220bp and any gene with less than 10 reads from both sexes aligning to it.

Sixty-six male-based genes ($p<0.05$) were common to both biological replicates (see Appendix). I chose four of these genes to confirm their male-biased expression by qPCR (Table 4.4). For one of these genes, *PKD2*, qPCR did not reveal any significant difference in expression level between males and females, or between stickleback embryos (stages 22-24) and larvae (stage 26) (Figure 4.5). The putative transcription factor *GTF2IRD2* showed a significant male bias in expression ($p<0.05$). Although there is no difference in expression between male and female embryos, *GTF2IRD2* expression decreases significantly in female larvae after hatching (Figure 4.5). For the Buster3 transposase-like gene *C5orf54*, qPCR confirmed a significant male bias in expression after hatching (stage 26) (Figure 4.5). But, the respective qPCR efficiencies for the *C5orf54* and *PKD2* oligonucleotide pairs fell outside of our quality control standards ($1.9<E<2.1$) (Table 4.4), which are similar to standards used in other qPCR verifications of RNA-Seq data (Nagalakshmi *et al.*, 2008). Thus, I consider qPCR results for these two genes to be preliminary, and subject to confirmation by future independent investigations. For the gene *PAQR4*, qPCR did not confirm male-biased expression. Instead, *PAQR4* showed a significant ($p<0.05$) female bias in expression after hatching (stage 26) (Figure 4.5).

I uncovered 301 genes that showed a consistent and significant ($p<0.05$) female expression bias in both biological replicates by RNA-Seq (see Appendix). Of those 301 genes, 275 reside on the X chromosome (Chr19). These 275 genes were expressed at levels approximately two-fold higher in females compared to males for both RNA-Seq experiments; I calculated a mean female:male expression bias of 2.1144 in one RNA-Seq experiment (median = 2.0341), and 2.0464 in the second RNA-Seq experiment (median = 1.9676). This two-fold female bias among X

chromosome genes likely reflects a lack of dosage compensation in *G. aculeatus*, in agreement with previous data (Leder *et al.*, 2010). A single X chromosome gene *CCDC34* showed a female bias much greater than the mean of all other X chromosome genes for both RNA-Seq experiments (3.4829 in one experiment, and 2.7695 in the other). I confirmed the female bias of *CCDC34* expression (regardless of age) by qPCR ($p<0.001$), though *CCDC34* also showed a significant decrease in expression in *G. aculeatus* larvae (stage 26) compared to stickleback embryos (stages 22-26) (Figure 4.6).

Twenty-six autosomal genes showed a consistent and significant ($p<0.05$) female bias in expression in both RNA-Seq experiments (see Appendix). I selected five genes to confirm their significant difference by qPCR (Table 4.4). Although I found no significant expression differences between males and females for the gene *OVGP1* by qPCR, there was a significant increase in expression between males after hatching (Figure 4.7). I confirmed a female bias in expression regardless of age for genes *ZAR1*, *EPS8L2*, and *MYCBPAP*, but not for *ZFP106* (Figure 4.6; Figure 4.7). However, *ZFP106* did show a significant female bias in expression among embryos (stages 22-24), while expression levels were lower among larvae (stage 26) of both sexes (Figure 4.6). I consider *MYCBPAP* qPCR results preliminary, since the qPCR efficiency (E) of the *MYCBPAP* oligonucleotides fell outside of my quality control standards ($1.9<E<2.1$) (Table 4.4). Thus, by qPCR, I confirmed male-biased expression for 2 out of 4 genes identified by RNA-Seq, and confirmed female-biased for 5 out of 6 female-biased genes.

Expressed Y Chromosome Genes in *G. aculeatus*

My colleague Matt Fitzgibbon took two parallel approaches to identify RNA-Seq reads from Y chromosome genes. Both approaches made use of the *de novo* transcript assembly tools. In his first approach, Matt used Cufflinks (Trapnell *et al.*, 2010) to assemble longer transcript sequences from reads from males or females that aligned to the 27 Y chromosome BACs (Table 4.3) by Bowtie and Tophat. In his second approach, Matt used Velvet (Zerbino and Birney, 2008) to assemble contiguous sequences among reads that did not align to either the *G. aculeatus*

female genome or the 27 Y BAC sequences. He then aligned these Velvet-assembled contigs to the *G. aculeatus* genome and the Y chromosome BACs. Finally, he performed a BLAST search of all Velvet and Cufflinks-assembled contiguous sequences against the NCBI non-redundant protein database.

Matt's *de novo* transcript assembly yielded 10,278 new contiguous sequences (contigs) from females and 12,370 new contigs from males. Among female contigs, 7,783 out of 10,278 contigs (75.72%) had BLAST results that returned at least one homologous sequence, as did 9,552 out of 12,370 (77.22%) of male contigs (data not shown). Only 137 female contigs aligned to Y chromosome BAC sequences (including 119 contigs with BLAST results), while 653 male contigs (555 with BLAST results) aligned to Y chromosome BAC sequences (data not shown).

DISCUSSION

Conservation of Sexual Differentiation Genes in *G. aculeatus*

To reveal which genes are differentially expressed between the sexes during the first steps of sexual differentiation in *G. aculeatus*, I chose to sequence cDNA fragment libraries from mRNA derived from midsection tissue of embryos and larvae (Figure 4.1). Female and male *G. aculeatus* are morphologically distinguishable by stages 27 and 28 (Lewis *et al.*, 2008), and possibly as soon as late stage 26 (Bruner, unpublished). This first sign of morphological differentiation between the sexes is an increase in primordial germ cell number in females (Lewis *et al.*, 2008). However, a burst of sexually dimorphic gene expression likely precedes morphological differentiation of the sexes (Sekido and Lovell-Badge, 2009; Shibata *et al.*, 2010; Wang *et al.*, 2010; Okubo *et al.*, 2011). This sexually dimorphic gene expression should include the elusive Y-linked sex determination gene, as well as its immediate downstream gene targets. I designed this screen to investigate the gene expression patterns of sticklebacks in their late embryonic stages (22, 23, and 24) and early larval stage (26).

This study reveals that there are sexually dimorphic genes in *G. aculeatus* in late embryonic and early larval stages. These include both male-biased and female-

biased genes on the autosomes and the sex chromosomes. However, they do not include genes from conserved sexual differentiation pathways in other vertebrates. For example, I saw no difference in expression between the sexes for genes such as *DMRT1*, *SOX9a*, *SOX9b*, *Cyp19a*, or *DAX1* (Morrish and Sinclair, 2002; Piprek, 2009) in *G. aculeatus* embryos and larvae (data not shown). However, the developing gonad is a relatively small portion of the tissue sample I used for RNA-Seq and qPCR (Figure 4.1). Thus, it is possible that any sexually dimorphic gene expression patterns in the gonad may be masked by gene expression patterns from surrounding somatic tissues. Future studies could attempt to address this point by selecting a smaller tissue target for RNA-Seq, preferably just the developing gonad. This approach may also identify the *G. aculeatus* sex determination gene once a complete Y chromosome sequence is available.

Sexually Dimorphic Gene Expression in *G. aculeatus*

By RNA-Seq and qPCR, I have identified *GTF2IRD2* as an early male-biased gene in *G. aculeatus*. There is no significant difference in expression between male and females during the late embryonic stages of development. But, by the early larval stage, *GTF2IRD2* expression is significantly lower in females compared to males (Figure 4.5). *GTF2IRD2* is a putative transcriptional regulator and is conserved among vertebrates; in mammals, it is widely expressed in developing embryos (Enkhmandakh *et al.*, 2004; Makeyev *et al.*, 2004). However, no reports to date indicate a role for this gene in sexual differentiation. Though *GTF2IRD2* is autosomal, there is evidence that a copy of this gene may have also transposed to the Y chromosome (Peichel, unpublished). This transposition, if true, may help explain the gene's male-biased detection by qPCR and RNA-Seq. Future studies should investigate whether the larval male-biased expression of *GTF2IRD2* is confined to the developing gonad. If so, that may indicate a role for this gene in sexual differentiation in *G. aculeatus*, and possibly other vertebrates.

My qPCR experiments have shown that five other genes are likely female-biased in late embryonic or larval stages. *ZAR1* orthologues may have roles in germ cell development in mammals (Uzbekova *et al.*, 2006), and in *G. aculeatus* its

expression is significantly higher in females than males. The four remaining genes show similar patterns of expression both by age and sex. Females in the late embryonic stages show significantly higher levels of *PAQR4*, *CCDC34*, *ZFP106*, and *EPS8L2* than males. However, by early larval development, expression levels of *PAQR4* and *ZFP106* equalize between the sexes (Figure 4.5; Figure 4.6). In addition, expression of the preliminary female-biased gene *MYCBPAP* follows a similar pattern to *ZFP106* and *PAQR4*. For *CCDC34* and *EPS8L2*, the expression differences between male and female larvae are still significant, but have narrowed compared to expression in late embryonic stages (Figure 4.6).

These expression patterns are the opposite of the expression pattern for the male-biased gene *GTF2IRD2*, as well as the preliminary male-biased gene *C5orf54* (Figure 4.5). Though I have only surveyed a handful of genes, these expression patterns may indicate that the earliest stages of sexual differentiation consist of an upregulation of certain genes in female embryos. However, by hatching, male larvae begin to upregulate a different set of genes. If so, then the initial molecular function of the Y-linked sex determination gene in males may be to repress targets like *CCDC34*, *EPS8L2*, *ZFP106*, and *PAQR4* in late embryonic stages, preventing the initiation of female differentiation. However, as with *GTF2IRD2*, the molecular roles of these female-biased genes are not fully known, making it difficult to hypothesize the roles these genes could be playing in *G. aculeatus* female differentiation. *PAQR4* is a member of a family of vertebrate adiponectin and progestin transmembrane receptions, but no ligand for *PAQR4* has been reported for any species (Tang *et al.*, 2005). Nucleolar-localized *ZFP106* has been proposed to help regulate testis development in mammals (Grasberger and Bell, 2005). *EPS8L2* has been implicated in actin-mediated cytoskeleton remodeling in response to growth factors, but not yet linked to sexual differentiation (Offenhauser *et al.*, 2004). Finally, *CCDC34* has no reported roles in development. Thus, as with male-biased *GTF2IRD2*, future studies of these genes should first confirm their differential expression, specifically in the developing gonad. A parallel RNA-Seq screen of the developing gonad could identify additional partner genes in these pathways.

Dosage Imbalance on the Sex Chromosomes of *G. aculeatus*

A majority (275/301) of the female-biased genes identified by my RNA-Seq screen are from the X chromosome (see Appendix). Among these X chromosome genes, expression is on average twofold higher in females compared to males. These findings suggest that *G. aculeatus* lacks a global dosage compensation mechanism to equalize X gene dosage between the sexes. Instead, for many X chromosome genes, female expression is twofold higher than male expression.

Another group has reported this dosage imbalance of X chromosome genes between the sexes for *G. aculeatus*, using a microarray approach instead of RNA-Seq (Leder *et al.*, 2010). However, dosage imbalance in *G. aculeatus* will be difficult to study until the complete Y chromosome sequence is available. Future studies should map the divergence of X and Y chromosome alleles to determine which portions of the X chromosome are most likely to contain genes for which males have no functional Y chromosome copy. Then, we should be able to revisit this RNA-Seq dataset to confirm that male *G. aculeatus* have only one copy of those loci (on their single X chromosome), while females have two.

G. aculeatus is not the only organism which appears to lack a global dosage compensation mechanism between the sexes. The silkworm moth, the chicken, and the trematode parasite *Schistosoma mansoni* also do not have global dosage compensation mechanisms (Ellegren *et al.*, 2007; Vicoso and Bachtrog, 2011; Walters and Hardcastle, 2011). Interestingly, these three species have ZZ-ZW sex chromosome systems, leading to speculation that global dosage compensation may be limited to XX-XY systems. *G. aculeatus* may be the first reported species with heteromorphic XX-XY sex chromosomes to lack a global dosage compensation mechanism. However, there is evidence for the regulation of sex-specific expression levels on a gene-by-gene basis in *G. aculeatus* (Leder *et al.*, 2010). For example, in my own RNA-Seq dataset, at least nine genes from the X chromosome were upregulated in males ($p<0.05$) compared to females. Thus, in *G. aculeatus*, there may be mechanisms in place to regulate X gene dosage spatially or temporally, based on the gene and its required function. This may be similar to the gene-by-gene local dosage compensation mechanism proposed for chickens (Mank

and Ellegren, 2009). Future studies on expression of X chromosome genes in *G. aculeatus* across different tissues and ages will allow us to fully understand the broad mechanisms covering X chromosome dosage in this species. However, it appears that *G. aculeatus* is the latest species, and possibly first species with XX-XY sex chromosomes, to lack a global dosage compensation mechanism between females and males.

Y Chromosome Genes in *G. aculeatus*

Through the efforts of my colleague Matt Fitzgibbon, we are using this RNA-Seq study to begin to identify Y chromosome genes expressed in male *G. aculeatus* during these late embryonic and early larval stages. A larger number of male-derived *de novo* contigs (653) aligned to Y chromosome BAC sequences than female-derived contigs (137). However, among contigs that did not align to the Y chromosome BACs or the female *G. aculeatus* genome, there remain nearly equal numbers of male (12,370) and female (10,278) contigs. Thus, without a complete Y chromosome sequence to which we can align these contigs and identify expressed Y chromosome genes, the only tool available to sort out Y chromosome gene contigs from low-quality "junk" contigs are the results of the BLAST search of these contigs against the NCBI non-redundant protein database. These BLAST results should help discard low-quality contigs.

A larger portion male-specific contigs (including some that align to Y chromosome BAC sequences) have BLAST results indicating that that are derived from transposable elements (TEs) compared to female-specific contigs (4.29% of male contigs with BLAST results compared to 2.18% of female-specific contigs with BLAST results) (data not shown). Several prior studies have documented that TEs invade the sex-specific regions of evolving Y or W chromosomes in a variety of lineages, including *G. aculeatus* (Liu *et al.*, 2004; Peichel *et al.*, 2004; Kondo *et al.*, 2006; Bachtrog *et al.*, 2008; Marais *et al.*, 2008). Our data indicate that, at least in *G. aculeatus* male embryos and larvae, many of these TE-derived genes are expressed, particularly since our cDNA library synthesis protocol first removed RNA molecules that were not poly-adenylated (Figure 4.2). Another line of evidence that

Y chromosome TE-derived genes are expressed comes from the preliminary qPCR results for the Buster3 transposase-like gene *C5orf54* (Figure 4.5). Though I originally identified this gene as autosomal, evidence recently surfaced indicating that the Y chromosome contains at least one copy of *C5orf54*. If confirmed, this finding may help explain why *C5orf54* showed the largest male-biased expression among any autosomal or X chromosome gene between my two biological replicate RNA-Seq experiments. Thus, the *G. aculeatus* Y chromosome likely contains a number of genes required for male sex determination and sexual differentiation, but there is increasing evidence that the Y chromosome also harbors a number of TE-derived genes that are transcribed.

The experiments described in this chapter indicate that, by late embryonic development, there are already genes in the *G. aculeatus* genome that are differentially expressed between the sexes. This may indicate that the Y chromosome-based *G. aculeatus* sex determination gene acts during these developmental stages, if not sooner. The preliminary results described here are best seen in the context of a future survey to characterize Y chromosome genes expressed in male embryos and larvae, and to identify candidate sex determination genes. Future studies will use the complete Y chromosome sequence, and begin to map degenerate loci between the X and Y chromosomes. These studies may also clarify the extent to which gene-by-gene dosage compensation has evolved in *G. aculeatus*. Thus, my experiments and preliminary results will fuel future studies of the gene content of the *G. aculeatus* sex chromosomes. Once we know the sex determination gene in *G. aculeatus*, we should be able to determine whether other stickleback species with divergent sex chromosomes have evolved their own sex determination genes.

Table 4.1. Summary of the sources of RNA samples used for RNA-Seq and qPCR.

Sample	Experiment	Stage 22	Stage 23	Stage 24	Stage 26
Female cDNA fragment library 1	RNA-Seq	17	25	43	24
Male cDNA fragment library 1	RNA-Seq	26	24	35	15
Female cDNA fragment library 2	RNA-Seq	25	25	25	19
Male cDNA fragment library 2	RNA-Seq	26	25	28	21
Pre-hatching female samples 1 through 8	qPCR	2	2	2	0
Pre-hatching male samples 1 through 7	qPCR	2	2	2	0
Post-hatching female samples 1 through 8	qPCR	0	0	0	6
Post-hatching male samples 1 through 8	qPCR	0	0	0	6

The table indicates the number of individual *G. aculeatus* embryos (pre-hatching, stages 22-24) or larvae (post-hatching, stage 26) dissected (Figure 4.1) for each RNA sample.

Table 4.2. A summary of all RNA-Seq experiments conducted.

Flow Cell	Library	Number of Technical Replicates	Total number of reads	Reads aligned to *G. aculeatus* XX genome	Reads unaligned
1	Female cDNA fragment library 1	3	35,112,569	15,918,712 (45.34%)	19,193,857 (54.66%)
1	Male cDNA fragment library 1	4	47,894,649	22,261,474 (46.48%)	25,633,175 (53.52%)
2	Female cDNA fragment library 2	2	50,361,561	24,821,218 (49.29%)	25,540,343 (50.71%)
2	Male cDNA fragment library 2	2	58,188,210	27,748,939 (47.69%)	30,439,271 (52.31%)

Two flow cells (labeled "1" and "2") were used to sequence female and male cDNA fragment libraries. The number of technical replicates is the number of lanes on a single flow cell dedicated to a particular sample. Reads from each library were aligned to the *G. aculeatus* female (XX) genome (BROAD S1 assembly, February 2006, available at http://www.ensembl.org/Gasterosteus_aculeatus/index.html) using the BWA tool.

