Turun yliopisto
University of Turku

# MOLECULAR MECHANISMS OF SEXUAL DIMORPHISM IN THREESPINE STICKLEBACK

Heidi Viitaniemi

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
Sarja - ser. A II osa - tom. 297 | Biologica - Geographica - Geologica | Turku 2015

## University of Turku

Faculty of Mathematics and Natural Sciences
Department of Biology
Laboratory of Genetics
Biology
Evolutionary Biology


## Supervised by

Academy Researcher Erica Leder
University of Turku
Finland

## Reviewed by

Professor Outi Savolainen
University of Oulu
Finland

Doctor Christoph Haag
Centre D'Ecologie Fonctionelle & Evolutive,
Montpellier
France

## Opponent

Associate Professor Bengt Hansson
Lund University
Sweden

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.


ISBN 978-951-29-5966-2 (PRINT)
ISBN 978-951-29-5967-9 (PDF)
ISSN 0082-6979
Painosalama Oy - Turku, Finland 2015

# ABSTRACT

Sexual dimorphism is commonly understood as differences in external features, such as morphological features or coloration. However, it can more broadly encompass behavior and physiology and at the core of these differences is the genetic mechanism – mRNA and protein expression. How, and which, molecular mechanisms influence sexually dimorphic features is not well understood thus far. DNA, RNA and proteins are the template required to create the phenotype of an individual, and they are connected to each other via processes of transcription and translation. As the genome of males and females are almost identical with the exception of the few genes on the sex chromosome or the sex-determining alleles (in the case of organisms without sex chromosomes), it is likely that many of the downstream processes resulting in sexual dimorphism are produced by changes in gene regulation and result from a regulatory cascade and not from a vastly different gene composition. Thus, in this thesis a systems biology approach is used to understand sexual dimorphism at all molecular levels and how different genomic features, e.g. sex chromosome evolution, can affect the interplay of these molecules.

The threespine stickleback, *Gasterosteus aculeatus*, is used as the model to investigate molecular mechanisms of sexual dimorphism. It has well-characterized ecology and behavior, especially in the breeding season when sexual dimorphism is high. Moreover, threespine stickleback has a recently evolved Y chromosome in the early stages of sex chromosome evolution, characterized by a lack of recombination leading to degeneration (i.e. gene loss).

The aim of my thesis is to investigate how the genotype links to the molecular phenotype and relates to differences in molecular expression between males and females. Based on previous research on sex differences in mRNA expression, I investigated sex-biased protein expression in adult fish outside the breeding season to see if differences persisted after translation. As sex-biased expression also prevailed in the proteome and previous transcription expression seemed to be related to the sex chromosomes, I investigated the genome level with a particular focus on the sex-chromosomes. I characterized the status of Y chromosome degeneration in the threespine stickleback and its effects on gene function. Furthermore, since the degeneration process leaves genes in a single copy in males, I examined whether the resulting dosage difference of messenger RNA for hemizygous genes is compensated as it is in other organisms. In addition, threespine sticklebacks have well-characterized behavioral differences related to the male's social status during the breeding season. To understand the connection between the genotype and behavior, I examined gene expression patterns related to breeding behavior using dominant and subordinate males as well as females.

# TABLE OF CONTENTS

**ABSTRACT** ........ **3**

**LIST OF ORIGINAL PUBLICATIONS** ........ **6**

**1. INTRODUCTION** ........ **7**

**2. AIMS OF THE THESIS** ........ **12**

2.1. Threespine Stickleback as model for molecular mechanisms of sexual dimorphism ........ 12

**3. MATERIAL AND METHODS** ........ **16**

3.1. Threespine stickleback sample collection ........ 16

3.2. Molecular methods ........ 17
3.2.1. Extraction methods ........ 18
3.2.2. Next generation Sequencing based methods ........ 18
3.2.2.1. Genome sequencing based methods (study II) ........ 18
3.2.2.2. Transcriptome sequencing based methods (study II, III) ........ 20
3.2.3. Microarray based methods (study IV) ........ 21
3.2.4. Mass spectrometry based methods ........ 22
3.2.4.1. Protein quantitation (study I) ........ 23
3.2.4.2. Y-specific protein identification (study II) ........ 24

3.3. Bioinformatics methods ........ 24
3.3.1. Biomart and BLAST in gene functional annotation ........ 24
3.3.2. Gene functional enrichment and pathway analysis ........ 25
3.3.3. Gene variant effect prediction (II) ........ 26

3.4. Evolutionary analysis of the sex chromosome (study II) ........ 27

**4. MAIN RESULTS AND DISCUSSION** ........ **28**

4.1. Sexual dimorphism at the protein level in the threespine stickleback (I) ........ 28

4.2. Effects of Y chromosome evolution on gene loss and persistence patterns (II) ........ 29

4.3. Dosage compensation in threespine stickleback (III) ........ 34

4.4. Hypothalamic transcriptome expression during reproduction in males and females (IV) ........ 36

**5. IMPLICATIONS AND FUTURE DIRECTIONS** ........ **40**

**6. ACKNOWLEDGEMENTS ........ 42**

**7. REFERENCES ........ 44**

**ORIGINAL PUBLICATIONS ........ 51**

# LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following studies and publications which are referred to in the text by their Roman numeral:

I Viitaniemi, HM and EH Leder. (2011). Sex-Biased Protein Expression in Threespine Stickleback, *Gasterosteus aculeatus*. *Journal of Proteomic Research* 10: 4033-4040

II Viitaniemi, HM, R Schultheiß, T Sävilammi, and EH Leder. Withering Ys - Patterns of gene persistence and loss on a novel sex chromosome. *Manuscript*

III Schultheiß, R, HM Viitaniemi, and EH Leder. Spatial dynamics of evolving dosage compensation in a young sex chromosome system. *Genome Biology and Evolution*, published online January 23, 2015

IV Viitaniemi, HM and Erica Leder. Hypothalamic mRNA expression in the threespine stickleback: effects of sex and reproductive status. *Manuscript*

Throughout the thesis, individual studies will be referred to with the Roman numeral assigned for each one. The following articles are reprinted with permission from:

Sex-Biased Protein Expression in Threespine Stickleback, *Gasterosteus aculeatus*. *J. Proteome Res.*, 2011, 10 (9), pp 4033–4040, **DOI:** 10.1021/pr200234a. Copyright 2011 American Chemical Society.