Table 4.3. Y chromosome BAC sequences used for alignment of RNA-Seq reads.

Y Chromosome BAC	Length (bp)
CH215-01O17	161,859
CH215-06A12	171,681
CH215-12C22	180,975
CH215-21M12	152,065
CH215-21O20	189,583
CH215-25A07	168,731
CH215-26O02	170,861
CH215-29O20	168,736
CH215-39H13	180,941
CH215-44C17	183,121
CH215-50F06	161,386
CH215-50O2	161,473
CH215-54G05	171,986
CH215-203C15	216,892
CH215-207L20	132,739
CH215-207M13	177,715
CH215-220M23	233,335
CH215-221F22	112,315
CH215-230L07	98,281
CH215-238B03	221,572
CH215-240I17	126,283
CH215-240N11	200,302
CH215-241K12	110,211
CH215-244B17	217,617
CH215-245B07	202,475
CH215-247K17	201,048
CH215-250P02	204,128

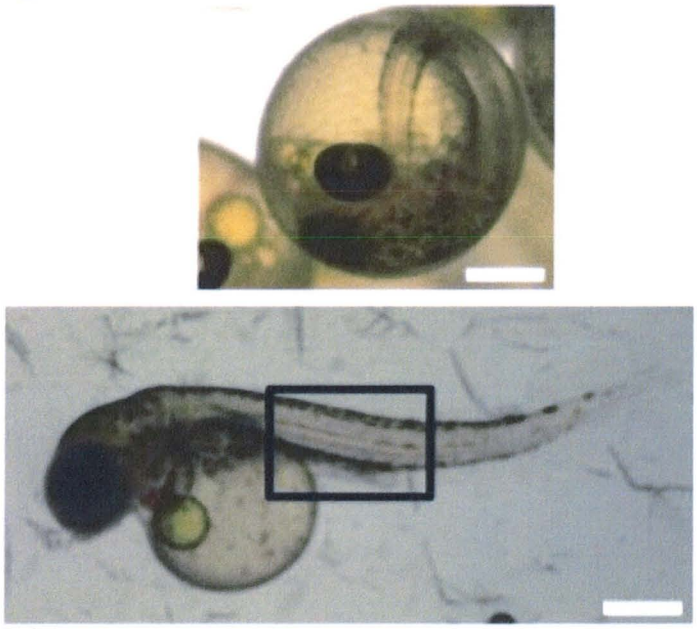
All Y chromosome BAC sequences come from the CHORI 215 BAC library (Kingsley *et al.*, 2004).

Table 4.4. A summary of genes tested by qPCR.

Gene	Sex bias	Molecular or physiological role	Chr	Ensembl gene ID	Oligonucleotides (5' to 3')	E
GTF2IRD2	Male**	Transcription factor from TFII-I family; target genes unknown	Un	ENSGACG00000013474	For: CAATGCTACTGCATCTGAGGAAA Rev: GCTGATCAAGGTCACAAAACCTATC	1.93
PKD2	Male*	Transmembrane ER cation channel; Ca^{2+} signaling; vascular integrity	16	ENSGACG00000002153	For: CGCGACATGACCTACTACGAAA Rev: CCAGGAAGGCGACCAGATAC	1.76
PAQR4	Male*	Related to transmembrane receptors for adiponectin and progestin; ligands unknown	11	ENSGACG00000013570	For: ACTGCATCAGGAGCCTGTTCTAC Rev: GGATGCCGTGCGTGTAAAT	2.05
C5orf54	Male**	Transposon-derived Buster3 transposase-like protein	18	ENSGACG00000006487	For: TGCCTGTTCTCGGCTTTGT Rev: TTTGCGATCGATGCTGTGA	2.51
CCDC34	Female**	Coiled-coil domain containing 34; function unknown	19/X	ENSGACG00000010963	For: GCTGGTGAACAAGGCCAAA Rev: CTGCTCCTCATCTGCTTGTTTTT	2.08
OVGP1	Female*	Oviductal glycoprotein 1; secreted; possible regulation by estradiol	12	ENSGACG00000003535	For: AGGGAAAATACCCCCTCATTAGATA Rev: CAATGGAGGTGGAAAATCTGAAT	1.94
ZAR1	Female*	Zygote arrest 1; reported sex-specific gonad expression; possible transcription regulator	08	ENSGACG00000014170	For: CGTACGTGTGGTGCGTTCAG Rev: TTTGGCATTTCCTGCAGAACT	2.02
ZFP106	Female*	Zinc finger protein 106; possible roles in mammalian testes development, transcriptional regulation	Un	ENSGACG00000000693	For: CACATATGGGAACCACGTTTTACA Rev: TCAGGCCGTCAGAGAAGTCA	2.03
EPS8L2	Female*	EPS8-like 2; actin remodeling in response to growth factors	Un	ENSGACG00000018690	For: CAACCAGCAGACAGGCTACAGT Rev: AAGCTGTCTCGATGGCTCAGA	1.95
MYCBPAP	Female**	c-Myc binding protein; possible role in mammalian spermatogenesis	09	ENSGACG00000018886	For: GCGGAAGCCCGTCACA Rev: GATTTCCTGATTCCTGGCGTAT	2.21
EEF1B2	None	Eukaryotic translation elongation factor 1 beta 2	01	ENSGACG00000015402	For: CCGCTGGTACAACCACATCA Rev: ACTGACCCAGAGGCTTCTTCAC	2.01

Ten genes were tested by qPCR, while *EEF1B2* was used as a reference gene to normalize expression of the test genes.
E: qPCR oligonucleotide efficiency (Pfaffl, 2001); E values that fall outside of quality control standards (1.9<E<2.1) are indicated in red
Significance of sex bias in gene expression (by RNA-Seq): *p<0.05 or **p<0.01 from two biological replicates
Un: unassembled; scaffold in the *G. aculeatus* genome that has not yet been incorporated into a chromosome assembly

Figure 4.1. *G. aculeatus* embryos.



Upper panel shows a stage 24 (Swarup, 1958) *G. aculeatus* embryo in its chorion. Lower panel shows a stage 24 *G. aculeatus* embryo after being euthanized and removed from its chorion casing. The black box indicates the region of tissue (which includes the developing gonad) from which total RNA was extracted for RNA-Seq and qPCR. White scale bars = 0.5mm.

Figure 4.2. An overview of the construction of cDNA fragment libraries for RNA-Seq.

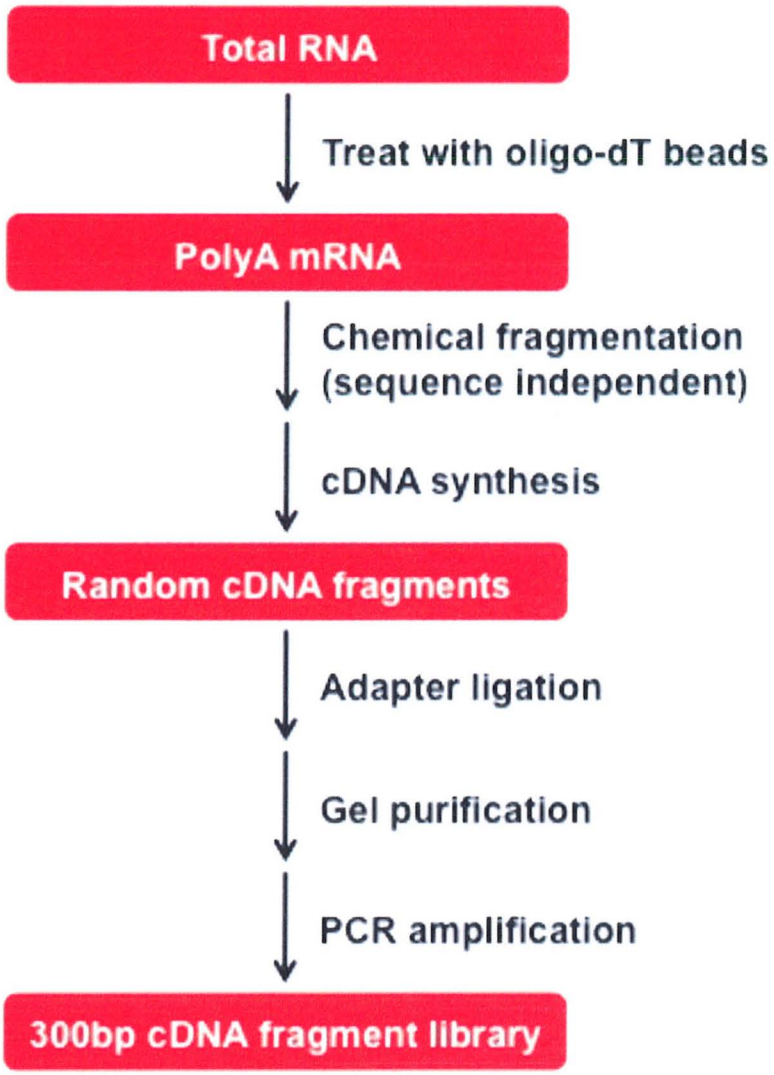


Figure 4.3. An overview of the analysis of the RNA-Seq datasets to identify differentially expressed genes and Y chromosome genes.

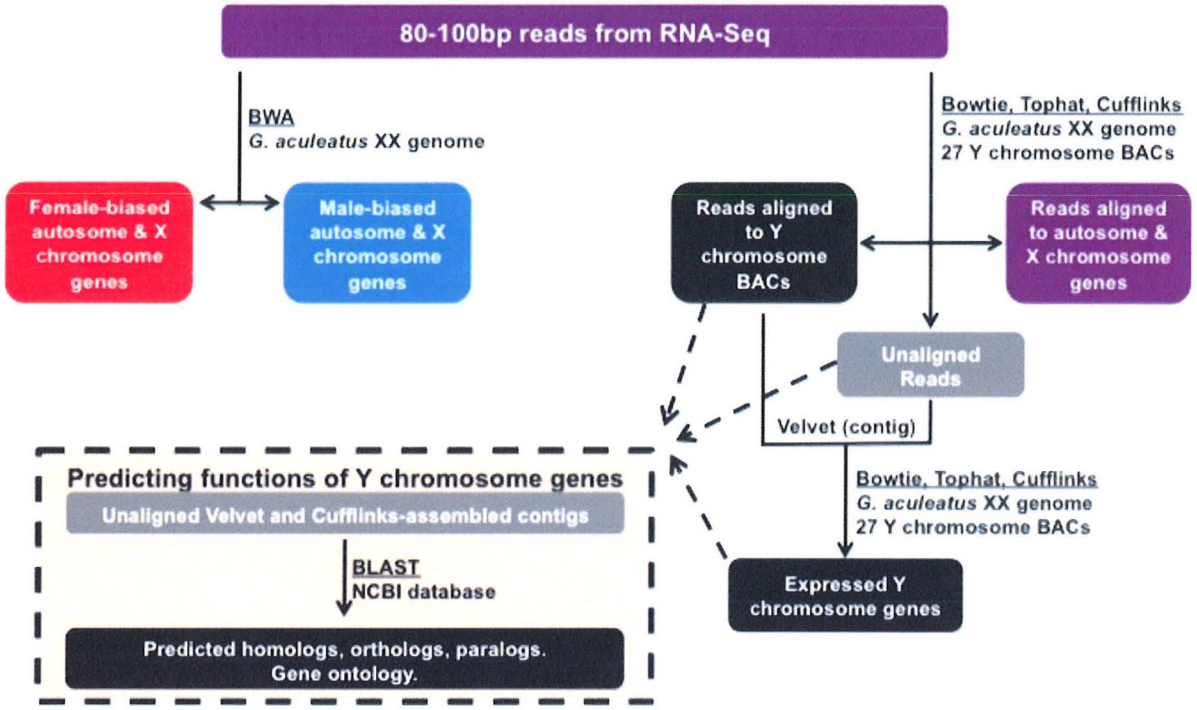
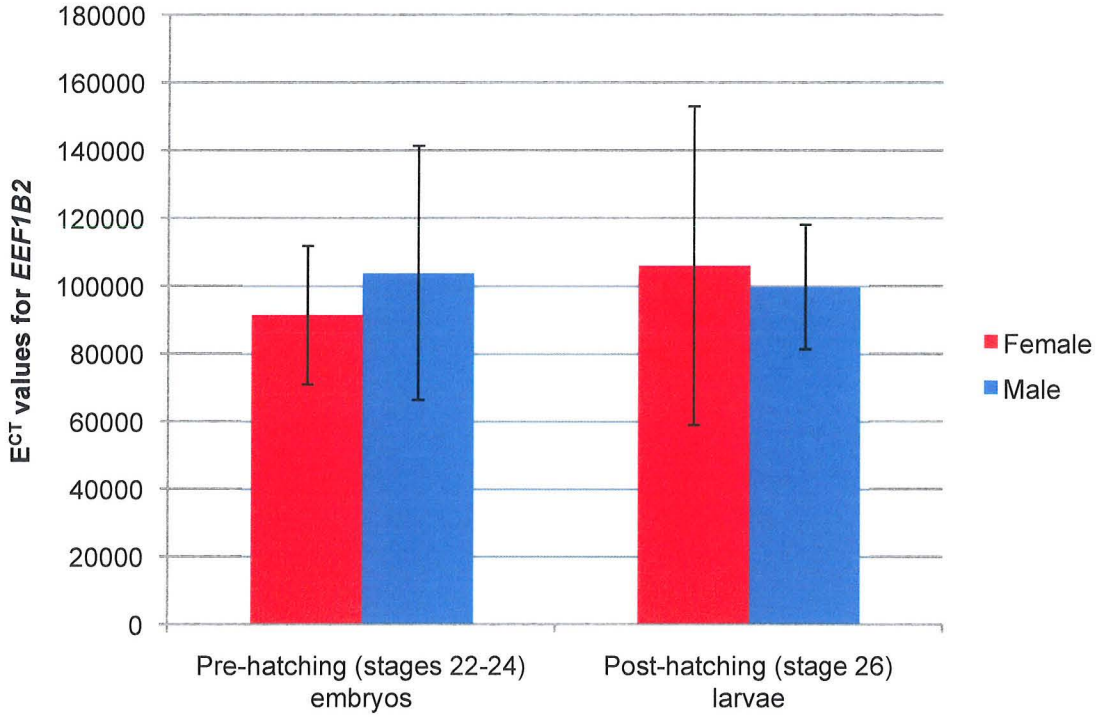
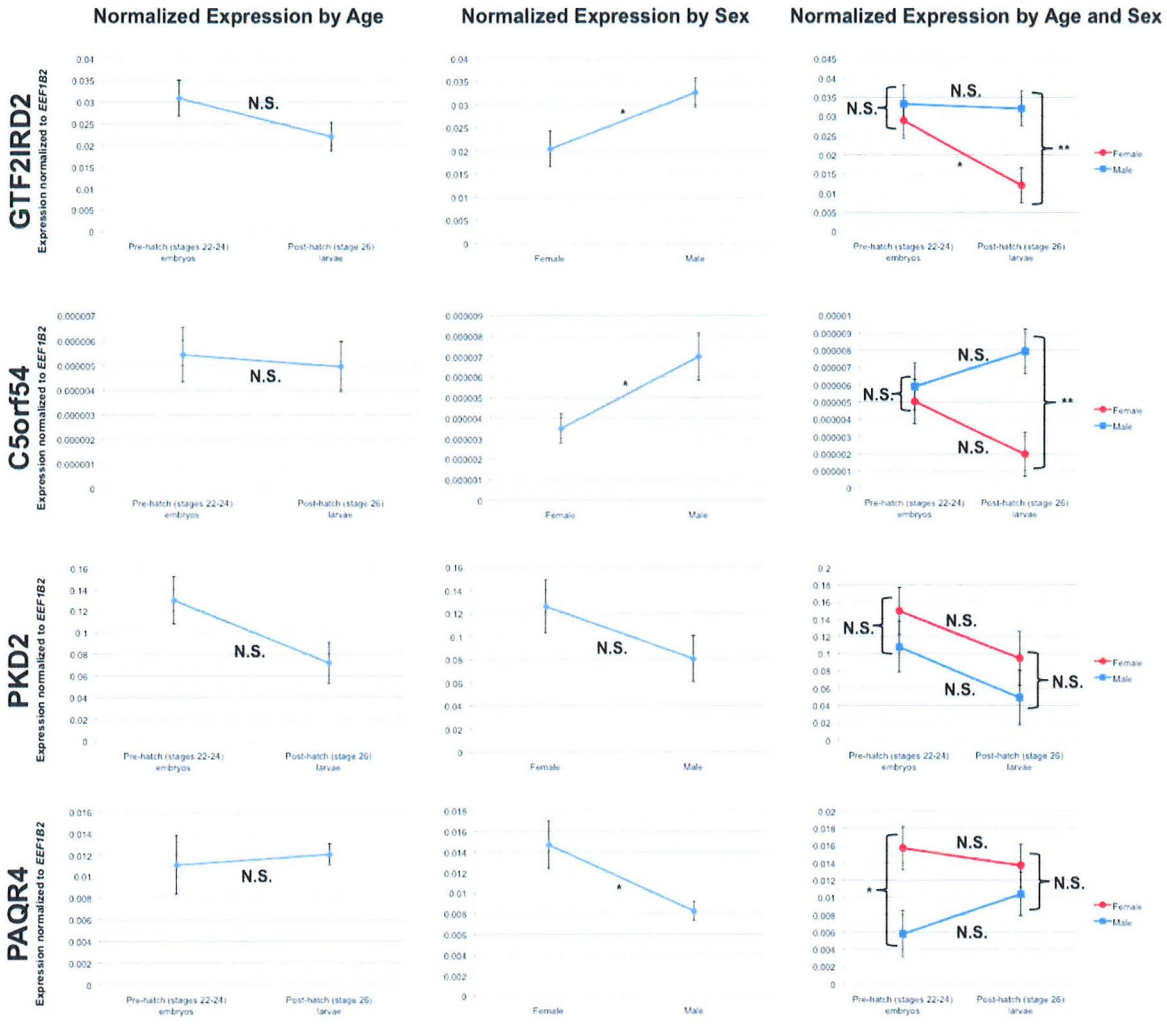


Figure 4.4. Expression of reference gene *EEF1B2*.



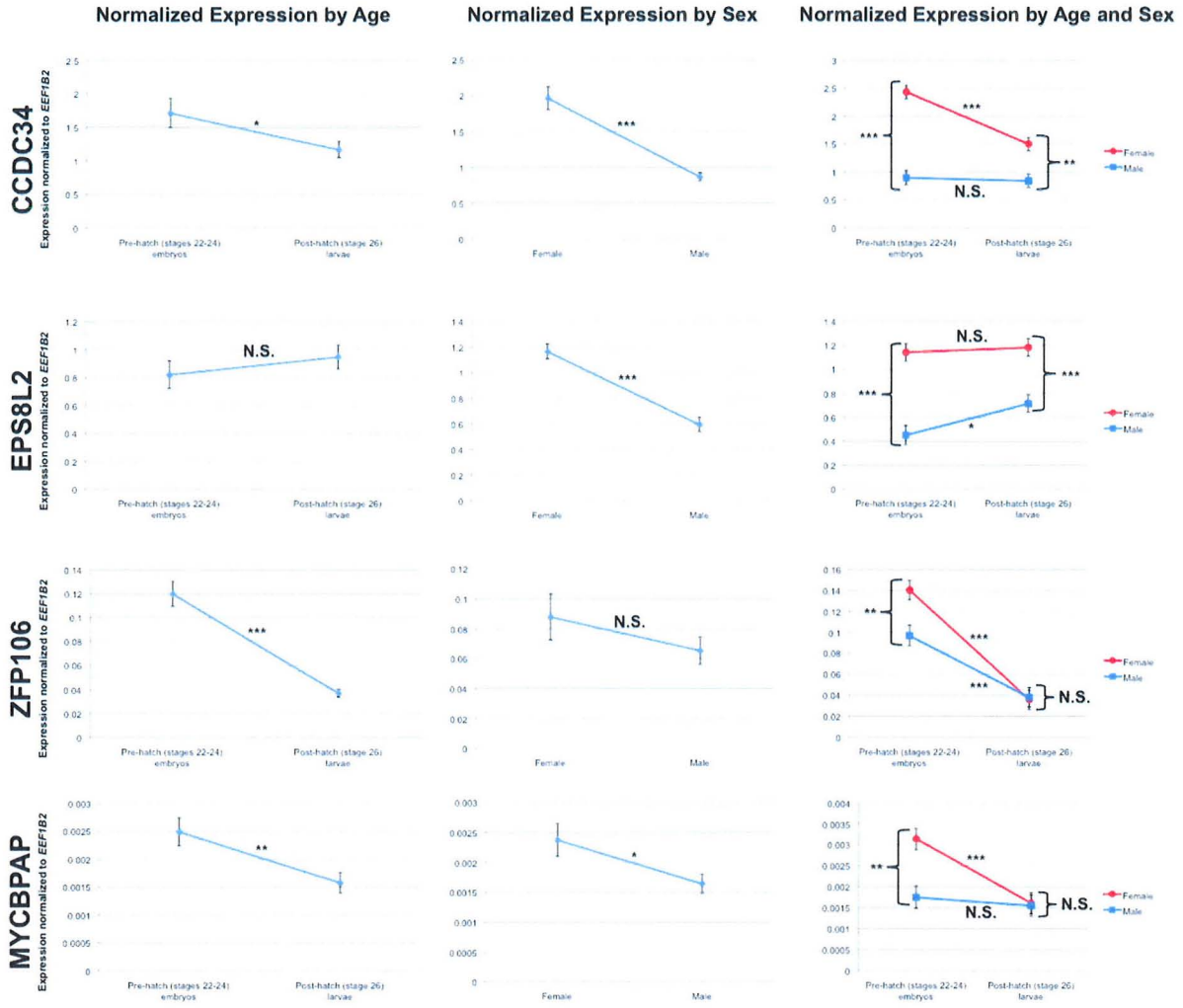
Expression of reference gene *EEF1B2* does not vary significantly between sexes or by age in *G. aculeatus* embryos and larvae (stages 22-24 and 26) by qPCR. Error bars indicate standard deviation.

Figure 4.5. Expression of *GTF2IRD2*, *C5orf54*, *PKD2*, and *PAQR4* by qPCR.



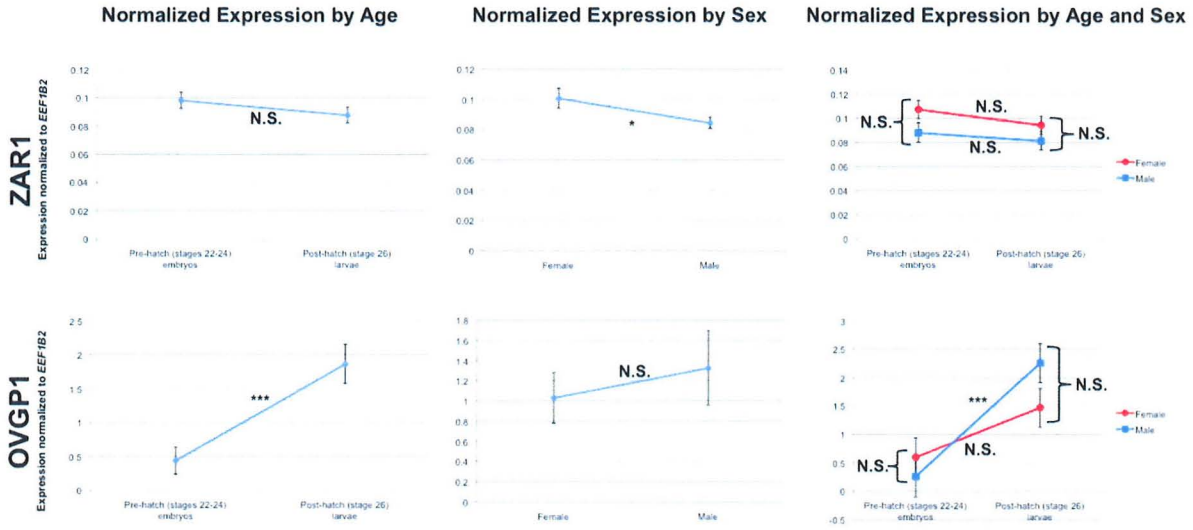
I used independent t-tests to identify significant differences by age (left panels) and sex (middle panels), while I used a two-way ANOVA to identify significant differences by both sex and age. Error bars indicate standard error. I note significance as follows: N.S. (not significant), *p<0.05, **p<0.01, and ***p<0.001.

Figure 4.6. Expression of *CCDC34*, *EPS8L2*, *ZFP106*, *and MYCBPAP* by qPCR.



I used independent t-tests to identify significant differences by age (left panels) and sex (middle panels), while I used a two-way ANOVA to identify significant differences by both sex and age. Error bars indicate standard error. I note significance as follows: N.S. (not significant), *p<0.05, **p<0.01, and ***p<0.001.

Figure 4.7. Expression of *ZAR1* and *OVGP1* by qPCR.



I used independent t-tests to identify significant differences by age (left panels) and sex (middle panels), while I used a two-way ANOVA to identify significant differences by both sex and age. Error bars indicate standard error. I note significance as follows: N.S. (not significant), *p<0.05, **p<0.01, and ***p<0.001.

CHAPTER FIVE

Conclusions and Future Directions

SUMMARY

The experiments I have described in this dissertation have revealed greater sex chromosome diversity in sticklebacks than previously known. Sex chromosome systems have evolved independently at least three times in the stickleback family; in some species, sex chromosome systems even vary among populations. Since there are isolated populations among most stickleback species, there is an even greater potential for sex chromosome diversity in this family. Thus, future studies should look for evidence of additional population differences in sex chromosome morphology or sex chromosome system. My experiments have also revealed some of the earliest differentially expressed genes in the threespine stickleback, *Gasterosteus aculeatus*. The RNA-Seq dataset I have generated will aid in the characterization of the *G. aculeatus* Y chromosome sequence, since I have also identified potentially hundreds of Y chromosome genes that are expressed in male embryos and larvae. Future studies will determine if the *G. aculeatus* Y chromosome does harbor a large number of transcribed transposable element-like loci, in addition to the elusive sex determination gene.

A REVISED VIEW OF SEX CHROMOSOME EVOLUTION IN STICKLEBACKS

How Many Sex Chromosome Systems Have Evolved in the Stickleback Family?

At least three sex chromosome systems have arisen independently in the stickleback family in within the past 40 million years. I based this statement on the major findings from Chapters 2 and 3 of this dissertation. The ZZ-ZW sex chromosome system of the fourspine stickleback (*Apeltes quadracus*) is not related to the chromosome (Chr) 12-based XX-XY sex chromosome system of the ninespine stickleback (*Pungitius pungitius*) or the Chr19-based sex chromosome systems in the genus *Gasterosteus* (Figure 1.1) (Ross *et al.*, 2009). Thus, in at least three times in stickleback history, three different autosomal pairs became nascent sex chromosomes. This event could have happened more than three times, since there have been no surveys for sex chromosomes in the fifteenspine stickleback (*Spinachia spinachia*), and only a cursory screen for sex chromosomes in the brook stickleback, *Culaea inconstans* (Ross *et al.*, 2009). If *C. inconstans* does have a sex chromosome system, it is also likely arose independent of the *Gasterosteus* and *P. pungitius* sex chromosome systems, based on evidence presented in Chapter 2 and Ross *et al.*, 2009. Thus, the potential for between-species sex chromosome diversity in this family is even greater than I can currently conclude. However, we will need additional genetic and cytogenetic surveys to conclude exactly how many times sex chromosome systems have arisen in this family.

Do Sticklebacks Share the Same Sex Determination Gene?

In Chapter 4, I described a screen to identify the sex determination gene in the threespine stickleback, *Gasterosteus aculeatus*. This search is still ongoing. The results of my screen, when combined with the complete sequence of the *G. aculeatus* Y chromosome (which is expected within the next year), should help form a list of candidate sex determination genes for this species. These candidates must be screened through a series of experiments to find the candidate gene with the precise spatial and temporal characteristics of a sex determination gene, as well as the functional role of a sex determination gene. A good candidate gene will be

expressed within the male bipotential gonad prior to the morphological divergence of the sexes. In addition, the candidate gene should be both necessary and sufficient to transform genetic females (XX) into phenotypic males when this candidate is expressed using transgenics prior to the determination of gonad fate (Koopman *et al.*, 1991; Matsuda *et al.*, 2007). Finally, once we know the *G. aculeatus* sex determination gene, we can screen for its presence or absence in other stickleback species, and begin to understand whether the turnover of sex chromosomes in this family is correlated with the turnover of sex determination genes. These experiments will take several years to execute, but I hope the process will begin now, with the data I have presented in Chapter 4.

Despite the wait required to identify and validate candidate sex determination genes, my high-throughput screen has already yielded data that are transforming our view of how the gene content of sex chromosomes has evolved in *G. aculeatus*. I have confirmed that transposable element (TE)-like genes have invaded the *G. aculeatus* Y chromosome (Peichel *et al.*, 2004). Surprisingly, at least some of these TE-like loci are expressed during late embryonic and early larval male development (Figure 4.5) (Peichel, unpublished). I have also confirmed the absence of a global dosage compensation mechanism in *G. aculeatus*. Gene dosage can be regulated on a case-by-case basis, but over 200 X chromosome genes were still expressed at twice the level in females compared to males. These findings should lead to new studies on gene regulation of sex chromosomes, including the regulation of invasive TE-like elements and the imbalance of X chromosome gene dose between the sexes.

FUTURE DIRECTIONS

The Search for the Sex Determination Gene in *G. aculeatus*

The search for the *G. aculeatus* sex determination gene has been an ongoing project in the Peichel laboratory for the past eight years. It initially took the form of positional cloning and "chromosome walking" experiments, before new data on the immense size of the non-recombining region of the Y chromosome made these approaches impractical. Other members of the laboratory have investigated specific

genes on the sex chromosomes to look for diverged Y chromosome alleles. However, this gene-by-gene approach was always limited by its focus on the genes known to be on the X chromosome. We rarely had the opportunity to look first at genes on the Y chromosome itself; we were always searching instead for diverged Y copies of X chromosome genes. Thus, this approach ignored the potentially rich assortment of genes which could have transposed to the Y chromosome following the loss of recombination with the X.

The search for the *G. aculeatus* sex determination gene has been most hampered by the lack of a complete Y chromosome sequence. The *G. aculeatus* sex chromosome pairs are among the largest chromosome pairs in the karyotype (Ross and Peichel, 2008). The X chromosome itself is over 20Mbp in length, with over 1,000 known or predicted protein-coding genes. The non-recombining region corresponds to approximately 16Mbp of X chromosome sequence. Thus, the *G. aculeatus* Y chromosome differs vastly from another model fish to which it is so often compared, the medaka fish (*Oryzias latipes*). The male-specific region of the medaka Y chromosome is less than 0.3Mbp in length, and the sex determination gene *DMY* is the only gene in this region (Matsuda *et al.*, 2002; Nanda *et al.*, 2002). But, *G. aculeatus* is not medaka. All evidence to date suggests that the male-specific region of the *G. aculeatus* Y chromosome is much larger, with a rich and complex assortment of genes. Some genes are Chr19 genes with alleles on the X chromosome, while others are copies of genes potentially transposed from autosomes (including *GTF2IRD2* and *C5orf54* in Figure 4.5; Table 4.5) (Peichel, unpublished). In short, the *G. aculeatus* Y chromosome is simply too large to navigate without a map.

Thus, the search for the *G. aculeatus* sex determination gene must enter the age of genomics and bioinformatics. The Y chromosome sequence is in progress, and will likely be completed within the next year. I believe the best approach to identifying the sex determination gene in this species is to identify predicted genes on the Y chromosome itself. This approach can make use of transcriptome datasets generated by myself (Chapter 4) and Catherine Peichel (unpublished) by RNA-Seq. Both of our RNA-Seq datasets include transcript sequences from expressed Y

chromosome genes, which should aid in mapping the gene content of the Y chromosome. This approach will help navigate the Y chromosome spatially, identifying Y chromosome genes of interest that we would not have known about otherwise, including genes that were transposed to the Y chromosome from autosomes. My RNA-Seq dataset may help further identify candidate sex determination genes, since this dataset includes all genes expressed in males during early stages of gonadal differentiation.

Dosage Compensation in Sticklebacks

G. aculeatus lacks a global mechanism to equalize the dose of X chromosome genes between the sexes (Chapter 4) (Leder *et al.*, 2010; Peichel, unpublished). However, gene-by-gene regulation of gene dose still occurs for some X chromosome genes; this regulation can take the form of equalizing gene expression between the sexes, or even boosting expression in males relative to females (Chapter 4) (Leder *et al.*, 2010; Peichel, unpublished).

While three independent experiments have demonstrated the absence of global dosage compensation in this species, I believe future experiments should explore the patterns of gene-by-gene dosage regulation among X chromosome genes. Both RNA-Seq and microarray expression studies could be used to investigate the spatial and temporal patterns of the X chromosome dosage imbalance. These investigations should examine whether the same sets of genes escape this (possibly default) dosage imbalance across different tissues and ages. I would also like to know if there are trends for the types of genes that do consistently show dosage imbalance over genes that do not.

If there is a small subset of genes that consistently escape dosage imbalance across tissue types and ages, this may indicate a nascent X chromosome specific dosage compensation regulatory mechanism for this handful of genes. If so, it would be interesting to learn the molecular details of this regulatory network, specifically if there are *cis* regulatory elements common to all genes in this network. However, it is equally possible that genes escaping the dosage imbalance could

each have their own tissue and age-specific regulatory elements governing their expression.

In either case, it appears that *G. aculeatus* has joined the ranks of a growing list of sex chromosome-bearing organisms that lack global dosage compensation mechanisms. Most prominent on this list is the chicken, and other recent additions include the silkworm moth and trematode *Schistosoma mansoni* (Ellegren *et al.*, 2007; Vicoso and Bachtrog, 2011; Walters and Hardcastle, 2011). However, these three species have ZZ-ZW sex chromosomes, while *G. aculeatus* has a XX-XY system. To date, we know of no other species of with XX-XY sex chromosomes (with a degenerate Y chromosome) that lacks a global dosage compensation mechanism. As additional studies of dosage compensation in other organisms continue, we will eventually learn whether *G. aculeatus* is an oddity among XX-XY species, or merely the first of many to join the league of species with sex chromosome dosage imbalance.