# 1. INTRODUCTION

Sexual dimorphism is a term used to describe features that discriminate between the two sexes, males and females. These features can be diverse ranging from external to internal and they are often linked with reproduction (Barnett and Bellwood, 2005; Frayer and Wolpoff, 1985; Hedrick and Temeles, 1989; Isensee and Noppinger, 2007; McPherson and Chenoweth, 2012); this is why sexual dimorphism is often most obvious during the time of breeding. Elaborate morphological features, such as the tail feathers of a peacock (Petrie et al., 1991) or the mating colors of African Cichlid species (Maan and Seehausen, 2010; Miyagi and Terai, 2013), are most pronounced and well-known. These morphological traits aim at conveying superiority as a mate (Williams, 1978) and ensuring mating with conspecifics (Genner and Turner, 2005; Moller and Cuervo, 1998; Parker and Partridge, 1998). Sexual dimorphism in behavioral traits, such as song characteristics of male birds (McGregor et al., 1981; Buchanan and Catchpole, 2000; Woodgate et al., 2012), are usually most striking at the breeding season and function in attracting a mate and thus ensuring successful reproduction. During the non-reproductive period the sex of an individual is maintained in the functions of many organs which express sex-specific genes (Isensee and Noppinger, 2007; Nishida et al., 2005a; Yang, 2006). Thus sexual dimorphism extends also beneath the external appearance with differences in physiology, such as metabolizing alcohol and steroid hormones in liver ((Mode and Gustafsson, 2006; Waxman and Holloway, 2009). Sexual dimorphism has received a lot of interest within the scientific community (Barnett and Bellwood, 2005; Isensee and Noppinger, 2007; McPherson and Chenoweth, 2012; Packer, 1983; Petrie et al., 1991). Additionally, many have examined the relationship among sexual dimorphism, sexual selection and reproduction (Fairbairn and Roff, 2006; Lande, 1980; Lindholm and Breden, 2002; Parsch and Ellegren, 2013; Rice, 1984; Valenzuela, 2008). However, the molecular mechanisms responsible for many sexually dimorphic traits have remained mostly undetermined. Examination of the underlying molecular mechanisms of sexual dimorphism has become possible with the increased availability of molecular methods, particularly for non-model organisms (e.g. Johnston et al., 2011; Pointer et al., 2013; Takahashi et al., 2013).

Sex-specific differences have been suggested to arise due to sexual selection (Darwin, 1859), such as horn size in African antelope species (Packer, 1983) or coloration in cichlid fish (Miyagi and Terai, 2013). A trait beneficial for one sex (e.g. bright coloration in males to attract females) can, when expressed in the opposite sex (e.g. colorful female conspicuous to predators), lower the fitness of the other sex (Rice, 1992). These types of traits which affect the fitness of males and females in opposite directions are called sexually antagonistic as the sexes have different selective optima for the trait (Bonduriansky and Chenoweth, 2009). Specifically, traits related to reproduction are under sexual selection as a tug of war

exists between the sexes over the investments made to assure the success of the following generation; this causes a sexual conflict (Chapman et al., 2003; Rice and Chippindale, 2001). A key question researchers have been trying to answer for decades is how to resolve these differing selective optima between the sexes (Griffin et al., 2013; Roberts et al., 2009; Stewart et al., 2010; Wright and Mank, 2013). Commonly the conflict due to an antagonistic trait is resolved by making the trait sex limited (i.e. sexually dimorphic). This can be achieved by modifying when and where the trait is expressed, and ultimately, this is controlled by the DNA.

The most important molecular mechanisms for determining a phenotype, external or internal, are transcription and translation of the DNA template and the interplay between these two processes (Carroll, 2009). Gene regulation involves the modulation of what genes are transcribed in which tissue and at what time during the day or life or an organism (Lin et al., 2010; Wittkopp et al., 2004). The modification steps between transcription and translation then further determines which of the transcripts eventually end up as functional proteins (Kozak, 1992; Valencia-Sanchez, 2006).

Gene regulation involves transcription factors that promote transcription of a gene by binding to specific regulatory DNA sequences upstream or within introns of a gene. Composition of these regulatory areas adjacent to the gene (*cis*-acting regulation) and transcription factors acting on these sites (*trans*-acting regulation) influence which genes and the quantity of those genes that are ultimately transcribed. Recently, mutations in regulatory regions have been proposed to be the main force in morphological evolution (Carroll, 2008). In an analysis of genome-wide variation in mRNA expression in humans, both *cis*- and *trans*-acting factors were found to be responsible for mRNA expression variation (Morley et al., 2004). Further evidence that gene regulation has a major role in creating sex-specific differences comes from Drosophila melanogaster where sexually dimorphic abdominal pigmentation was found to be due to differences in regulation between males and females (Williams et al., 2008).

As the study by (Williams et al., 2008) shows, one way to generate/maintain sexually dimorphic traits is by sex-specific gene regulation and expression. Given that the genome of males and females are almost identical with the exception of the few genes on the Y- (or W-) chromosome or the sex-determining alleles (in the case of organisms without sex chromosomes), it is likely that many of the downstream processes resulting in sexual dimorphism are produced by changes in gene regulation (Griffin et al., 2013; Meiklejohn et al., 2014; Pennell and Morrow, 2013). By regulating genes differentially, expression of a trait can be restricted to the sex that benefits and thus detrimental effects on the other sex can be avoided/diminished (Rice, 1984; Wang and Purisima, 2005; Ellegren and Parsch, 2007; Parsch and Ellegren, 2013). Nevertheless, sexual antagonism does not always need to be the prerequisite for sex-biased expression (Parsch and Ellegren, 2013) nor are sex