Diversity of Stickleback Sex Chromosome Systems

In Chapter 3, I described a Connecticut population of *A. quadracus* that lacks the ZZ-ZW heteromorphic sex chromosomes reported in two other populations of this species. Two hypotheses explain this result in the Connecticut population: either this population has homomorphic ZZ-ZW sex chromosomes, or this population lacks the ZZ-ZW pair entirely. In either case, I argue that this population represents an unexpected population-specific diversity of sex chromosome systems in sticklebacks. This diversity either lies in the differential degeneration of the W chromosome (making it identical to the Z in the Connecticut population, but smaller than the W in Massachusetts and Maine populations), or the absence of the ZZ-ZW system entirely in the Connecticut population. Either way, *A. quadracus* now joins the list of stickleback species with population-specific differences in sex chromosomes.

In addition to *A. quadracus*, there are two other stickleback species with population-based differences in sex chromosome architecture or sex chromosome system. While most populations of *G. aculeatus* have an XX-XY sex chromosome

system (Chr19), *G. aculeatus* from the Sea of Japan have a Y-autosome fusion (Chr19 and Chr09), yielding an $X_1X_1X_2X_2$-X_1X_2Y sex chromosome system (Kitano *et al.*, 2009). A *G. wheatlandi* population from Massachusetts has an $X_1X_1X_2X_2$-X_1X_2Y sex chromosome system from a separate Y-autosome fusion (Chr19 and Chr12) (Ross *et al.*, 2009). However, this Y-autosome fusion has not been documented in another *G. wheatlandi* population. In fact, Chen and Reisman (1970) report heteromorphic XX-XY sex chromosomes in *G. wheatlandi* from Reid State Park, Maine, USA. In this population, the Y chromosome is the smallest chromosome in the male karyotype, and males and females have identical chromosome numbers (Chen and Reisman, 1970).

Diversity in sex chromosome systems between populations has not been surveyed properly among sticklebacks, despite the potential role for this diversity in population divergence and speciation. In *G. aculeatus* from the Sea of Japan, for example, at least one divergent behavioral trait in this population has been linked to the neo-X chromosome (Kitano *et al.*, 2009), and this Y-autosome fusion may have occurred when the Sea of Japan was an isolated inland sea approximately 2 million years ago. We do not know of any morphological, ecological, behavioral, or physiological differences between either *G. wheatlandi* or *A. quadracus* populations that may explain these differences in sex chromosome architecture. For *A. quadracus*, I investigated the geologic history of the Connecticut West River and Long Island Sound and found no evidence for a similar period of geographic isolation during the most recent period of glaciation (Lewis and Stone, 1991; Stone *et al.*, 1998). Thus, we currently lack an explanation for reports of sex chromosome diversity for *G. wheatlandi* and *A. quadracus.*

I would encourage future surveys of multiple stickleback populations and species to search for additional evidence of sex chromosome diversity. These surveys should take different forms, depending on the species in question. For the *Gasterosteus* species and *P. pungitius*, we could screen genomic DNA samples from multiple populations to verify the presence of *SEX*-linked Chr19 or Chr12 markers. Those screens would easily identify populations that have a different (non-Chr19 or non-Chr12) sex chromosome system, or lack sex chromosomes

completely. However, more complex cytogenetic surveys are required to identify populations with different degrees of sex chromosome degeneration or sex chromosome rearrangements. Cytogenetic methods are also the primary means of surveying species without *SEX*-linked markers, including *A. quadracus*. Since cytogenetic techniques are more time-consuming, these methods should only be deployed to survey populations where there has already been a reported ecological, morphological, behavioral, or physiological difference, as inspired our previous investigations of the Sea of Japan *G. aculeatus* population (Kitano *et al.*, 2007).

These surveys are superficial in scope, and would merely identify potential population differences in sex chromosome architecture or sex chromosome system. However, these surveys are the necessary first steps to identify populations where sex chromosome diversity is linked to reproductive isolation, which may indicate a role for sex chromosome diversity in speciation. Thus, these surveys should lay the foundation for the next great report, linking sex chromosome divergence between populations to the first stages of the origin of a new species.

FINAL THOUGHTS

This is an exciting time in the field of sex determination and sex chromosomes. Within the past two decades a number of new model sex chromosome systems and sex determination systems have emerged. These systems include a rich diversity of sex chromosomes in different stages of evolution. In medaka and their close relatives, we see rapid turnover of sex determination genes among species with homomorphic sex chromosomes. In sticklebacks, sex chromosome diversity is coupled with sex chromosome complexity, and large heteromorphic sex chromosomes dominate karyotypes. New findings are challenging old paradigms, from the paucity of $Z_1Z_1Z_2Z_2$-Z_1Z_2W systems in fishes to the surprising absence of global dosage compensation in some heteromorphic sex chromosome systems.

With a little luck (and some well-executed experiments), the next decade should bring with it a new generation of sex determination genes in groups like sticklebacks and salmonids, while careful developmental studies should link how

environmental sex determination mechanisms manipulate sexual differentiation pathways. Right on the heels of these discoveries will be new hypotheses to explain the transitions between genetic and environmental sex determination mechanisms.

Thus, though I leave sticklebacks and sex determination behind, I look forward to monitoring this rich and diverse field from the outside. Today's mysteries in the field of sex determination are not easy to address. They are complex issues at the intersection of genetics, development, molecular biology, and evolution. But, in my core, I am an optimist, and I have confidence that patience and careful experiments will prevail. So, I will be interested to see how the answers to today's complex questions spawn new questions for tomorrow's research. All in all, the future looks bright.

REFERENCES

Andreata AA, de Almeida Toledo LF, Oliveira C and Toledo Filho SA (1992). "Choromosome Studies in Hypoptopomatinae (Pisces, Siluriformes, Loricariidae): I. XX/XY Sex Chromosome Hetermomorphism in Pseudotocinclus tietensis." *Cytologia* **57**: 369-372.

Ayala FJ and Coluzzi M (2005). "Chromosome speciation: humans, *Drosophila*, and mosquitoes." *Proceedings of the National Academy of Sciences of the United States of America* **102 Suppl 1**: 6535-6542.

Bachtrog D (2006). "A dynamic view of sex chromosome evolution." *Current Opinion in Genetics and Development* **16**(6): 578-585.

Bachtrog D, Hom E, Wong KM, Maside X and de Jong P (2008). "Genomic degradation of a young Y chromosome in *Drosophila miranda*." *Genome Biology* **9**(2): R30-R30.10.

Bell MA and Foster SA (1994). Introduction to the Evolutionary Bioogy of the Threespine Stickleback. *The Evolutionary Biology of the Threespine Stickleback.* Bell MA and Foster SA. Oxford; New York, Oxford University Press: 1-27.

Bergero R and Charlesworth D (2009). "The evolution of restricted recombination in sex chromosomes." *Trends in Ecology and Evolution* **24**(2): 94-102.

Berta P, Hawkins JR, Sinclair AH, Taylor A, Griffiths BL, Goodfellow PN and Fellous M (1990). "Genetic evidence equating *SRY* and the testis-determining factor." *Nature* **348**(6300): 448-450.

Bertollo LA, Born GG, Dergam JA, Fenocchio AS and Moreira-Filho O (2000). "A biodiversity approach in the neotropical erythrinidae fish, *Hoplias malabaricus*. Karyotypic survey, geographic distribution of cytotypes and cytotaxonomic considerations." *Chromosome Research* **8**(7): 603-613.

Bertollo LA, Oliveira C, Molina WF, Margarido VP, Fontes MS, Pastori MC, Falcao JN and Fenocchio AS (2004). "Chromosome evolution in the erythrinid fish, *Erythrinus erythrinus* (Teleostei: Characiformes)." *Heredity* **93**(2): 228-233.

Britton-Davidian J, Catalan J, da Graca Ramalhinho M, Ganem G, Auffray JC, Capela R, Biscoito M, Searle JB and da Luz Mathias M (2000). "Rapid chromosomal evolution in island mice." *Nature* **403**(6766): 158.

Bruner AM, unpublished.

Bull JJ (1983). *Evolution of Sex Determining Mechanisms*. Menlo Park, Benjamin/Cummings Publishing Company.

Callard GV, Tchoudakova AV, Kishida M and Wood E (2001). "Differential tissue distribution, developmental programming, estrogen regulation and promoter characteristics of cyp19 genes in teleost fish." *Journal of Steroid Biochemistry and Molecular Biology* **79**(1-5): 305-314.

Chan YF, Marks ME, Jones FC, Villarreal G, Jr., Shapiro MD, Brady SD, Southwick AM, Absher DM, Grimwood J, Schmutz J, Myers RM, Petrov D, Jonsson B, Schluter D, Bell MA and Kingsley DM (2010). "Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer." *Science* **327**(5963): 302-305.

Charlesworth B (1978). "Model for evolution of Y chromosomes and dosage compensation." *Proceedings of the National Academy of Sciences of the United States of America* **75**(11): 5618-5622.

Charlesworth D and Charlesworth B (1980). "Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes." *Genetics Research* **35**(2): 205-214.

Charlesworth D, Charlesworth B and Marais G (2005). "Steps in the evolution of heteromorphic sex chromosomes." *Heredity* **95**(2): 118-128.

Cheadle C, Vawter MP, Freed WJ and Becker KG (2003). "Analysis of microarray data using Z score transformation." *Journal of Molecular Diagnostics* **5**(2): 73-81.

Chen TR and Reisman HM (1970). "A comparative chromosome study of the North American species of sticklebacks (Teleostei: Gasterosteidae)." *Cytogenetics* **9**(5): 321-332.

Cioffi MB and Bertollo LA (2010). "Initial steps in XY chromosome differentiation in *Hoplias malabaricus* and the origin of an X(1)X(2)Y sex chromosome system in this fish group." *Heredity* **105**(6): 554-561.

Cioffi MB, Martins C and Bertollo LA (2010). "Chromosome spreading of associated transposable elements and ribosomal DNA in the fish *Erythrinus erythrinus*. Implications for genome change and karyoevolution in fish." *BMC Evolutionary Biology* **10**: 271-279.

Conover DO and Heins SW (1987). "Adaptive variation in environmental and genetic sex determination in a fish." *Nature* **326**(6112): 496-498.

Costa V, Angelini C, De Feis I and Ciccodicola A (2010). "Uncovering the complexity of transcriptomes with RNA-Seq." *Journal of Biomedicine and Biotechnology* **2010**: 853916.

de Almeida-Toledo LF, Daniel-Silva MF, Moyses CB, Fonteles SB, Lopes CE, Akama A and Foresti F (2002). "Chromosome evolution in fish: sex

chromosome variability in *Eigenmannia virescens* (Gymnotiformes: Sternopygidae)." *Cytogenetic and Genome Research* **99**(1-4): 164-169.

de Almeida-Toledo LF, Foresti F, Pequignot EV and Daniel-Silva MF (2001). "XX:XY sex chromosome system with X heterochromatinization: an early stage of sex chromosome differentiation in the Neotropic electric eel *Eigenmannia virescens*." *Cytogenetics and Cell Genetics* **95**(1-2): 73-78.

DiNapoli L and Capel B (2008). "*SRY* and the standoff in sex determination." *Molecular Endocrinology* **22**(1): 1-9.

Ellegren H, Hultin-Rosenberg L, Brunstrom B, Dencker L, Kultima K and Scholz B (2007). "Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes." *BMC Biology* **5**: 40-51.

Enkhmandakh B, Bitchevaia N, Ruddle F and Bayarsaihan D (2004). "The early embryonic expression of TFII-I during mouse preimplantation development." *Gene Expression Patterns* **4**(1): 25-28.

Ezaz T, Quinn AE, Miura I, Sarre SD, Georges A and Marshall Graves JA (2005). "The dragon lizard *Pogona vitticeps* has ZZ/ZW micro-sex chromosomes." *Chromosome Research* **13**(8): 763-776.

Ezaz T, Sarre SD, O'Meally D, Graves JA and Georges A (2009). "Sex chromosome evolution in lizards: independent origins and rapid transitions." *Cytogenetic and Genome Research* **127**(2-4): 249-260.

Fraser JA and Heitman J (2005). "Chromosomal sex-determining regions in animals, plants and fungi." *Current Opinion in Genetics and Development* **15**(6): 645-651.

Godwin J, Luckenbach JA and Borski RJ (2003). "Ecology meets endocrinology: environmental sex determination in fishes." *Evolution and Development* **5**(1): 40-49.

Graham P, Penn JK and Schedl P (2003). "Masters change, slaves remain." *Bioessays* **25**(1): 1-4.

Grasberger H and Bell GI (2005). "Subcellular recruitment by TSG118 and TSPYL implicates a role for zinc finger protein 106 in a novel developmental pathway." *International Journal of Biochemistry and Cell Biology* **37**(7): 1421-1437.

Haag ES and Doty AV (2005). "Sex determination across evolution: connecting the dots." *PLoS Biology* **3**(1): e21.

Hahlbeck E, Katsiadaki I, Mayer I, Adolfsson-Erici M, James J and Bengtsson BE (2004). "The juvenile three-spined stickleback (*Gasterosteus aculeatus* L.) as

a model organism for endocrine disruption II--kidney hypertrophy, vitellogenin and spiggin induction." *Aquatic Toxicology* **70**(4): 311-326.

Hamaguchi S, Toyazaki Y, Shinomiya A and Sakaizumi M (2004). "The XX-XY sex-determination system in *Oryzias luzonensis* and *O. mekongensis* revealed by the sex ratio of the progeny of sex-reversed fish." *Zoological Science* **21**(10): 1015-1018.

Handley LJ, Ceplitis H and Ellegren H (2004). "Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution." *Genetics* **167**(1): 367-376.

Henning F, Moyses CB, Calcagnotto D, Meyer A and de Almeida-Toledo LF (2011). "Independent fusions and recent origins of sex chromosomes in the evolution and diversification of glass knife fishes (*Eigenmannia*)." *Heredity* **106**(2): 391-400.

Herpin A, Braasch I, Kraeussling M, Schmidt C, Thoma EC, Nakamura S, Tanaka M and Schartl M (2010). "Transcriptional rewiring of the sex determining *dmrt1* gene duplicate by transposable elements." *PLoS Genetics* **6**(2): e1000844.

Hibbeler S, Scharsack JP and Becker S (2008). "Housekeeping genes for quantitative expression studies in the three-spined stickleback *Gasterosteus aculeatus*." *BMC Molecular Biology* **9**: 18-27.

Hoffmann AA and Rieseberg LH (2008). "Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation?" *Annual Review of Ecology, Evolution, and Systematics* **39**: 21-42.

Hornaday K, Alexander S and Breden F (1994). "Absence of Repetitive DNA Sequences Associated with Sex Chromosomes in Natural Populations of the Trinidad Guppy (*Poecilia reticulata*)." *Journal of Molecular Evolution* **39**: 431-433.

Hoseman KE, Colosimo PF, Summers BR and Kingsley DH (2004). "A Simple and Efficient Microinjection Protocol for Making Transgenic Sticklebacks." *Behaviour* **141**: 1345-1355.

Iwamatsu T (1999). "Convenient Method for Sex Reversal in a Freshwater Teleost, the Medaka." *Journal of Experimental Zoology* **283**: 210-214.

Kawahara R, Miya M, Mabuchi K, Near TJ and Nishida M (2009). "Stickleback phylogenies resolved: evidence from mitochondrial genomes and 11 nuclear genes." *Molecular Phylogenetics and Evolution* **50**(2): 401-404.

King M (1993). *Species Evolution: the Role of Chromosome Change*. Cambridge; New York, Cambridge University Press.

Kingsley DM and Peichel CL (2007). The Molecular Genetics of Evolutionary Change in Sticklebacks. *Biology of the Three-spined Stickleback.* Östlund-Nilsson S, Mayer I and Huntingford F. Boca Raton, CRC Press: 41-82.

Kingsley DM, Zhu B, Osoegawa K, de Jong P, Schein J, Marra M, Peichel CL, Amemiya C, Schluter D, Balabhadra S, Friedlander B, Cha YM, Dickson M, Grimwood J, Schmutz J, Talbot WS and Myers R (2004). "New genomic tools for molecular studies of evolutionary change in threespine sticklebacks." *Behaviour* **141**: 1331-1344.

Kitano J, Mori S and Peichel CL (2007). "Phenotypic divergence and reproductive isolation between sympatric forms of Japanese threespine sticklebacks." *Biological Journal of the Linnean Society* **91**: 671-685.

Kitano J, Ross JA, Mori S, Kume M, Jones FC, Chan YF, Absher DM, Grimwood J, Schmutz J, Myers RM, Kingsley DM and Peichel CL (2009). "A role for a neo-sex chromosome in stickleback speciation." *Nature* **461**(7267): 1079-1083.

Klinkhardt MB (1998). "Some Aspects of Karyoevolution in Fishes." *Animal Research and Development* **47**: 7-36.

Kondo M, Hornung U, Nanda I, Imai S, Sasaki T, Shimizu A, Asakawa S, Hori H, Schmid M, Shimizu N and Schartl M (2006). "Genomic organization of the sex-determining and adjacent regions of the sex chromosomes of medaka." *Genome Research* **16**(7): 815-826.