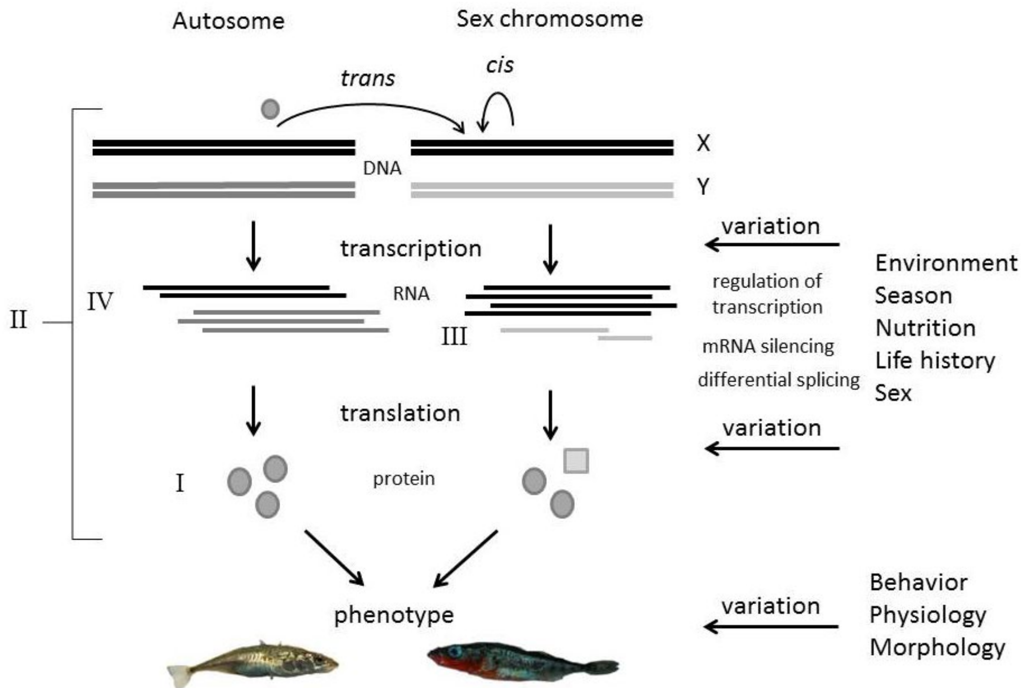



**Figure 1.** There are three molecular levels, DNA, RNA and proteins, which constitute the basis of gene expression. These levels are connected to each other via mechanisms of transcription and translation. In both of these intermediate steps there can be variation due to regulation and processing of the mRNA or peptide. Different factors such as sex or environment (on the left hand side) affect what is expressed and where. Ultimately, gene expression leads to the phenotype and the sexually dimorphic differences we observe regarding behavior, physiology and morphology. To understand how the sexually dimorphic features are formed, I studied different parts of this chain. The four studies are marked with the corresponding Roman numerals in this figure.

chromosomes always needed to resolve sexual conflict as it can also be resolved by sex-specific expression patterns in autosomal genes (Fairbairn and Roff, 2006; Mank et al., 2006). Sex-specific expression is achieved by differential gene regulation, as discussed above, and the sex-specific regulation is expected to originate from the sex chromosomes (Bellott et al., 2014; Charlesworth et al., 1987; Rice, 1984). However, in organisms with genetic sex determination, the apparent contradiction between sexually antagonistic traits and shared genome can be overcome by tight linkage between the antagonistic trait and sex (sex-linkage) (Parnell and Streelman, 2013); sex-linkage thus further creates the potential for the evolution of chromosomal sex determination (Rice, 1987).

According to theory, development of chromosomal sex determination begins on the chromosome where the sex-determining locus resides. The next step involves the repression of recombination which results in the accumulation of mutations and repetitive DNA (Charlesworth, 1991). Accumulation of sex-specific alleles or sexually antagonistic alleles further drives the separation of the sexes (Rice, 1987). However, despite the difference in

gene composition between the X and the Y, a degenerated Y carries only a small number of unique genes and therefore the role of gene regulation is likely still an important factor even in organisms with developed sex chromosomes (Bellott et al., 2014; Charlesworth et al., 1987; Cortez et al., 2014; Williams et al., 2008).

Three important steps are commonly recognized when heteromorphic sex chromosomes evolve from autosomes: suppression of XY recombination, Y degeneration, and dosage compensation (Charlesworth, 1996). The first step towards evolving sex chromosomes is local suppression of recombination around the sex determining locus in an autosome to produce the proto-X and proto-Y (Charlesworth, 1991; Rice, 1996). Recombination is reduced to avoid fitness reduction due to unfit allele combinations which could produce individuals with an intermediate sex phenotype. Loci in vicinity of this locus are also affected by the repression of recombination (Rice, 1992). Suppression of recombination gradually spreads until nearly the whole chromosome has stopped recombining. Once recombination is suppressed, selection acting against deleterious mutations is also reduced in the Y (Charlesworth and Charlesworth, 2000).

Once recombination suppression between the Y and X has started, the Y is free to accumulate mutations and repetitive DNA as well as structural changes (inversions, deletions and insertions) as a result of the the suppressed recombination; over time these can result in expansion of the non-recombining region (Charlesworth, 1991; Charlesworth and Charlesworth, 2000; Charlesworth et al., 2005; Rice, 1987, 1996). This stepwise suppression of recombination can also been seen in the age of the Y chromosome genes; blocks where recombination has stopped at a similar time comprise an evolutionary stratum in the Y (Lahn, 1999). Also, accumulation of sex-specific alleles or sexually antagonistic alleles can further drive the separation of the X and Y chromosomes (Rice, 1987). Over time, loss of recombination and the subsequent changes in the Y chromosome lead to loss of genes either due to non-functionalization or loss of sequence from the Y. The process of loss of Y chromosome sequence and genes is called Y degeneration (Charlesworth and Charlesworth, 2000). The genes which are retained in the Y chromosome tend to accumulate male-specific alleles and functions (Wilson Sayres and Makova, 2012; Zhou and Bachtrog, 2012) whereas in the X chromosome the fate of male-beneficial alleles depends on their dominance; recessive male-beneficial alleles are more likely to be X-linked than alleles with even partial dominance (Rice, 1984). Although it is possible for a sexually antagonistic allele to exist in an autosome, sex-linkage is expected since the spread of an antagonistic allele is better facilitated in the X chromosome even when the allele is rare in frequency (Rice, 1984).