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**(5): 416-420.

Koopman P (2005). "Sex determination: a tale of two Sox genes." *Trends in Genetics* **21**(7): 367-370.

Koopman P, Gubbay J, Vivian N, Goodfellow P and Lovell-Badge R (1991). "Male development of chromosomally female mice transgenic for *Sry*." *Nature* **351**(6322): 117-121.

Koopman P, Munsterberg A, Capel B, Vivian N and Lovell-Badge R (1990). "Expression of a candidate sex-determining gene during mouse testis differentiation." *Nature* **348**(6300): 450-452.

Lahn BT and Page DC (1999). "Four evolutionary strata on the human X chromosome." *Science* **286**(5441): 964-967.

Leder EH, Cano JM, Leinonen T, O'Hara RB, Nikinmaa M, Primmer CR and Merila J (2010). "Female-biased expression on the X chromosome as a key step in

sex chromosome evolution in threespine sticklebacks." *Molecular Biology and Evolution* **27**(7): 1495-1503.

Levan A, Fredga K and Sandberg AA (1964). "Nomenclature for Centromeric Position on Chromosomes." *Hereditas* **52**(2): 201-220.

Lewis RS and Stone JR (1991). "Late Quaternary Stratigraphy and Depositional History of the Long Island Sound Basin: Connecticut and New York." *Journal of Coastal Research* **SI**(11): 1-23.

Lewis ZR, McClellan MC, Postlethwait JH, Cresko WA and Kaplan RH (2008). "Female-specific increase in primordial germ cells marks sex differentiation in threespine stickleback (*Gasterosteus aculeatus*)." *Journal of Morphology* **269**(8): 909-921.

Li H and Durbin R (2009). "Fast and accurate short read alignment with Burrows-Wheeler transform." *Bioinformatics* **25**(14): 1754-1760.

Li Y, Hill JA, Yue GH, Chen F and Orban L (2002). "Extensive search does not identify genomic sex markers in *Tetraodon nigroviridis*." *Journal of Fish Biology* **61**: 1314-1317.

Liu Z, Moore PH, Ma H, Ackerman CM, Ragiba M, Yu Q, Pearl HM, Kim MS, Charlton JW, Stiles JI, Zee FT, Paterson AH and Ming R (2004). "A primitive Y chromosome in papaya marks incipient sex chromosome evolution." *Nature* **427**(6972): 348-352.

Lowry DB and Willis JH (2010). "A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation." *PLoS Biology* **8**(9): e1000500.

Lucchesi JC, Kelly WG and Panning B (2005). "Chromatin remodeling in dosage compensation." *Annual Review of Genetics* **39**: 615-651.

Makeyev AV, Erdenechimeg L, Mungunsukh O, Roth JJ, Enkhmandakh B, Ruddle FH and Bayarsaihan D (2004). "*GTF2IRD2* is located in the Williams-Beuren syndrome critical region 7q11.23 and encodes a protein with two TFII-I-like helix-loop-helix repeats." *Proceedings of the National Academy of Sciences of the United States of America* **101**(30): 11052-11057.

Mank JE and Ellegren H (2009). "All dosage compensation is local: gene-by-gene regulation of sex-biased expression on the chicken Z chromosome." *Heredity* **102**(3): 312-320.

Marais GA, Nicolas M, Bergero R, Chambrier P, Kejnovsky E, Moneger F, Hobza R, Widmer A and Charlesworth D (2008). "Evidence for degeneration of the Y chromosome in the dioecious plant *Silene latifolia*." *Current Biology* **18**(7): 545-549.

Marin I and Baker BS (1998). "The evolutionary dynamics of sex determination." *Science* **281**(5385): 1990-1994.

Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, Morin H, Pitrat M, Dogimont C and Bendahmane A (2009). "A transposon-induced epigenetic change leads to sex determination in melon." *Nature* **461**(7267): 1135-1138.

Martinez PA, Ezaz T, Valenzuela N, Georges A and Marshall Graves JA (2008). "An XX/XY heteromorphic sex chromosome system in the Australian chelid turtle *Emydura macquarii*: a new piece in the puzzle of sex chromosome evolution in turtles." *Chromosome Research* **16**(6): 815-825.

Matsuda M (2005). "Sex determination in the teleost medaka, *Oryzias latipes*." *Annual Review of Genetics* **39**: 293-307.

Matsuda M, Nagahama Y, Shinomiya A, Sato T, Matsuda C, Kobayashi T, Morrey CE, 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**(6888): 559-563.

Matsuda M, Sato T, Toyazaki Y, Nagahama Y, Hamaguchi S and Sakaizumi M (2003). "*Oryzias curvinotus* has *DMY*, a gene that is required for male development in the medaka, *O. latipes*." *Zoological Science* **20**(2): 159-161.

Matsuda M, Shinomiya A, Kinoshita M, Suzuki A, Kobayashi T, Paul-Prasanth B, Lau EL, Hamaguchi S, Sakaizumi M and Nagahama Y (2007). "*DMY* gene induces male development in genetically female (XX) medaka fish." *Proceedings of the National Academy of Sciences of the United States of America* **104**(10): 3865-3870.

Mattern MY (2004). "Molecular phylogeny of the Gasterosteidae: the importance of using multiple genes." *Molecular Phylogenetics and Evolution* **30**(2): 366-377.

Matthey R (1949). *Les Chromosomes des Vertébrés*. Lausanne, F. Rouge.

Modi WS and Crews D (2005). "Sex chromosomes and sex determination in reptiles." *Current Opinion in Genetics and Development* **15**(6): 660-665.

Moran P, Martinez JL, Garcia-Vazquez E and Pendas AM (1996). "Sex chromosome linkage of 5S rDNA in rainbow trout (*Oncorhynchus mykiss*)." *Cytogenetics and Cell Genetics* **75**(2-3): 145-150.

Morrish BC and Sinclair AH (2002). "Vertebrate sex determination: many means to an end." *Reproduction* **124**(4): 447-457.

Mortazavi A, Williams BA, McCue K, Schaeffer L and Wold B (2008). "Mapping and quantifying mammalian transcriptomes by RNA-Seq." *Nature Methods* **5**(7): 621-628.

Nachman MW and Searle JB (1995). "Why is the house mouse karyotype so variable?" *Trends in Ecology and Evolution* **10**(10): 397-402.

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M and Snyder M (2008). "The transcriptional landscape of the yeast genome defined by RNA sequencing." *Science* **320**(5881): 1344-1349.

Nakamura D (1989). Evolutionary Aspects of Vertebrate Sex Determination. *Evolutionary Mechanisms in Sex Determination*. Wachtel SS. Boca Raton, CRC Press: 3-14.

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*." *Proceedings of the National Academy of Sciences of the United States of America* **99**(18): 11778-11783.

Nanda I, Volff JN, Weis S, Korting C, Froschauer A, Schmid M and Schartl M (2000). "Amplification of a long terminal repeat-like element on the Y chromosome of the platyfish, *Xiphophorus maculatus*." *Chromosoma* **109**(3): 173-180.

Ocalewicz K, Fopp-Bayat D, Woznicki P and Jankun M (2008). "Heteromorphic sex chromosomes in the ninespine stickleback *Pungitius pungitius*." *Journal of Fish Biology* **73**: 456-462.

Ocalewicz K, Woznicki P, Furgala-Selezniow G and Jankun M (2011). "Chromosomal location of Ag/CMA3-NORs, 5S rRNA an dtelomeric repeats in two stickleback species." *Italian Journal of Zoology* **78**(1): 12-19.

Offenhauser N, Borgonovo A, Disanza A, Romano P, Ponzanelli I, Iannolo G, Di Fiore PP and Scita G (2004). "The eps8 family of proteins links growth factor stimulation to actin reorganization generating functional redundancy in the Ras/Rac pathway." *Molecular Biology of the Cell* **15**(1): 91-98.

Okada E, Yoshimoto S, Ikeda N, Kanda H, Tamura K, Shiba T, Takamatsu N and Ito M (2009). "*Xenopus* W-linked *DM-W* induces *Foxl2* and *Cyp19* expression during ovary formation." *Sexual Development* **3**(1): 38-42.

Okubo K, Takeuchi A, Chaube R, Paul-Prasanth B, Kanda S, Oka Y and Nagahama Y (2011). "Sex differences in aromatase gene expression in the medaka brain." *Journal of Neuroendocrinology* **23**(5): 412-423.

Paar V, Gluncic M, Basar I, Rosandic M, Paar P and Cvitkovic M (2011). "Large tandem, higher order repeats and regularly dispersed repeat units contribute substantially to divergence between human and chimpanzee Y chromosomes." *Journal of Molecular Evolution* **72**(1): 34-55.

Peichel CL, unpublished.

Peichel CL and Mills MG, unpublished.

Peichel CL, Nereng KS, Ohgi KA, Cole BL, Colosimo PF, Buerkle CA, Schluter D and Kingsley DM (2001). "The genetic architecture of divergence between threespine stickleback species." *Nature* **414**(6866): 901-905.

Peichel CL, Ross JA, Matson CK, Dickson M, Grimwood J, Schmutz J, Myers RM, Mori S, Schluter D and Kingsley DM (2004). "The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome." *Current Biology* **14**(16): 1416-1424.

Pfaffl MW (2001). "A new mathematical model for relative quantification in real-time RT-PCR." *Nucleic Acids Research* **29**(9): e45.

Piprek RP (2009). "Genetic mechanisms underlying male sex determination in mammals." *Journal of Applied Genetics* **50**(4): 347-360.

Pomiankowski A, Nothiger R and Wilkins A (2004). "The evolution of the *Drosophila* sex-determination pathway." *Genetics* **166**(4): 1761-1773.

Quinn AE, Georges A, Sarre SD, Guarino F, Ezaz T and Graves JA (2007). "Temperature sex reversal implies sex gene dosage in a reptile." *Science* **316**(5823): 411.

Raymond CS, Shamu CE, Shen MM, Seifert KJ, Hirsch B, Hodgkin J and Zarkower D (1998). "Evidence for evolutionary conservation of sex-determining genes." *Nature* **391**(6668): 691-695.

Rice WR (1987). "Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome." *Genetics* **116**(1): 161-167.

Ross JA and Peichel CL (2008). "Molecular cytogenetic evidence of rearrangements on the Y chromosome of the threespine stickleback fish." *Genetics* **179**(4): 2173-2182.

Ross JA, Urton JR, Boland J, Shapiro MD and Peichel CL (2009). "Turnover of sex chromosomes in the stickleback fishes (Gasterosteidae)." *PLoS Genetics* **5**(2): e1000391.

Schrader O, Ahne R, Fuchs J and Schubert I (1997). "Karyotype analysis of *Helianthus annuus* using Giemsa banding and fluorescence in situ hybridization." *Chromosome Research* **5**(7): 451-456.

Schultheis C, Bohne A, Schartl M, Volff JN and Galiana-Arnoux D (2009). "Sex determination diversity and sex chromosome evolution in poeciliid fish." *Sexual Development* **3**(2-3): 68-77.

Schutt C and Nothiger R (2000). "Structure, function and evolution of sex-determining systems in Dipteran insects." *Development* **127**(4): 667-677.

Searle JB (1993). Chromosomal Hybrid Zones in Eutherian Mammals. *Hybrid Zones and the Evolutionary Process*. Harrison RG. New York, Oxford University Press: 309-353.

Sekido R and Lovell-Badge R (2009). "Sex determination and *SRY*: down to a wink and a nudge?" *Trends in Genetics* **25**(1): 19-29.

Shapiro MD, Summers BR, Balabhadra S, Aldenhoven JT, Miller AL, Cunningham CB, Bell MA and Kingsley DM (2009). "The genetic architecture of skeletal convergence and sex determination in ninespine sticklebacks." *Current Biology* **19**(13): 1140-1145.

Shibata Y, Paul-Prasanth B, Suzuki A, Usami T, Nakamoto M, Matsuda M and Nagahama Y (2010). "Expression of gonadal soma derived factor (*GSDF*) is spatially and temporally correlated with early testicular differentiation in medaka." *Gene Expression Patterns* **10**(6): 283-289.

Smith CA, Roeszler KN, Ohnesorg T, Cummins DM, Farlie PG, Doran TJ and Sinclair AH (2009). "The avian Z-linked gene *DMRT1* is required for male sex determination in the chicken." *Nature* **461**(7261): 267-271.

Stone JR, Schafer JP, London EH, DiGiacomo-Cohen ML, Lewis RS and Thompson WB (1998). "Quaternary Geologicl Map of Connecticut and Long Island Sound Basin." United States Geological Survey, United States Department of the Interior.

Swarup H (1958). "Stages in the development of the stickleback *Gasterosteus aculeatus* (L.)." *Journal of Embryology and Experimental Morphology* **6**(3): 373-383.

Swarup H (1959a). "Production of Triploidy in *Gasterosteus aculeatus* (L)." *Journal of Genetics* **56**(2): 129-142.

Swarup H (1959b). "Effect of Triploidy on the Body Size, General Organization and Cellular Structure in *Gasterosteus aculeatus* (L.)." *Journal of Genetics* **56**(2): 143-155.

Takehana Y, Demiyah D, Naruse K, Hamaguchi S and Sakaizumi M (2007). "Evolution of different Y chromosomes in two medaka species, *Oryzias dancena* and *O. latipes*." *Genetics* **175**(3): 1335-1340.

Tanaka K, Takehana Y, Naruse K, Hamaguchi S and Sakaizumi M (2007). "Evidence for different origins of sex chromosomes in closely related *Oryzias* fishes: substitution of the master sex-determining gene." *Genetics* **177**(4): 2075-2081.

Tang YT, Hu T, Arterburn M, Boyle B, Bright JM, Emtage PC and Funk WD (2005). "PAQR proteins: a novel membrane receptor family defined by an ancient 7-transmembrane pass motif." *Journal of Molecular Evolution* **61**(3): 372-380.

Trapnell C, Pachter L and Salzberg SL (2009). "TopHat: discovering splice junctions with RNA-Seq." *Bioinformatics* **25**(9): 1105-1111.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ and Pachter L (2010). "Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation." *Nature Biotechnology* **28**(5): 511-515.

Uzbekova S, Roy-Sabau M, Dalbies-Tran R, Perreau C, Papillier P, Mompart F, Thelie A, Pennetier S, Cognie J, Cadoret V, Royere D, Monget P and Mermillod P (2006). "Zygote arrest 1 gene in pig, cattle and human: evidence of different transcript variants in male and female germ cells." *Reproductive Biology and Endocrinology* **4**: 12-25.

Valenzuela N (2009). "Co-evolution of genomic structure and selective forces underlying sexual development and reproduction." *Cytogenetic and Genome Research* **127**(2-4): 232-241.

van Doorn GS and Kirkpatrick M (2007). "Turnover of sex chromosomes induced by sexual conflict." *Nature* **449**(7164): 909-912.

van Doorn GS and Kirkpatrick M (2010). "Transitions between male and female heterogamety caused by sex-antagonistic selection." *Genetics* **186**(2): 629-645.

van Ooijen JW, Moer MP, Jansen RC and Maliepaard C (2002). MapQTL 4.0: Software for the Calculation of QTL Positions on Genetic Maps. Wagenenin, the Netherlands, Plant Research International.

van Ooijen JW and Voorips RE (2001). JoinMap 3.0: Software for the Calculation of Genetic Linkage Maps. Wagenenin, the Netherlands, Plant Research International.

Veyrunes F, Waters PD, Miethke P, Rens W, McMillan D, Alsop AE, Grutzner F, Deakin JE, Whittington CM, Schatzkamer K, Kremitzki CL, Graves T, Ferguson-Smith MA, Warren W and Marshall Graves JA (2008). "Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes." *Genome Research* **18**(6): 965-973.

Vicoso B and Bachtrog D (2011). "Lack of global dosage compensation in *Schistosoma mansoni*, a female-heterogametic parasite." *Genome Biology and Evolution* **3**: 230-235.

Volff JN, Zarkower D, Bardwell VJ and Schartl M (2003). "Evolutionary dynamics of the DM domain gene family in metazoans." *Journal of Molecular Evolution* **57 Suppl 1**: S241-S249.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and et al. (1995). "AFLP: a new technique for DNA fingerprinting." *Nucleic Acids Research* **23**(21): 4407-4414.

Walters JR and Hardcastle TJ (2011). "Getting a full dose? Reconsidering sex chromosome dosage compensation in the silkworm, *Bombyx mori*." *Genome Biology and Evolution* **3**: 491-504.

Wang DS, Zhou LY, Kobayashi T, Matsuda M, Shibata Y, Sakai F and Nagahama Y (2010). "*Doublesex*- and *Mab-3*-related transcription factor-1 repression of aromatase transcription, a possible mechanism favoring the male pathway in tilapia." *Endocrinology* **151**(3): 1331-1340.

Wang Z, Gerstein M and Snyder M (2009). "RNA-Seq: a revolutionary tool for transcriptomics." *Nature Reviews Genetics* **10**(1): 57-63.

White MJD (1978). *Modes of Speciation*. San Francisco, W. H. Freeman.

Wilkins AS (1995). "Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway." *Bioessays* **17**(1): 71-77.