Once a gene, or its function, becomes lost from the Y chromosome, a male faces a problem with maintaining the required amount of the transcribed or translated gene product to sustain the expression balance in the pathway to which the gene belongs (Charlesworth et al., 2005). As these pathways contain several genes, expression level balance is often highly

evolved in autosomes and based on having two functioning chromosomes; when loss occurs in the male Y, the resulting imbalance should be compensated to ensure proper functioning of the pathway to be expressed (Charlesworth, 1996; Disteche, 2012; Mank, 2013). The mechanism which corrects the imbalance is called dosage compensation. To date, several different ways of how dosage compensation can be achieved have been reported and they range from up-regulation of the single X chromosome in the male to complete silencing of one of the female X chromosomes (Disteche, 2012).

As sex chromosomes have evolved from autosomes, understanding the trajectory from an autosome to a sex chromosome requires using taxa which enable investigations of the incipient degeneration process. However, many well-studied taxa (e.g. mammals, birds, and *Drosophila* spp) have old sex chromosomes where the degeneration process has already resulted in highly diverged sex chromosomes (Berlin and Ellegren, 2005; Bernardo Carvalho et al., 2009; Kaiser and Bachtrog, 2010; Kejnovsky and Vyskot, 2010; Kondo, 2006; Lahn, 1999; Wang et al., 2012; Zhou et al., 2008)2010; Kejnovsky and Vyskot, 2010; Kondo, 2006; Lahn, 1999; Wang et al., 2012; Zhou et al., 2008. Neo-sex chromosomes, where autosomal regions have been fused to an already existing sex chromosome, can offer valuable insights into the degeneration process (Yoshida et al., 2014; Zhou and Bachtrog, 2012; Zhou et al., 2008). Nonetheless, the degeneration of novel sex chromosomes that originate directly from autosomes likely differs from that of neo-sex chromosomes as many mechanisms associated with sex chromosomes (e.g. suppressed recombination and/or dosage compensation) already exist at the time of the chromosome fusion (Bachtrog, 2011). As already discussed, a tight link between sex chromosomes and sexual dimorphism is expected and thus combining research of sex chromosomes and sexual dimorphism can facilitate our understanding of the molecular background of sexual dimorphism. Furthermore, as the evolution of sexual dimorphism can be influenced by dosage imbalances and degeneration of the Y as well as sex-specific expression in the autosomes, it will important to understand how these two mechanisms shape the genetics of sexual dimorphism (Dean and Mank, 2014; Fairbairn and Roff, 2006).

# 2. AIMS OF THE THESIS

Since it is the molecular expression that eventually leads to a sex-specific phenotype in morphology, physiology, and behavior, the aim of my thesis is to investigate how males and females differ at the molecular level. As sex chromosomes are proposed to be important in development and persistence of sexual dimorphism (Connallon and Clark, 2010; Rice, 1984; Roberts et al., 2009; Vallender and Lahn, 2004; Zhou and Bachtrog, 2012), I will add further depth to the investigation by examining a species with nascent sex chromosome evolution. This allows for investigating molecular dynamics and evolution of both sexual dimorphism and sex chromosomes at the same time. Leder et al. (2010) showed that the liver transcriptome is sexually biased in adult fish outside the breeding season and my first aim was to investigate whether this bias extends also to the proteome (study I). After examining the transcriptome and proteome and identifying sex-bias in both levels, the next aim was to understand the causes of sex-biased expression in the DNA level. Towards this aim I characterized and defined the status of Y chromosome degeneration in the threespine stickleback and its effects on gene persistence and loss (study II). I further investigated how Y degeneration links to the dose of messenger RNA (mRNA) between the sexes since some genes in males appear in a single copy (study III). As a major component of sexual dimorphism in threespine stickleback is observed in morphology and behavior during breeding, I investigated mRNA expression patterns in the hypothalamus to see what differences are observed in the brain related to male and female reproductive status (study IV).

## 2.1. Threespine Stickleback as model for molecular mechanisms of sexual dimorphism

To accomplish the aim of my thesis, I chose a species which has the features mentioned above, namely sexually dimorphic traits and in the early stages of sex chromosome evolution. The threespine stickleback, *Gasterosteus aculeatus*, has established status as a model organism in evolutionary studies, and its full potential as a model species has been realized with the sequenced genome allowing the use of modern molecular techniques to address a variety of evolutionary and ecological questions (Jones et al., 2012). Besides the sequenced genome, there are other features in this species, such as evolutionary history (Aguirre et al., 2008; Leinonen et al., 2006; Mäkinen et al., 2006; Mäkinen et al., 2008) and traits related to reproduction (FitzGerald and Wootton, 1993; Kitano et al., 2007; Tinbergen, 1951) which make it suitable for our study.

The threespine stickleback is a small teleost distributed around the northern hemisphere. It can be encountered in marine habitats as well as in freshwater habitats. The latter have

been colonized by marine individuals at the end of the last glacial period as freshwater habitats became available (Aguirre et al., 2008; Leinonen et al., 2006; Mäkinen et al., 2006; Mäkinen et al., 2008). Since the establishment of freshwater populations, significant phenotypic divergence between marine and freshwater forms has developed (Aguirre et al., 2008; Leinonen et al., 2006), and a similar degree of morphological difference has also been shown to exist between the sexes (Aguirre and Akinpelu, 2010; Albert et al., 2007; Leinonen et al., 2010). It has been proposed that the phenotypic differences between marine and freshwater forms would reflect the phenotypic differences between a male and a female (Albert et al., 2007). In other words, a freshwater form resembles a male phenotype and a marine form resembles the female phenotype.

Sexually dimorphic traits related to reproduction are well characterized in threespine stickleback (FitzGerald and Wootton, 1993; Foster et al., 2008; Kitano et al., 2007; Tinbergen, 1951). One extensively studied trait is the sexually selected carotenoid-based thoracic red-pigmentation (Pike et al., 2007; Sparkes et al., 2008; Tinbergen, 1951) which is displayed for females by the males during a courtship dance. Parental care displayed by the male is also characteristic for the threespine stickleback. It includes the male being responsible for nest-building, for guarding the eggs and ventilating the eggs by fanning their fins (McKinnon, 1996; Pressley, 1981; Slijkhuis et al., 1984). Interestingly, these behaviors differ to some degree between the ecological morphs. For example the marine morph and benthic form share less showy courtship behaviors, e.g. zigzag swimming, due to cannibalistic behaviors within the populations, whereas the benthic morphs lacking cannibalism are more elaborate in their courtship (Foster et al. 2008). Hormonal regulation is important for controlling the cyclicity and onset of the breeding season in males (Páll et al., 2005). The major hormonal player in males is 11-ketotestosterone (11KT) which increases at the start of the breeding season and also controls the development of secondary sexual characteristics such as spiggin production (Jakobsson et al., 1999; Jones, 2001), kidney hypertrophy (Jakobsson et al., 1999) and male nest building and courting behavior (Hellqvist et al., 2008; Páll et al., 2002a, 2002b).

In addition to morphological and behavioral differences observed within threespine stickleback, linkage analyses have revealed that several quantitative trait loci (QTL) associated with sexual dimorphism in body shape map to chromosome XIX (Albert et al., 2007). This chromosome is the nascent sex chromosome and males are heterogametic (i.e. XY system) (Peichel et al., 2004). Comparative studies have shown that XIX is the sex chromosome shared by the two extant *Gasterosteus* species (*G. aculeatus* and *G. wheatlandi*) but not by *Pungitius* (Ross et al., 2009). The phylogeny of the stickleback fish species is relatively well known (Bell and Foster, 1994) and thus the age of the sex chromosomes can be estimated. *Gasterosteus* diverged from *Pungitius* approximately 16-13 million years ago (Bell and Foster, 1994; Bell et al., 2009). Within the *Gasterosteus* two extant species are encountered, *G. aculeatus* and *G. wheatlandi* and their estimated split happened approximately 10 million

years ago (Bell and Foster, 1994). Therefore, the sex chromosome has risen after the split of *Pungitius* and *Gasterosteus* but before the split between *G. aculeatus* and *G. wheatlandi* (Bell and Foster, 1994; Bell et al., 2009) placing the age the age of the *Gasterosteus* sex chromosome around 13-16 MY. Subsequently, one further split within *G. aculeatus* has resulted in two lineages approximately two million years ago; a lineage found in Japan Sea and the lineage encountered elsewhere in the Northern hemisphere (Higuchi and Goto, 1996). This Japan Sea lineage has recently been observed to contain a neo-sex chromosome formed by the fusion of chromosome IX to chromosome XIX (Ross et al., 2009).

In the threespine stickleback, the metaphase size of X and Y are only slightly different; one major deletion and multiple pericentric inversions have been revealed by fluorescent in situ hybridization experiments (Ross and Peichel, 2008). Additionally, sequencing of bacterial artificial chromosome (BAC) clones identified multiple insertions and deletions; thus it is likely that the accumulation of repetitive DNA in the Y contributes to the similarity in size during metaphase (Peichel et al., 2004; Ross and Peichel, 2008). Altogether the threespine sex chromosome has been shown to consist of three regions: the recombining pseudoautosomal region, the non-recombining region (stratum I) and the non-homologous region (stratum II) (Roesti et al., 2013; Ross and Peichel, 2008). These features indicate that the threespine stickleback Y chromosome is undergoing degeneration.

Compared to other recently investigated sex chromosome systems, the threespine stickleback system has several advantages. Firstly, the threespine stickleback has well-established ecology and sexually dimorphic traits that have been described in detail (see above). Secondly, as the estimated age of the threespine stickleback Y chromosome is around 13-16 MY, it is young compared to the over hundred million year old, vastly degenerated mammalian sex chromosomes (Livernois et al., 2011; Wilson Sayres and Makova, 2012). Although young neo-sex chromosomes, where an autosome has fused to an existing sex chromosome, are useful for studying Y chromosome evolution, the steps in neo-sex chromosomes can be different due to the fusion occurring in an established sex chromosome with already existing features of a sex chromosome such as dosage compensation and suppressed recombination (e.g. the *D. miranda* neo-Y, Zhou and Bachtrog, 2012). Thirdly, the availability of an assembled genome sequence allows visualizing patterns of gene loss and silencing in a spatial framework when genome sequencing is combined with transcriptomics and proteomics; this spatial scale cannot be achieved as efficiently with just transcriptome sequence (e.g. the *Silene latifolia*, Bergero and Charlesworth, 2011). Lastly, differences in life history characteristics, such as mating system, generation time or alteration of generations, are expected to influence the Y degeneration process (Bachtrog, 2011).

In summary, investigation of species with sex chromosomes of recent autosomal origin and with varied life history characteristics will benefit our understanding of the early stages

of Y chromosome evolution. Also as sexual dimorphism is theoretically related to sex chromosomes, combining these two lines of investigation can yield a deeper understanding of the link between genotype and phenotype. Thus, the threespine stickleback provides a valuable system to investigate the early stages of Y chromosome evolution and for relating Y chromosome evolution to the molecular background of sexually dimorphic traits within the species.

# 3. MATERIAL AND METHODS

## 3.1. Threespine stickleback sample collection

As the focus of the research was on sexual dimorphism, and it is most pronounced during the reproductive period, most of the sample collection was done during the breeding season when the individuals were characterized with visible reproductive traits. These included male nuptial coloration (i.e. blue eye and red throat) and gravid females (studies II, III and IV). For study I, the samples were not collected at the reproductive season but this was adequate as the investigation was about sexually dimorphic protein expression between sexes and mRNA expression was previously shown to differ between the sexes in adult non-reproductively active individuals (Leder et al., 2010). A variety of tissues, which are known to have impact on sexual dimorphism were collected. For example liver has a role in metabolism of steroids, and males and females metabolize some compounds with different efficiency (Ahluwalia, 2003; Kwekel et al., 2010; Roy and Chatterjee, 1983; Valle et al., 2007). Brain, on the other hand, is known to control onset of sexual maturation and behavior related to reproduction (Almeida et al., 2012; Schlinger et al., 1999; Schulz and Sisk, 2006).