Withler RE, McPhail JD and Devlin RH (1986). "Electrophoretic polymorphism and sexual dimorphism in the freshwater and anadromous threespine sticklebacks (*Gasterosteus aculeatus*) of the Little Campbell River, British Columbia." *Biochemical Genetics* **24**(9-10): 701-713.

Wootton RJ (1976). *The Biology of the Sticklebacks*. London; New York, Academic Press.

Yoshimoto S, Ikeda N, Izutsu Y, Shiba T, Takamatsu N and Ito M (2010). "Opposite roles of *DMRT1* and its W-linked paralogue, *DM-W*, in sexual dimorphism of *Xenopus laevis*: implications of a ZZ/ZW-type sex-determining system." *Development* **137**(15): 2519-2526.

Zerbino DR and Birney E (2008). "Velvet: algorithms for de novo short read assembly using de Bruijn graphs." *Genome Research* **18**(5): 821-829.

APPENDIX

Supplementary Table of Differentially Expressed Genes

I designed and executed a next generation RNA-Seq screen to identify genes that are differentially expressed between male and female threespine sticklebacks (*Gasterosteus aculeatus*) just as sexual differentiation begins. I report the major findings of this screen in Chapter 4. This appendix is a supplementary table listing all autosomal and X-chromosome genes that show a significant expression bias between the sexes ($p<0.05$) in two independent RNA-Seq experiments. The 301 female-biased genes and 66 male-biased genes are listed by chromosome. For 10 genes, I also used quantitative PCR to compare expression levels between the sexes (Table 4.4). In this supplementary table, I have highlighted the Ensembl designations of these genes in grey.

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
	01	ENSGACG00000006908	0.0833	0.0477	0.1042	0.0478	1.7481	2.1800
	01	ENSGACG00000006915	0.0855	0.0476	0.0914	0.0463	1.7980	1.9729
Wdr69	01	ENSGACG00000007395	0.2097	0.1105	0.2194	0.1203	1.8979	1.8240
Trpc6	01	ENSGACG00000012155	0.1083	0.2066	0.1081	0.2279	0.5244	0.4743
Slc35f2	01	ENSGACG00000013315	0.0817	0.1502	0.0973	0.1941	0.5438	0.5012
	01	ENSGACG00000013401	0.1181	0.2196	0.0758	0.2168	0.5379	0.3494
Nfkbiz	01	ENSGACG00000013710	0.2374	0.1045	0.2108	0.1257	2.2725	1.6769
Gucy1b2	01	ENSGACG00000015355	0.0809	0.1736	0.0276	0.1450	0.4661	0.1901
Cd59	01	ENSGACG00000015409	0.0662	0.2839	0.0990	0.2910	0.2331	0.3402
Rgs13	03	ENSGACG00000016116	0.1550	0.4157	0.2485	0.6225	0.3729	0.3993
Cyr61	03	ENSGACG00000017235	0.2162	0.4637	0.2079	0.4650	0.4661	0.4472
Aqp1	03	ENSGACG00000017380	0.0052	0.0485	0.0100	0.0389	0.1076	0.2580
	04	ENSGACG00000016919	0.0931	0.1748	0.0821	0.1602	0.5327	0.5124
Clnk	04	ENSGACG00000017564	0.0755	0.1560	0.1022	0.1781	0.4841	0.5741
Slc26a3	04	ENSGACG00000019389	0.1216	0.0669	0.1799	0.0697	1.8180	2.5799
	04	ENSGACG00000019494	0.0759	0.1466	0.0779	0.1568	0.5179	0.4969
	04	ENSGACG00000019864	0.5352	1.0525	0.2574	0.5373	0.5085	0.4791
Retstat	05	ENSGACG00000003399	0.1283	0.2752	0.1337	0.2575	0.4661	0.5190
	05	ENSGACG00000008680	0.0159	0.0729	0.0470	0.0822	0.2185	0.5714
Lrrc18	06	ENSGACG00000007028	1.0647	0.5857	1.3132	0.7987	1.8180	1.6440
Ogdhl	06	ENSGACG00000008169	0.0508	0.0290	0.0326	0.0087	1.7481	3.7265
Ankrd23	06	ENSGACG00000011020	0.8234	1.6860	0.7561	1.3311	0.4883	0.5680
	07	ENSGACG00000018954	0.0731	0.3399	0.2345	0.5664	0.2151	0.4141
Epm2aip1	07	ENSGACG00000018984	0.6017	1.1633	0.6288	1.0611	0.5172	0.5926
Tmem151a	07	ENSGACG00000019708	0.0785	0.6175	0.2769	0.7205	0.1271	0.3843
	07	ENSGACG00000020099	0.0442	0.2213	0.0709	0.3677	0.1998	0.1928

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
	07	ENSGACG00000020240	0.0612	0.1375	0.0897	0.1604	0.4450	0.5590
Unc119	07	ENSGACG00000020475	0.0326	0.0699	0.0209	0.0514	0.4661	0.4065
Prss7	07	ENSGACG00000020914	0.0059	0.0276	0.0114	0.0358	0.2151	0.3194
Barhl2	08	ENSGACG00000004346	0.1766	0.3632	0.1416	0.3420	0.4864	0.4141
Zar1	08	ENSGACG00000014170	0.3482	0.0996	0.2457	0.1398	3.4961	1.7568
Slc22a7	09	ENSGACG00000017814	0.1688	0.0966	0.2112	0.0339	1.7481	6.2286
Mycbpap	09	ENSGACG00000018886	0.1651	0.0590	0.1376	0.0568	2.7969	2.4222
Limd1	10	ENSGACG00000003072	1.5185	0.8026	1.5666	0.8900	1.8920	1.7602
Rxfp4	10	ENSGACG00000003931	0.2541	0.5450	0.2037	0.4008	0.4661	0.5082
Atad4	11	ENSGACG00000005720	0.4451	0.2273	0.4322	0.2553	1.9578	1.6929
Aga	11	ENSGACG00000008055	4.3469	1.9553	3.1475	1.6088	2.2232	1.9564
Accn1	11	ENSGACG00000009558	0.0150	0.0398	0.0147	0.0259	0.3774	0.5699
	11	ENSGACG00000012558	0.5264	0.1506	0.6302	0.3523	3.4961	1.7887
Paqr4	11	ENSGACG00000013570	0.2878	0.5439	0.2900	0.4953	0.5291	0.5856
Ovgp1	12	ENSGACG00000003535	1.5922	0.9211	1.6751	0.9750	1.7286	1.7181
Sema3e	12	ENSGACG00000005697	0.0132	0.0292	0.0092	0.0179	0.4524	0.5160
	12	ENSGACG00000013371	0.0621	0.4443	0.0398	0.4277	0.1398	0.0932
Ydjc	13	ENSGACG00000005098	1.2768	2.4652	1.3102	2.2707	0.5179	0.5770
C2orf39	13	ENSGACG00000007338	0.3453	0.1679	0.1683	0.1030	2.0565	1.6339
Dusp21	13	ENSGACG00000011200	1.3263	2.6728	2.8612	4.9802	0.4962	0.5745
Pla2g1b	13	ENSGACG00000012644	0.8054	0.3839	0.6456	0.3080	2.0977	2.0962
Tal2	13	ENSGACG00000013904	0.1497	0.4013	0.2160	0.5581	0.3729	0.3870
Hspb11	14	ENSGACG00000017813	0.4287	0.0951	0.3888	0.1018	4.5061	3.8197
Hsbp11	14	ENSGACG00000017815	1.3785	0.0704	1.1998	0.2824	19.5783	4.2482
	14	ENSGACG00000018393	0.0870	0.2696	0.1209	0.2080	0.3227	0.5813
Gabrr2	15	ENSGACG00000013033	0.0114	0.0586	0.0073	0.0635	0.1951	0.1157

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Pkd-2	16	ENSGACG00000002153	0.0869	0.6342	0.0673	0.1329	0.1370	0.5066
Q2ped4	16	ENSGACG00000002202	0.7588	1.4564	1.2806	2.3826	0.5210	0.5375
Lypd1	16	ENSGACG00000002758	0.0646	0.0294	0.0528	0.0303	2.1976	1.7390
	16	ENSGACG00000006493	0.0309	0.0568	0.0368	0.0734	0.5438	0.5012
Ankrd44	16	ENSGACG00000008817	0.2387	0.4589	0.2959	0.5233	0.5201	0.5655
	17	ENSGACG00000003551	0.5279	0.9909	0.5925	1.1735	0.5327	0.5049
Abhd6	17	ENSGACG00000003935	0.6706	0.3836	0.4623	0.2404	1.7481	1.9229
Cnih3	18	ENSGACG00000004246	0.0195	0.0428	0.0250	0.0448	0.4560	0.5590
	18	ENSGACG00000004653	0.0705	0.5042	0.2713	0.6067	0.1398	0.4472
C5orf54	18	ENSGACG00000006487	0.0429	4.4461	0.3300	5.5594	0.0096	0.0594
Det1	19/X	ENSGACG00000003152	1.7675	0.7693	1.5717	0.9307	2.2974	1.6887
C16orf57	19/X	ENSGACG00000003257	3.0353	1.3188	2.1561	1.0139	2.3016	2.1265
	19/X	ENSGACG00000003286	0.0742	0.0335	0.1503	0.0695	2.2142	2.1638
	19/X	ENSGACG00000003338	0.0646	0.1316	0.0642	0.1667	0.4907	0.3851
	19/X	ENSGACG00000003348	0.6977	0.3859	0.6403	0.3643	1.8080	1.7575
Olfml1	19/X	ENSGACG00000003368	0.8277	0.4526	1.1319	0.5656	1.8287	2.0013
Mical2	19/X	ENSGACG00000003448	0.8460	0.4457	0.7211	0.3981	1.8982	1.8115
Lingo1	19/X	ENSGACG00000003462	3.8192	1.8599	5.0043	2.5876	2.0534	1.9340
Hmg20a	19/X	ENSGACG00000003466	2.1308	0.8915	1.8032	0.9079	2.3903	1.9861
	19/X	ENSGACG00000003482	7.6132	3.5683	6.5524	3.3605	2.1336	1.9498
Syt12	19/X	ENSGACG00000003487	2.0145	1.0296	1.8553	0.9776	1.9565	1.8979
Cat	19/X	ENSGACG00000003491	11.1774	5.2707	14.7498	6.1552	2.1207	2.3963
	19/X	ENSGACG00000003509	11.8905	5.8050	10.0912	4.9447	2.0483	2.0408
Ptdss2	19/X	ENSGACG00000003546	1.2022	0.5435	1.1122	0.5311	2.2121	2.0941
C21orf110	19/X	ENSGACG00000003573	1.4588	0.8360	1.3265	0.6964	1.7451	1.9047
Avpr1a	19/X	ENSGACG00000003589	1.0541	0.5304	2.0532	0.8287	1.9873	2.4776

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Akr1d1	19/X	ENSGACG00000003639	3.3493	1.7741	3.3678	1.4825	1.8879	2.2717
Trim24	19/X	ENSGACG00000003650	0.9361	0.3816	0.8865	0.4768	2.4530	1.8593
Fbln1	19/X	ENSGACG00000003661	2.5206	1.3452	2.4130	1.3064	1.8737	1.8471
Ppara	19/X	ENSGACG00000003703	1.2815	0.5461	0.9090	0.5198	2.3465	1.7488
Yars2	19/X	ENSGACG00000003720	4.9401	2.3550	5.4374	2.5148	2.0977	2.1621
Mrps35	19/X	ENSGACG00000003757	10.5440	4.8231	5.7367	2.8899	2.1862	1.9851
Tead4	19/X	ENSGACG00000003761	0.6946	0.3353	0.6031	0.3694	2.0712	1.6329
Dennd2a	19/X	ENSGACG00000003792	1.7800	0.8011	1.4553	0.6427	2.2219	2.2643
Pus7	19/X	ENSGACG00000003840	5.2481	2.6685	4.4990	2.4812	1.9667	1.8133
Kiaa1644	19/X	ENSGACG00000003905	0.0894	0.0389	0.1097	0.0558	2.2974	1.9676
Plxnb2	19/X	ENSGACG00000003911	1.5959	0.7727	1.0823	0.5936	2.0652	1.8232
Tubgcp6	19/X	ENSGACG00000003928	1.7021	0.9788	1.5213	0.8238	1.7391	1.8466
Appl2	19/X	ENSGACG00000003947	0.8440	0.4124	0.6315	0.3793	2.0465	1.6650
Nuak1	19/X	ENSGACG00000003957	2.0449	1.1194	1.7319	0.9794	1.8267	1.7682
Mov10l1	19/X	ENSGACG00000003977	0.1900	0.0631	0.1306	0.0740	3.0120	1.7652
Tmem117	19/X	ENSGACG00000004020	0.2161	0.1163	0.2386	0.1202	1.8582	1.9861
Adamts20	19/X	ENSGACG00000004063	0.8120	0.3809	0.9046	0.4345	2.1319	2.0820
Prickle1	19/X	ENSGACG00000004088	1.1085	0.5813	1.1104	0.6662	1.9070	1.6668
Pphln1	19/X	ENSGACG00000004098	1.2913	0.6596	1.3651	0.7489	1.9578	1.8227
Yaf2	19/X	ENSGACG00000004103	3.4847	1.5490	3.9664	2.3773	2.2497	1.6685
Glt8d3	19/X	ENSGACG00000004112	1.0419	0.5631	1.2355	0.7228	1.8503	1.7092
Pdzrn4	19/X	ENSGACG00000004122	0.7653	0.3850	0.5935	0.3037	1.9879	1.9541
Pnpla8	19/X	ENSGACG00000004140	0.6937	0.3526	0.5173	0.2997	1.9673	1.7261
Dnm1l	19/X	ENSGACG00000004145	3.9057	2.1698	3.5391	1.5697	1.8000	2.2546
	19/X	ENSGACG00000004200	26.5735	13.8467	14.1368	8.1025	1.9191	1.7448
Bpgm	19/X	ENSGACG00000004215	5.4150	2.5032	4.6129	2.2592	2.1632	2.0418

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
	19/X	ENSGACG00000004229	69.7580	38.5688	77.6095	40.5403	1.8087	1.9144
	19/X	ENSGACG00000004247	1.4373	0.7019	1.5063	0.7441	2.0477	2.0244
Kcna6	19/X	ENSGACG00000004310	3.7621	2.0959	4.3486	2.1833	1.7950	1.9918
Chchd3	19/X	ENSGACG00000004357	15.0286	7.6909	13.4978	8.0063	1.9541	1.6859
Net1	19/X	ENSGACG00000004380	1.9072	1.0357	1.6616	0.8515	1.8414	1.9513
Dhtkd1	19/X	ENSGACG00000004439	7.1861	3.6674	6.9991	3.9343	1.9595	1.7790
Camk1d	19/X	ENSGACG00000004485	0.5947	0.1772	0.5827	0.2938	3.3563	1.9835
Ucma	19/X	ENSGACG00000004488	10.1462	5.1520	10.1740	4.1901	1.9694	2.4281
Pctk2	19/X	ENSGACG00000004498	0.8600	0.3884	0.4313	0.2281	2.2142	1.8905
	19/X	ENSGACG00000004521	2.6643	1.2984	2.4429	1.4039	2.0521	1.7400
Hal	19/X	ENSGACG00000004528	0.9001	0.3879	1.0912	0.5871	2.3202	1.8588
Ntn4	19/X	ENSGACG00000004555	0.7858	0.4436	0.9615	0.4864	1.7714	1.9769
	19/X	ENSGACG00000004593	23.5787	10.5840	23.8969	9.9720	2.2278	2.3964
Man2c1	19/X	ENSGACG00000004613	2.7838	1.1164	1.9603	1.0521	2.4936	1.8633
Neil1	19/X	ENSGACG00000004670	3.1786	1.6449	3.5406	1.5595	1.9324	2.2703
Commd4	19/X	ENSGACG00000004675	3.5150	1.5897	2.2543	1.2409	2.2111	1.8167
	19/X	ENSGACG00000004691	21.4837	10.5883	17.7557	10.7244	2.0290	1.6556
Stra6	19/X	ENSGACG00000004695	0.5011	0.2067	0.4985	0.1732	2.4240	2.8781
Stoml1	19/X	ENSGACG00000004724	4.2516	1.7510	3.7104	1.8679	2.4281	1.9864
Hexa	19/X	ENSGACG00000004744	4.5742	2.4322	5.0247	2.4626	1.8807	2.0404
Ppcdc	19/X	ENSGACG00000004795	1.5662	0.8095	1.5912	0.7739	1.9348	2.0560
C15orf44	19/X	ENSGACG00000004827	5.7096	3.3032	6.4643	3.4291	1.7285	1.8851
Dennd4a	19/X	ENSGACG00000004867	0.3629	0.2053	0.4450	0.2242	1.7675	1.9849
Megf11	19/X	ENSGACG00000004885	0.0849	0.0469	0.0869	0.0513	1.8116	1.6956
Zwilch	19/X	ENSGACG00000005043	2.9769	1.4745	2.8498	1.2912	2.0189	2.2071
Lctl	19/X	ENSGACG00000005057	1.0786	0.4435	1.3748	0.5414	2.4321	2.5393