The fish used in study I were unrelated individuals from a laboratory experiment conducted at the University of Helsinki. Parents for the individuals originated from the Baltic Sea collected in Helsinki, Finland. Details of the individuals, rearing conditions and sampling are described in (Leinonen et al., 2010). Liver tissue was used for the experiment.

Unlike study in I, the fish used in studies II, III and IV were of wild origin. Fish from lakes in Alaska, USA, Bear Paw (61.6141109448°, -149.7534433°), Lynne (61.712°, -150.039°), South Rolly (61.401°, -150.073°) and Corcoran (61.574°, -149.688°) were used for studies II and III: two males and two females from Bear Paw for study II, and five males and five females from all four populations for study III. All fish were caught using minnow traps during breeding season in June 2010. Additionally for study II fish from the Baltic Sea at the island of Seili (FIN, 60.241321°, 21.962849°) were collected. These two sampling locations, Alaska (freshwater) and Finland (brackish water), were chosen to add geographic dimension to the investigation of sex chromosome evolution as these two populations have diverged since the last ice age (Bell and Foster, 1994). Also 25 additional individuals from Bear Paw and 10 from Baltic Sea were collected and used in study II to verify with more individuals the allele calls for the first mentioned eight individuals from Bear Paw and Seili.

For study IV the samples were collected from a marine population near the Sven Löven research station, (Fiskebäckskil, Sweden). Upon capture, the individuals were separated to

males and females, and maintained in aquaria at the research station with circulating marine water (~14 degrees) until the experiment was started. For allowing the establishment of social status one female and three males (displaying nuptial coloration at collection time) were put into the experimental aquaria and were observed until one of the males achieved breeding coloration, was attacking the other males, and was courting the female while the other males remained subordinate. The other two males were determined as subordinate based on their behavior of trying to avoid the dominant male and not displaying breeding coloration. One of the two subordinate males in the tank was randomly chosen for the microarray experiment. All of the 16 males (eight dominant and eight subordinate) and five females in the experiments were decapitated and sampled at the research station immediately after the experiment.

**Table 1.** Samples used in the four studies presented in the thesis.

| Study | Sampling location | Sampling time | Origin | Tissue |
|---|---|---|---|---|
| I | Helsinki University | 2009 | Laboratory | Liver |
| II | Alaska and Baltic Sea | 2010 June, Alaska<br>2011 June, Seili | Field | Brain, Liver, Testis, Muscle |
| III | Alaska | 2010 June | Field | Brain |
| IV | Sweden | 2013 June | Field | Hypothalamus |

## 3.2. Molecular methods

In the era of the ‘omics science, development of techniques for gathering and analyzing genomic, transcriptomic, proteomic or other types of data has made it feasible to address many ecological and evolutionary questions in a high-throughput manner even in non-model species (Joyce and Palsson, 2006; McLean, 2013). Although availability of genome is no longer a limitation when selecting a model for testing ecological and evolutionary questions, there are benefits for having a well-assembled genome; e.g. it is possible to adopt methods from model organisms such as targeted genome capture. Throughout my thesis I am using high-throughput techniques either alone (studies I, III and IV), or in a combination such that the separate techniques complement each other (study II). For example the route from mRNA to protein is not always straightforward (Anderson and Seilhamer, 1997; Cox et al., 2005; Ning et al., 2012) as transcribed mRNA can be subjected to further regulation by microRNA (Valencia-Sanchez, 2006). In addition, proteins go through post-translational modifications such as phosphorylation or glycosylation before being functional or are destined for degradation by ubiquitination. Thus interdisciplinary interpretation of data is encouraged (Ning et al., 2012). As there are numerous methodologies available for genomics, transcriptomics and proteomics, I will discuss in more detail the methodological choices for each study in the following sections.

### 3.2.1. Extraction methods

As all the molecular levels, DNA, RNA and proteins, were investigated in the studies, I primarily used extraction methods which allow simultaneous collection of more than one type of molecule in high quality. DNA used in study II was extracted following a traditional salt extraction method with RNAse treatment included to ensure sufficient quality and quantity for next generation sequencing. Choice of tissue occasionally limits the extraction method. For example, when the tissue is relatively homogeneous and plentiful (e.g. liver) it is possible to do several extractions from the same tissue and the amount of tissue is not limited. On the other hand, when the tissue is functionally organized or small (e.g. brain) only one extraction is possible and thus consideration for the extraction method given the desired downstream processes is required. Therefore I chose a phenol-based phase separation for extracting RNA and proteins used in studies I, II, III and IV. This protocol yields RNA and protein suitable for transcriptome approaches or mass spectrometry based proteomics, respectively, and can also provide DNA, yet quantity tends to be limited. All RNA samples were extracted twice with DNAse treatment with RQ1 RNAse-free DNAse (Promega) after the first isolation to ensure no traces of DNA or DNAse are left in the RNA.

### 3.2.2. Next generation Sequencing based methods

Recent advances and development of fast, accurate and affordable sequencing methodologies have opened a new view to evolutionary research. Collectively these methods that enable rapid high throughput sequencing are called next generation sequencing (NGS) (Metzker, 2009). For transcriptome research, NGS overcomes some of the problems associated with microarrays and simultaneously enables acquiring data for millions of molecules without prior knowledge of the sequence (Wang et al., 2009; Ozsolak and Milos, 2010). The NGS platform used in studies II and III is Illumina HiSeq. For both DNA and RNA 100 bases of paired-end sequencing was chosen. Paired-end sequencing has benefits for identifying structural variants or fusion events as a result of read-through between neighboring transcripts in the genome (Maher et al., 2009).