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Snapc5	19/X	ENSGACG00000005078	3.4160	1.3112	2.5291	1.1095	2.6053	2.2795
Smad3	19/X	ENSGACG00000005092	1.7785	0.5504	1.1606	0.5966	3.2315	1.9452
Aagab	19/X	ENSGACG00000005119	9.2275	4.4904	5.7489	2.9424	2.0550	1.9538
	19/X	ENSGACG00000005181	3.5867	1.6137	4.0255	1.8432	2.2227	2.1839
Kif23	19/X	ENSGACG00000005226	4.1458	1.9843	3.3792	1.6629	2.0893	2.0321
Ighmbp2	19/X	ENSGACG00000005293	3.4014	1.5962	2.8631	1.5651	2.1310	1.8294
Chid1	19/X	ENSGACG00000005331	5.5181	2.4803	4.9770	2.2812	2.2248	2.1817
	19/X	ENSGACG00000005365	1.2969	0.7519	0.9554	0.5228	1.7248	1.8274
Efcab4a	19/X	ENSGACG00000005399	0.4835	0.2017	0.5039	0.2196	2.3973	2.2947
Them138	19/X	ENSGACG00000005406	3.8701	2.0856	5.1439	2.1236	1.8556	2.4222
C11orf10	19/X	ENSGACG00000005414	23.3848	13.4493	31.6032	12.8037	1.7387	2.4683
Eps8l2	19/X	ENSGACG00000005436	1.2871	0.5974	1.0282	0.6218	2.1544	1.6536
	19/X	ENSGACG00000005442	2.0977	0.9361	1.5476	0.6967	2.2409	2.2214
Syt7	19/X	ENSGACG00000005468	0.2698	0.1572	0.3000	0.1445	1.7163	2.0762
	19/X	ENSGACG00000005483	8.6818	4.3474	10.0057	4.3902	1.9970	2.2791
Slc5a22	19/X	ENSGACG00000005489	1.2743	0.6936	1.1100	0.6328	1.8372	1.7540
	19/X	ENSGACG00000005509	1.2281	0.6147	1.4375	0.7398	1.9978	1.9431
Lrdd	19/X	ENSGACG00000005514	0.5500	0.2596	0.4092	0.1956	2.1189	2.0917
	19/X	ENSGACG00000005541	1.4297	0.8046	1.1155	0.5922	1.7769	1.8837
Athl1	19/X	ENSGACG00000005561	5.0890	2.5766	4.1241	2.4614	1.9750	1.6755
Incenp	19/X	ENSGACG00000005590	11.8943	5.4162	9.5653	4.5221	2.1961	2.1152
Rab3il1	19/X	ENSGACG00000005613	1.9957	1.0212	1.9375	0.9978	1.9542	1.9417
Hps5	19/X	ENSGACG00000005632	2.1275	1.0511	2.1682	1.0030	2.0241	2.1617
Saal1	19/X	ENSGACG00000005737	3.4757	1.6634	3.7643	1.5439	2.0895	2.4381
Cd81	19/X	ENSGACG00000005809	1.2623	0.6497	1.6243	0.7749	1.9430	2.0962
Sigirr	19/X	ENSGACG00000005853	1.9657	1.1185	3.0907	1.6733	1.7575	1.8471

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Ano5	19/X	ENSGACG00000005889	2.7751	1.4343	2.8481	1.5700	1.9347	1.8141
Gas2	19/X	ENSGACG00000005909	0.4205	0.1835	0.4308	0.1754	2.2911	2.4555
Ano3	19/X	ENSGACG00000005915	0.2325	0.0950	0.2111	0.1005	2.4473	2.1010
Fibin	19/X	ENSGACG00000005940	85.2546	47.0416	66.9680	39.1351	1.8123	1.7112
Prc1	19/X	ENSGACG00000005957	5.6142	2.2845	4.7922	2.2775	2.4575	2.1042
Ap4e1	19/X	ENSGACG00000005988	1.7804	0.6996	1.6150	0.9431	2.5447	1.7124
Gnb5	19/X	ENSGACG00000005996	2.8900	1.3111	2.8898	1.4886	2.2043	1.9414
Myo5c	19/X	ENSGACG00000006001	0.7413	0.3928	0.6155	0.3189	1.8871	1.9298
	19/X	ENSGACG00000006025	3.0246	1.2904	3.0585	1.4025	2.3439	2.1807
Rsl24d1	19/X	ENSGACG00000006058	91.6233	40.1084	49.3903	22.3675	2.2844	2.2081
Tcf12	19/X	ENSGACG00000006101	0.6835	0.3070	0.4383	0.2564	2.2261	1.7098
	19/X	ENSGACG00000006110	1.0410	0.5379	0.6920	0.2833	1.9352	2.4423
Mtmr15	19/X	ENSGACG00000006141	1.9755	0.8328	1.5034	0.6935	2.3720	2.1677
	19/X	ENSGACG00000006219	3.6783	1.6847	3.2070	1.8274	2.1833	1.7550
Cep152	19/X	ENSGACG00000006224	0.7940	0.3877	0.5506	0.2740	2.0477	2.0093
Galk2	19/X	ENSGACG00000006232	3.0978	1.7117	3.4881	1.6661	1.8098	2.0936
Parp16	19/X	ENSGACG00000006315	4.0763	1.7846	3.9081	2.0998	2.2841	1.8611
	19/X	ENSGACG00000006340	13.4092	6.6704	13.8033	5.9087	2.0103	2.3361
	19/X	ENSGACG00000006351	3.1427	1.6344	5.2221	1.8848	1.9229	2.7706
Nptn	19/X	ENSGACG00000006370	1.4258	0.7304	1.7810	0.9949	1.9520	1.7901
Cox5a	19/X	ENSGACG00000006516	31.0289	15.6269	12.4771	6.4867	1.9856	1.9235
C11orf17	19/X	ENSGACG00000006521	2.7085	0.9464	1.2831	0.7769	2.8619	1.6515
	19/X	ENSGACG00000006644	0.2493	0.1161	0.2689	0.1083	2.1465	2.4819
Sox6	19/X	ENSGACG00000006649	0.4369	0.2334	0.3853	0.2186	1.8720	1.7628
C11orf58	19/X	ENSGACG00000006659	8.9017	4.4912	5.8993	3.4470	1.9821	1.7114
Plekha7	19/X	ENSGACG00000006687	0.3067	0.1523	0.2557	0.1458	2.0142	1.7540

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Pik3c2a	19/X	ENSGACG00000006738	1.5319	0.8848	1.3981	0.7715	1.7314	1.8121
Nucb2	19/X	ENSGACG00000006759	2.0360	0.9099	0.9521	0.5272	2.2375	1.8059
	19/X	ENSGACG00000006806	2.3550	1.2141	3.0910	1.2725	1.9398	2.4291
Alkbh3	19/X	ENSGACG00000006950	2.1585	0.8646	2.4460	0.9589	2.4966	2.5507
C15orf42	19/X	ENSGACG00000006985	1.5554	0.8176	1.6936	0.8823	1.9025	1.9195
Ckmt1a	19/X	ENSGACG00000006990	3.5947	1.9383	3.8257	1.9674	1.8546	1.9445
Isg20	19/X	ENSGACG00000007047	16.1038	8.4091	20.3579	9.8178	1.9151	2.0736
Rlbp1	19/X	ENSGACG00000007054	0.5317	0.2020	0.5115	0.1239	2.6324	4.1278
Abhd2	19/X	ENSGACG00000007084	0.6890	0.3653	0.7847	0.4293	1.8863	1.8278
Acan	19/X	ENSGACG00000007109	1.2645	0.5470	0.9912	0.5821	2.3117	1.7027
Hisppd2a	19/X	ENSGACG00000007114	0.7505	0.3742	0.5784	0.2908	2.0053	1.9891
	19/X	ENSGACG00000007133	12.0867	4.7764	10.8115	4.1968	2.5305	2.5762
Map1a	19/X	ENSGACG00000007135	2.7925	1.4744	2.1454	1.2242	1.8939	1.7525
Scamp2	19/X	ENSGACG00000007153	2.3744	1.1755	2.3290	1.1895	2.0200	1.9579
Mp1	19/X	ENSGACG00000007166	9.7065	4.8716	11.4215	5.4640	1.9924	2.0903
Psma4	19/X	ENSGACG00000007210	2.5450	1.0894	1.3334	0.6272	2.3362	2.1259
Oaz2	19/X	ENSGACG00000007234	1.4331	0.8257	1.4706	0.8127	1.7357	1.8094
Ppib	19/X	ENSGACG00000007238	75.8598	39.0403	59.3243	31.6828	1.9431	1.8724
Snx22	19/X	ENSGACG00000007247	0.8092	0.3019	0.3610	0.1615	2.6804	2.2359
	19/X	ENSGACG00000007281	0.8807	0.4056	0.6510	0.2911	2.1713	2.2359
	19/X	ENSGACG00000007340	40.6477	18.6522	51.1617	21.5102	2.1792	2.3785
Ireb2	19/X	ENSGACG00000007341	5.2563	2.4965	4.7574	2.7210	2.1055	1.7484
Slc25a44	19/X	ENSGACG00000007371	6.3760	3.7098	8.0767	4.1348	1.7187	1.9533
Wdr61	19/X	ENSGACG00000007380	4.2347	2.3741	3.4149	1.7825	1.7837	1.9158
Snupn	19/X	ENSGACG00000007450	3.0922	1.6115	3.0050	1.5695	1.9188	1.9147
Cspg4	19/X	ENSGACG00000007480	3.4530	1.9753	3.5479	2.1530	1.7481	1.6478

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Lamb4	19/X	ENSGACG00000007499	8.0586	4.2936	6.7871	3.7517	1.8769	1.8091
Nup160	19/X	ENSGACG00000007560	4.4881	2.5923	4.0584	2.1110	1.7313	1.9225
Tbc1d15	19/X	ENSGACG00000007682	2.4181	1.2534	1.4482	0.7486	1.9292	1.9345
	19/X	ENSGACG00000007729	1.0903	0.6027	0.8877	0.4569	1.8090	1.9428
Mphosph6	19/X	ENSGACG00000007779	2.3912	1.3885	2.0294	0.9489	1.7222	2.1387
Nox5	19/X	ENSGACG00000007833	0.3527	0.1238	0.3249	0.1545	2.8487	2.1028
	19/X	ENSGACG00000007854	1.9959	1.0247	1.6247	0.6752	1.9478	2.4060
	19/X	ENSGACG00000007952	3.7936	2.0079	2.4522	1.0453	1.8894	2.3459
Nedd1	19/X	ENSGACG00000007956	3.1110	1.4888	3.1770	1.6886	2.0896	1.8814
Prr5l	19/X	ENSGACG00000008007	0.1649	0.0885	0.2040	0.0743	1.8646	2.7441
Fam96a	19/X	ENSGACG00000008052	2.4626	1.3207	2.5524	0.9460	1.8646	2.6980
Spg11	19/X	ENSGACG00000008057	0.9581	0.5125	1.0360	0.5773	1.8694	1.7945
Znf277	19/X	ENSGACG00000008086	3.8157	1.9043	3.5177	1.8355	2.0037	1.9165
	19/X	ENSGACG00000008153	13.6466	6.9059	10.4447	5.8327	1.9761	1.7907
Dus2l	19/X	ENSGACG00000008159	1.5728	0.6640	1.0355	0.5327	2.3685	1.9438
Tnni2	19/X	ENSGACG00000008321	30.0868	16.7788	20.1940	11.1892	1.7931	1.8048
Tnni2	19/X	ENSGACG00000008333	148.7639	78.9453	101.8275	61.0437	1.8844	1.6681
Lsp1	19/X	ENSGACG00000008376	2.2229	1.0597	0.9072	0.5410	2.0977	1.6769
	19/X	ENSGACG00000008450	3.6138	1.5395	4.5367	2.3599	2.3474	1.9225
	19/X	ENSGACG00000008494	9.2307	4.6454	6.2002	3.1698	1.9871	1.9560
Zdhhc7	19/X	ENSGACG00000008550	3.1384	1.6289	3.1003	1.8585	1.9267	1.6682
C19orf40	19/X	ENSGACG00000008589	1.8969	0.8386	1.4869	0.7162	2.2620	2.0762
C19orf40	19/X	ENSGACG00000008599	3.4143	1.6755	2.9196	1.4338	2.0377	2.0363
Ptprz1	19/X	ENSGACG00000008617	0.5254	0.3056	0.4640	0.2392	1.7191	1.9396
Cadps2	19/X	ENSGACG00000008655	0.2846	0.1622	0.3153	0.1600	1.7546	1.9705
Iqub	19/X	ENSGACG00000008678	0.3553	0.1270	0.2393	0.1427	2.7969	1.6769

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Lmod2	19/X	ENSGACG00000008702	0.8791	0.3553	0.6619	0.3070	2.4742	2.1561
Pot1	19/X	ENSGACG00000008743	0.5708	0.2582	0.4284	0.2428	2.2104	1.7641
	19/X	ENSGACG00000008779	4.7496	1.7790	5.6087	2.6598	2.6698	2.1087
	19/X	ENSGACG00000008783	4.6046	1.8786	5.1008	1.8963	2.4511	2.6899
Pim3	19/X	ENSGACG00000008837	0.5696	0.2102	0.4360	0.1792	2.7095	2.4332
Ftsj1	19/X	ENSGACG00000008843	5.1773	2.4446	4.1695	2.0065	2.1179	2.0780
	19/X	ENSGACG00000008869	1.7679	0.9481	1.3606	0.6085	1.8646	2.2359
	19/X	ENSGACG00000008872	2.6127	1.2619	2.3370	1.1307	2.0704	2.0669
Mupcdh	19/X	ENSGACG00000008907	0.6477	0.2728	0.7113	0.3105	2.3741	2.2909
	19/X	ENSGACG00000008928	3.5453	1.7588	2.8564	1.3601	2.0158	2.1001
Atp2b1	19/X	ENSGACG00000008957	0.0390	0.0223	0.0325	0.0134	1.7481	2.4222
Gtse1	19/X	ENSGACG00000008981	1.7568	0.9703	1.6240	0.8966	1.8105	1.8113
Cftr	19/X	ENSGACG00000009039	0.3805	0.2102	0.3531	0.1571	1.8098	2.2477
Wdr51b	19/X	ENSGACG00000009335	0.2143	0.1221	0.2571	0.0911	1.7562	2.8228
	19/X	ENSGACG00000009373	1.1844	0.6776	0.7501	0.3992	1.7481	1.8791
Cep290	19/X	ENSGACG00000009388	0.6705	0.2902	0.4538	0.2754	2.3108	1.6481
Eif4g2	19/X	ENSGACG00000009454	10.6514	5.1111	8.4204	4.3416	2.0840	1.9395
Lrrk2	19/X	ENSGACG00000009572	1.0425	0.4667	1.0606	0.4913	2.2337	2.1588
Slc2a13	19/X	ENSGACG00000009605	1.2150	0.5198	1.3707	0.6161	2.3377	2.2246
Kif21a	19/X	ENSGACG00000009626	1.9892	1.0137	1.5338	0.8413	1.9623	1.8231
Mgat4c	19/X	ENSGACG00000009714	1.3020	0.4867	1.2335	0.5602	2.6753	2.2020
Rnf141	19/X	ENSGACG00000009717	1.1587	0.5662	1.2509	0.6868	2.0465	1.8212
Ampd3	19/X	ENSGACG00000009729	1.6513	0.7643	1.4442	0.7578	2.1605	1.9056
Swap70	19/X	ENSGACG00000009748	1.7054	0.7423	1.1952	0.5104	2.2974	2.3415
Csk	19/X	ENSGACG00000009794	1.1446	0.6366	1.3749	0.6115	1.7980	2.2486
Adal	19/X	ENSGACG00000009803	15.5190	5.9394	12.2378	5.6551	2.6129	2.1640

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Nat10	19/X	ENSGACG00000009849	10.9333	4.5342	9.6843	4.3183	2.4113	2.2426
Alx4	19/X	ENSGACG00000009885	0.1406	0.0714	0.1222	0.0677	1.9705	1.8059
Tspan18	19/X	ENSGACG00000009892	1.0194	0.4769	0.8541	0.3468	2.1374	2.4628
Cd82	19/X	ENSGACG00000009906	1.0310	0.5469	0.9773	0.4160	1.8853	2.3494
TP53i11	19/X	ENSGACG00000009918	0.3092	0.1273	0.1863	0.0591	2.4289	3.1506
	19/X	ENSGACG00000009927	8.6696	3.2108	6.8822	3.4344	2.7001	2.0039
Reln	19/X	ENSGACG00000009959	2.5269	1.4737	2.5004	1.3819	1.7147	1.8094
Plekha5	19/X	ENSGACG00000010014	0.6396	0.3350	0.6016	0.3218	1.9095	1.8694
Gpr22	19/X	ENSGACG00000010047	1.1029	0.5144	1.2917	0.4126	2.1443	3.1303
Nt5dc3	19/X	ENSGACG00000010077	4.6617	2.2897	3.5702	2.0446	2.0359	1.7461
Kcnd2	19/X	ENSGACG00000010094	0.2566	0.1084	0.2438	0.1311	2.3666	1.8595
Flnc	19/X	ENSGACG00000010114	11.4345	6.1562	9.0375	5.1996	1.8574	1.7381
Fgd6	19/X	ENSGACG00000010186	0.5157	0.1990	0.4095	0.1691	2.5912	2.4222
Fbxl22	19/X	ENSGACG00000010198	7.2734	3.5191	9.1321	4.8264	2.0668	1.8921
	19/X	ENSGACG00000010201	1.8554	1.0430	1.5441	0.8798	1.7789	1.7551
Lrrc61	19/X	ENSGACG00000010239	1.7607	0.9653	2.9862	1.2121	1.8241	2.4636
Idh3a	19/X	ENSGACG00000010244	2.6816	1.4456	2.9017	1.3727	1.8551	2.1138
Dnaja4	19/X	ENSGACG00000010346	72.4621	35.1211	65.6927	31.2630	2.0632	2.1013
Rab8b	19/X	ENSGACG00000010417	1.9821	1.1436	2.0657	1.0983	1.7333	1.8808
Kti12	19/X	ENSGACG00000010443	2.6898	1.5361	2.5637	1.5538	1.7511	1.6499
	19/X	ENSGACG00000010453	4.7102	2.1478	3.9183	1.7916	2.1930	2.1870
Tspan3	19/X	ENSGACG00000010459	13.4750	5.9786	14.5301	6.0680	2.2539	2.3946
Rcn2	19/X	ENSGACG00000010496	3.5597	1.6136	2.4460	1.3128	2.2060	1.8633
Scaper	19/X	ENSGACG00000010509	0.6695	0.3774	0.5294	0.2601	1.7739	2.0350
Etfa	19/X	ENSGACG00000010550	17.6959	8.6824	18.2493	7.9930	2.0381	2.2832
Lactb	19/X	ENSGACG00000010620	4.5732	2.5130	4.5656	2.5718	1.8198	1.7752