In the case of the threespine stickleback an assembled genome is available which can be used for mapping the sequences obtained with NGS. The reference genome used for mapping the DNA and mRNA reads from Illumina was Ensembl build v67 of the threespine stickleback. In 2008 Ross and Peichel reported that the last two supercontigs of chromosome XIX in the Ensembl assembly were inverted. I corrected for this reported inversion of the last two super contigs in chromosome XIX in my version of the reference genome.

#### 3.2.2.1. Genome sequencing based methods (study II)

Albeit the genome of the threespine is publicly available, it only has the X chromosomal sequence representing the sex chromosomes (Jones et al., 2012). In this case, whole

genome sequencing would not yield the same benefits as a reductionist approach targeting only the sex chromosome. Genome reduction methods provide the advantage of acquiring more data over a subset, e.g. chromosome, of the genome of interest (Mamanova et al., 2010). With the availability of the genome, it is possible to adopt techniques used in model organism and to remold it to suit the need of the experimental question. Since genome information for males (e.g. Y chromosome) is unavailable, DNA sequence capture was used to acquire sequence of the sex chromosomes for males and females (see 3.1. for sample details). Oligo probes (or baits) were designed from the genome sequence to cover the majority of chromosome XIX and some regions from the autosomes as a control. These probes are then used to capture corresponding sequences from the total DNA which are then consequently sequenced. However, Y-specific sequences which are too divergent to hybridize to the probes based on X sequence will not be observed in these data.

To confirm that the allele calls from capture study are representative of the population, genotyping by sequencing (GBS) (Elshire et al., 2011) was done to obtain sequence data for additional individuals. GBS is also a genome reduction method based on fragmentation of the genome with a restriction enzyme followed by ligation of adapters for sequencing and library amplification. This method also allows for the insertion of barcodes with the adapters for pooling tens of samples on one lane. In addition, the desired insert size is can be obtained by size selection of the library prior to sequencing. As GBS is based on random small fragments from the whole genome whereas the DNA capture is targeting specific chromosomal regions, the overlap between these two datasets is expected to be small. However, it provides an estimate of the accuracy for genotype calling of the capture data since it utilizes more individuals.

Reads from both the DNA capture and the GBS were trimmed using window adaptive trimming to remove low quality bases prior to read mapping. Due to read mapping algorithms performing alignment one read at a time against the reference genome while the previously aligned reads at the region are not considered at all, erroneous read alignment can occur and lead to erroneously called polymorphisms in the sample. Not considering this problem can lead to biased estimates of variation in the sample. There are tools available, such as the local re-alignment tool provided in GATK 2.3.9 (Broad Institute) (McKenna et al., 2010), specifically designed to identify and correct regions where erroneous read alignment has happened. In essence, the GATK re-alignment tool performs a local multiple alignment for regions with indels. Local re-alignment was done to improve variant calling in males and females and also the accuracy of assigning Y chromosome alleles. These Y-specific calls were also confirmed, when possible, with the allele calls made from GBS data (see paragraph above) and transcriptomes (see section below).

#### 3.2.2.2. Transcriptome sequencing based methods (study II, III)

In studies II and III, I used RNA sequencing (RNA-Seq) to investigate transcript expression across tissues. In RNA-Seq no prior genome information is required for sequencing as is needed for microarrays; the only limitation for detecting a transcript comes from whether the gene is expressed in the tissue at the time of sampling (Ozsolak and Milos, 2010; Wang et al., 2009). The major benefit of using transcriptome sequencing over hybridization-based microarray methods is obtaining the actual sequence of the transcript in question. With sequence information it is possible to identify variants between individuals and this information can further be used for assessing the level of alternative mRNA splicing and parental origin of expression (Gregg et al., 2010; Li et al., 2012) or, as in my case, to separate expression coming from homologous chromosomes in males, namely X and Y. Other benefits of RNA-Seq are low background, high dynamic range and reproducibility (Malone and Oliver, 2011).

In study II the acquired transcriptome reads were used to examine the propensity of the degenerating Y-copy genes to be transcribed. Mapping of the reads was done with the Tuxedo pipeline (Trapnell et al., 2012). This pipeline is designed for the specific needs of mapping mRNA reads to reference genome since a read can contain sequence from two different exons, and it further characterizes transcripts from the mapped read data. Variant loci identified from the transcribed genes on the sex chromosome were compared to the Y allele calls made from DNA capture. This was done to determine presence of Y-specific transcripts in males in order to ascertain the level of gene loss at the transcriptome level. I also examined whether expression levels for X and Y chromosomes in males were differing by determining allele-specific expression (ASE) patterns (Gregg et al., 2010; Li et al., 2012) for those variant loci in transcripts which overlapped Y-diagnostic loci from the capture data (i.e. Y alleles) and had a minimum depth of 20 reads to ensure accurate estimation of the allelic proportions.

In study III the focus was solely on expression patterns between the sexes and between autosomes and the sex chromosome in the genome in order to examine the status of dosage compensation. Mapping of the reads was done with the Tuxedo pipeline (Trapnell et al., 2012). Although the tuxedo pipeline can also be used to calculate differential expression, the R package DeSeq2 (Anders and Huber, 2012) offers more choice in statistical design. Differential expression was first determined between the sexes after which the expression levels between autosomal and sex-linked genes were calculated for both sexes. The latter analysis was also done separately for the two different strata on the sex chromosome (see 2.1.). As the Y chromosome potentially contributes to the overall sex chromosome expression level in males, all genes determined to have a transcribed Y-copy in males were removed therefore leaving only the hemizygous genes in the male (i.e. X:AA).