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Fam148a	19/X	ENSGACG00000010669	0.5541	0.2882	0.4523	0.2312	1.9229	1.9564
Narg2	19/X	ENSGACG00000010682	2.2136	1.0731	2.2917	1.1034	2.0629	2.0769
Anxa2	19/X	ENSGACG00000010689	16.9845	6.9026	10.7535	5.0969	2.4606	2.1098
St8sia2	19/X	ENSGACG00000010746	0.6584	0.2608	0.6505	0.3317	2.5240	1.9607
Pex11a	19/X	ENSGACG00000010771	1.3308	0.7258	1.8517	1.0869	1.8335	1.7035
Fkbp8	19/X	ENSGACG00000010785	0.0283	0.1127	0.0570	0.1113	0.2510	0.5124
Blm	19/X	ENSGACG00000010863	4.4099	2.1818	4.5869	2.5582	2.0212	1.7931
Ccdc34	19/X	ENSGACG00000010963	6.9858	2.0057	3.6996	1.3358	3.4829	2.7695
Cecr5	19/X	ENSGACG00000010978	8.8876	4.4983	8.1613	4.2709	1.9758	1.9109
Ccdc77	19/X	ENSGACG00000010988	1.6423	0.7766	1.1892	0.4711	2.1147	2.5244
Ush1c	19/X	ENSGACG00000011004	1.9527	0.8616	1.7511	0.8677	2.2663	2.0180
Ptprj	19/X	ENSGACG00000011062	1.7867	0.8161	2.1201	0.9030	2.1893	2.3477
Th	19/X	ENSGACG00000011104	0.1131	0.0368	0.1846	0.0944	3.0766	1.9564
Ccdc34	19/X	ENSGACG00000011173	3.8144	1.2360	3.2873	1.9489	3.0862	1.6867
Rab19	19/X	ENSGACG00000011175	2.1023	1.0088	1.8450	0.9204	2.0840	2.0046
Brsk2	19/X	ENSGACG00000011262	0.0843	0.1985	0.0578	0.1326	0.4247	0.4359
C15orf58	19/X	ENSGACG00000011544	5.3080	2.1590	6.5585	3.1009	2.4586	2.1150
Plekhg7	19/X	ENSGACG00000011708	0.1723	0.0719	0.1658	0.0989	2.3973	1.6769
	19/X	ENSGACG00000011725	1.7887	0.7106	1.5189	0.7126	2.5172	2.1316
	19/X	ENSGACG00000011784	4.3729	2.1183	6.8454	3.8708	2.0644	1.7685
	19/X	ENSGACG00000011879	1.4938	0.8105	1.7270	0.8306	1.8430	2.0791
	19/X	ENSGACG00000011906	1.2049	0.5875	1.3598	0.6777	2.0511	2.0064
Chrm2	19/X	ENSGACG00000011914	1.7002	0.8448	1.7557	0.5786	2.0124	3.0344
Lrrc17	19/X	ENSGACG00000012034	1.5153	0.8148	1.2882	0.6738	1.8598	1.9117
	19/X	ENSGACG00000012049	0.3952	0.1641	0.4415	0.2559	2.4084	1.7248
	19/X	ENSGACG00000012110	13.1699	25.0856	9.2685	19.9242	0.5250	0.4652

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Cpt1b	19/X	ENSGACG00000012306	4.4887	2.3284	4.2063	1.9413	1.9279	2.1668
Chkb	19/X	ENSGACG00000012349	6.9423	3.2038	9.8128	3.9099	2.1669	2.5098
Shank3	19/X	ENSGACG00000012458	0.0692	0.1638	0.0729	0.1546	0.4226	0.4718
Ripk3	19/X	ENSGACG00000012580	4.9441	2.7036	4.9294	2.9793	1.8287	1.6546
	19/X	ENSGACG00000012826	0.0877	0.1630	0.0337	0.1006	0.5379	0.3354
Kcp	19/X	ENSGACG00000012835	1.2970	0.7570	1.2465	0.6629	1.7134	1.8802
Miox	19/X	ENSGACG00000012900	0.3860	0.0812	0.1748	0.0912	4.7547	1.9165
Lmf2	19/X	ENSGACG00000012928	5.7741	2.7059	5.4555	2.8865	2.1339	1.8900
B4galnt3	19/X	ENSGACG00000012971	1.0809	0.5251	1.0775	0.5560	2.0585	1.9378
C12orf64	19/X	ENSGACG00000013089	0.0228	0.0490	0.0059	0.0367	0.4661	0.1597
Cd9	19/X	ENSGACG00000013122	9.7189	4.5804	10.0906	5.2122	2.1219	1.9360
C15orf26	19/X	ENSGACG00000013167	0.4444	0.1144	0.3762	0.1428	3.8846	2.6352
	19/X	ENSGACG00000013175	1.1139	0.4978	1.4883	0.6124	2.2375	2.4303
Lysmd4	19/X	ENSGACG00000013238	3.0950	1.5996	2.5745	1.3009	1.9348	1.9791
	19/X	ENSGACG00000013258	1.9901	0.9900	2.0810	0.8439	2.0103	2.4661
	19/X	ENSGACG00000013321	2.4578	1.3603	2.3599	1.3701	1.8068	1.7224
Cry1	19/X	ENSGACG00000013480	2.4593	1.2425	2.0573	1.2421	1.9793	1.6562
	19/X	ENSGACG00000013483	38.9986	21.2962	35.4363	19.3921	1.8312	1.8274
Chst1	19/X	ENSGACG00000013487	0.5397	0.1835	0.5845	0.2690	2.9416	2.1726
	19/X	ENSGACG00000013544	0.1965	0.3736	0.1776	0.3382	0.5259	0.5251
Ttc38	19/X	ENSGACG00000013552	3.8466	2.1199	4.2373	2.1458	1.8145	1.9747
Efcab4b	19/X	ENSGACG00000013617	0.3968	0.1951	0.2704	0.1280	2.0341	2.1117
Wfdc1	19/X	ENSGACG00000013768	6.8134	3.4777	5.1690	2.7110	1.9591	1.9067
Hsbp1	19/X	ENSGACG00000013784	48.0075	26.5161	45.7401	27.5255	1.8105	1.6617
Top1	19/X	ENSGACG00000013788	1.0041	0.3656	1.0428	0.4360	2.7465	2.3917
Anapc13	19/X	ENSGACG00000013894	4.6656	2.1448	6.5188	2.4375	2.1754	2.6743

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Arl2bp	19/X	ENSGACG00000013899	1.3778	0.6072	1.1508	0.4779	2.2692	2.4079
Pllp	19/X	ENSGACG00000013906	1.9182	1.0267	2.4152	1.0734	1.8684	2.2500
B3gnt9	19/X	ENSGACG00000013963	0.6097	2.0346	0.6126	2.1102	0.2997	0.2903
Cbfb	19/X	ENSGACG00000013972	0.7316	0.4074	0.3674	0.2125	1.7958	1.7286
	19/X	ENSGACG00000013996	1.9074	0.9651	1.8062	0.8019	1.9764	2.2524
	19/X	ENSGACG00000014068	5.4609	2.8698	6.7343	2.8489	1.9029	2.3638
Tmco7	19/X	ENSGACG00000014081	1.7758	1.0167	2.1438	1.1086	1.7466	1.9338
Dyx1c1	19/X	ENSGACG00000014118	0.6934	0.3110	0.6445	0.2832	2.2300	2.2758
Enpp2	20	ENSGACG00000004492	0.0470	0.0953	0.0603	0.1079	0.4936	0.5590
Adcy2	20	ENSGACG00000006520	0.0402	0.0738	0.0393	0.0763	0.5452	0.5148
	20	ENSGACG00000007546	0.0924	0.1839	0.1092	0.1941	0.5027	0.5624
	20	ENSGACG00000007557	0.0788	0.1508	0.0920	0.1611	0.5229	0.5713
Scrt1	20	ENSGACG00000008982	0.0441	0.1155	0.0659	0.1264	0.3814	0.5217
Cd226	21	ENSGACG00000002708	0.1854	0.3535	0.2378	0.4254	0.5244	0.5590
	Un	ENSGACG00000000119	0.2847	0.7464	0.3651	0.8166	0.3814	0.4472
Pif-1	Un	ENSGACG00000000172	0.1134	0.2121	0.1287	0.2202	0.5347	0.5844
	Un	ENSGACG00000000320	0.0383	0.2193	0.0984	0.1870	0.1748	0.5261
Sema5b	Un	ENSGACG00000000378	0.1223	0.2872	0.1568	0.3106	0.4256	0.5049
Pif-1	Un	ENSGACG00000000540	0.1303	0.3356	0.1616	0.3340	0.3885	0.4839
Zfp106	Un	ENSGACG00000000693	0.2297	0.1210	0.2404	0.1387	1.8979	1.7328
	Un	ENSGACG00000000796	0.2342	0.1288	0.3697	0.2274	1.8180	1.6261
	Un	ENSGACG00000000900	0.0530	0.4170	0.0340	0.3345	0.1271	0.1016
Tfr2	Un	ENSGACG00000001165	0.2267	0.0648	0.2132	0.1300	3.4961	1.6397
	Un	ENSGACG00000001482	0.0422	0.3316	0.0270	0.3870	0.1271	0.0699
Chek2	Un	ENSGACG00000010772	0.0995	0.2609	0.1914	0.3996	0.3814	0.4791
Gtf2ird2	Un	ENSGACG00000013474	0.0975	1.1509	0.3128	1.3990	0.0848	0.2236

			Normalized Number of Reads (RPKM)				Fold Difference	
			Flow Cell 1		*Flow Cell 2*		*Flow Cell 1*	*Flow Cell 2*
Gene	**Chr**	**Ensembl ID**	**Female**	**Male**	**Female**	**Male**	**Female:Male**	**Female:Male**
Kcna6	Un	ENSGACG00000015703	0.2807	0.6690	0.3300	0.6708	0.4195	0.4919
Eps8l2	Un	ENSGACG00000018690	1.7799	0.7487	1.2086	0.7408	2.3774	1.6316
Fads2	Un	ENSGACG00000018692	1.0333	0.5199	0.5890	0.2415	1.9873	2.4392

JAMES RALPH URTON

Curriculum Vitae

Research Experience and Education

2011- **University of Washington; Seattle, WA**
Department of Genome Sciences
Postdoctoral Fellow in the laboratory of Dr. Christine Queitsch.
Genetic and molecular mechanisms of plasticity, robustness, and evolution in the thale cress Arabidopsis thaliana.

2005-2011 **University of Washington; Seattle, WA**
Fred Hutchinson Cancer Research Center; Seattle, WA
Graduate Program in Molecular and Cellular Biology
Doctor of Philosophy, Molecular and Cellular Biology
- Advisor: Dr. Catherine Peichel, Human Biology Division
- Dissertation Title: Evolution of Sex Chromosomes and Sex Determination Mechanisms in Stickleback Fishes (Gasterosteidae)
- Ph.D. conferred August 2011

2003-2005 **Fred Hutchinson Cancer Research Center; Seattle, WA**
Basic Sciences Division
Research Technician I in the laboratory of Dr. Meng-Chao Yao.
Programmed DNA rearrangement in the ciliate Tetrahymena thermophila.

2002 **Fred Hutchinson Cancer Research Center; Seattle, WA**
Basic Sciences Division
Summer undergraduate intern in the laboratory of Dr. Meng-Chao Yao.
Chromatin organization in the ciliate Tetrahymena thermophila.

2001 **University of Iowa; Iowa City, IA**
Department of Microbiology
Summer undergraduate intern in the laboratory of Dr. David Weiss.
Coordination of cell wall synthesis and binary fission in Escherichia coli.

1999-2003 **Augustana College; Rock Island, IL**
Bachelor of Arts, *summa cum laude*
Major: Biology
Minors: Geology, Chemistry

Publications

Urton JR, McCann SR, and Peichel CL. "Karyotype Differentiation Between Two Stickleback Species (Gasterosteidae)." *Cytogenetic and Genome Research*. In press.

Ross JA, **Urton JR**, Boland J, Shapiro MD, and Peichel CL. "Turnover of Sex Chromosomes in the Stickleback Fishes (Gasterosteidae)." *PLoS Genetics*. 2009; 5(2) e1000391.

Oral Presentations

2010 **Urton JR**, Balcells R, and Peichel CL.
"The Search for the Sex Determination Gene in the Threespine Stickleback."
Evolution 2010; Portland, OR

2010 **Urton JR**
"Searching for the Sex Determination Gene in the Threespine Stickleback"
Friday Night Seminar Series, Basic Sciences Division & Human Biology Division, Fred Hutchinson Cancer Research Center; Seattle, WA

2010 **Urton JR**, Bruner AM, McCann SR, Balcells R, and Peichel CL.
"The Evolution of Sex Determination in Stickleback Fishes."
Society for Integrative and Comparative Biology 2010 Annual Meeting; Seattle, WA

2009 **Urton JR**, Ross JA, Kitano J, Bruner AM, McCann SR, and Peichel CL.
"Evolution of Sex Chromosomes and Sex Determination in Stickleback Fishes."
Fifth International Symposium on the Biology of Vertebrate Sex Determination; Kona, HI

2008 **Urton JR**, Ross JA, Boland J, Shapiro MD, and Peichel CL.
"Evolution of Sex Determination in Sticklebacks."
Evolution 2008; Minneapolis, MN

2008 **Urton JR**
"Evolution of Sex Determination in Stickleback Fishes"
Genetics and Genomics Group Meeting, Fred Hutchinson Cancer Research Center; Seattle, WA

Funding, Awards, and Honors

2006-2008	University of Washington: Genome Training Grant
2005	NSF Graduate Research Fellowship: Honorable Mention
2003	Phi Beta Kappa: Zeta Chapter of Illinois
2002	Lincoln Academy of Illinois: Student Laureate Award
1999-2003	Augustana College: Presidential Scholarship

Teaching and Mentoring Experience

2010 **Fred Hutchinson Cancer Research Center; Seattle, WA**
Mentored a local high school student in the laboratory for five weeks.

2009 **Fred Hutchinson Cancer Research Center; Seattle, WA**
Mentored an undergraduate intern in the laboratory for ten weeks.

2007 **Fred Hutchinson Cancer Research Center; Seattle, WA**
Science Education Partnership Program
High school curriculum development program on protein structure.

2007 **Milwaukee School of Engineering; Milwaukee, WI**
Center for BioMolecular Modeling
Participant in the inaugural PALM (Proteins in Active Learning Modules) Workshop for current and future science educators.

2007 **University of Washington; Seattle, WA**
Department of Biology
Teaching Assistant for BIOL200: Introduction to Cell, Molecular, and Developmental Biology.

Public Outreach

2011 **Fred Hutchinson Cancer Research Center; Seattle, WA**
Public presentation for the Science Friday Seminar Series
Title: "Sticklebacks and Sex Determination: the W, X, Y, and Z of it!"

2010 **Texada Stickleback Group; Gillies Bay, BC, Canada**
Public presentation for the 2010 Texada Stickleback Symposium
Title: "How sticklebacks determine sex, and Y you should care."

2010 **Fred Hutchinson Cancer Research Center; Seattle, WA**
Public presentation for the Science Friday Seminar Series
Title: "What do sticklebacks have to do with cancer?"

2009-2010 **Centers for Ocean Sciences Education Excellence: Ocean Learning Communities; Seattle, WA**
Poster sessions for the general public at the Seattle Aquarium

2009-2011 **Fred Hutchinson Cancer Research Center; Seattle, WA**
Presented ongoing research in the Peichel laboratory to new FHCRC employees at orientation meetings

2007-2008 **Northwest Association for Biomedical Research; Seattle, WA**
Mentor, 2008 Student Biotech Expo

Other Graduate School Activities

2009-2010 **University of Washington; Seattle, WA**
Committee for the Biosciences Career Seminar Series

2008-2010 **Fred Hutchinson Cancer Research Center; Seattle, WA**
Organizer, FHCRC Graduate Student Meetings