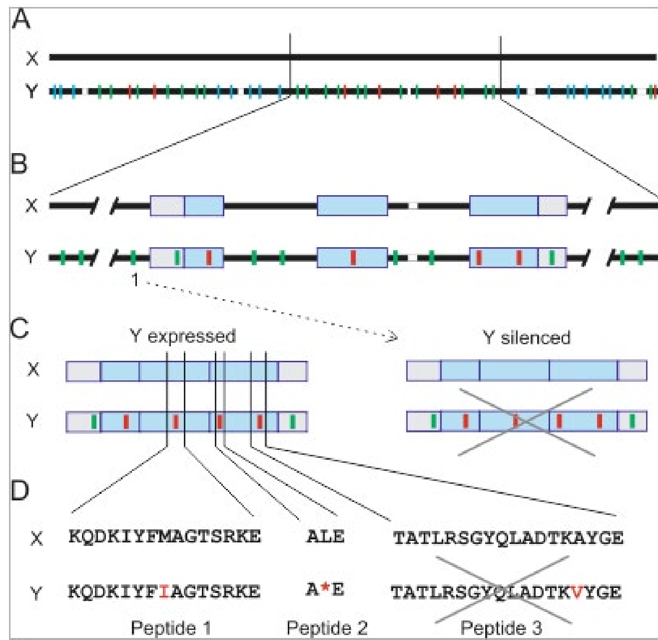



**Figure 2.** Workflow of study II illustrating how data from different methodologies was used to examine Y chromosome. The majority of mutations specific to the Y were found in the non-recombining region, and a hypothetical section of this region is shown (A). A high frequency of SNPs (colored lines) and indels (white boxes) were identified throughout this region. From the captured DNA data, Y-specific mutations were then categorized by their potential effect on the gene (B). They can be broadly classified into mutations in gene regions, both coding (red lines) and potentially regulatory (green lines) or intergenic (blue lines in A). Transcriptome sequencing was then carried out in order to confirm and further assess the impact of the mutations identified at the DNA level (C). In some cases, genes with Y-copy alleles at the DNA level were not observed at the transcriptome level. These silenced Y-copy genes presumably have been rendered non-functional due to mutation(s) in the regulatory regions (e.g. mutation 1 in B). Finally, LC-MS/MS of male and female proteomes was conducted to confirm that Y-copy alleles were being translated into proteins. If a male produced both X and Y copies of a protein, often it was observed in the peptides from proteomic data (peptide 1 in D). In cases where the mutation created a premature stop codon, as * in peptide 2, the Y-copy of peptide 3 would not be expected in the data. Incidentally, peptide 2 would not be observed from either the X or the Y due to its small size.

### 3.2.3. Microarray based methods (study IV)

In study IV, I investigated the transcriptome of hypothalamus tissue in male and female threespine stickleback with a DNA oligo array. Although the molecule under examination is the same as in RNA-Seq, the way in which it is analyzed is inherently different (Wang et al., 2009; Malone and Oliver, 2011). Probes used in oligo arrays are predesigned 60 base (in this case) oligos that are attached to a glass surface; as mentioned previously, some prior genome knowledge is needed in order to design the probes. To quantify expression, samples are fluorescently labelled with one or two dyes and then allowed to hybridize to the complementary target probes immobilized on the array. Expression level is determined

by measuring fluorescence of the probes by laser (Malone and Oliver, 2011). Although there are inherent problems and biases to consider with microarrays (see above for RNA-seq), a vast body of research has provided solutions to them (Malone and Oliver, 2011; Smyth, 2005; Yang et al., 2002). For example, in two-color arrays, the difference in the label florescence intensity needs to be taken into account when analyzing the results across array slides (Rosenzweig et al., 2004; Yang et al., 2002). A clear benefit of microarrays over sequencing methods is the cost per sample. One way to overcome cost issues with RNA-Seq is to pool several individuals together for the sequencing; however, then the information of individual level variation of expression is lost. Therefore, when individual level variation is required a microarray provides a more cost-effective way to examine expression for 24 samples compared to RNA Seq.

In brief, each male sample was labelled once, either by Cy3 or Cy5, and hybridized twice on an array; once against a female and once against a male having the opposite social status (i.e. dominant or subordinate). Female samples were labelled with both Cy3 and Cy5 because of this experimental design (i.e. the same female had to be hybridized to both the dominant male and the subordinate male which were in opposite colors). Differential expression between the dominant male-female (DMvF), subordinate male-female (SMvF) and dominant male-subordinate male (DMvSM) comparisons was analyzed using the Limma package (Smyth, 2005).

### 3.2.4. Mass spectrometry based methods

In studies I and II, I used mass spectrometry-based proteomics to acquire protein expression and identifications, respectively. Mass spectrometry based proteomics is at present a commonly used tool in model and non-model organism research (Buggiotti et al., 2008; Cox and Mann, 2007; Martínez-Fernández et al., 2008; Martyniuk and Denslow, 2009; Ong et al., 2003; Schulze and Usadel, 2010; Tyers and Mann, 2003). In addition to identifying which proteins are present in the sample, mass spectrometry can be used to determine protein expression levels similarly to microarrays. Furthermore, it is important to study the proteome (all expressed proteins in tissue) alongside the transcriptome (all expressed mRNA in tissue) in order to know which genes are actually expressed as their functional end products - the proteins (Sun et al., 2010). Studies have shown that that there does not exist a perfect one-to-one ratio between mRNA and protein abundance (Anderson and Seilhamer, 1997; Ning et al., 2012). This result highlights the importance of examining the transcriptome and the proteome together when studying gene expression in order to acquire a comprehensive picture of the entire process of gene expression.

In principle, mass spectrometry-based detection of proteins requires reduction of the sample complexity, enzymatic digestion of the full length protein to shorter peptide sequences, ionization of the peptide molecules followed by mass analysis of the peptide

molecules (Aebersold and Mann, 2003). As the sample complexity usually exceeds the detection limits of the instrumentation, reduction of the complexity is commonly done based on electrophoretic or chromatographic separation by size before (for proteins) or after digestion (for peptides) (Corthals et al., 2000). For the following steps of ionization and mass analysis, several types of methods are available. Given the scope of this thesis I will only discuss the methods used in my studies. The method used in both studies I and II was liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). Tandem mass spectrometry combines two rounds of mass detection; first round chooses a peptide and identifies its mass, and for the second round the chosen peptide is fractionated followed by mass analysis of the fractioned ions from the peptide. The mass spectrum which is obtained with the mass spectrometer is next queried against a database with theoretical mass spectra of proteins to identify the proteins present in the sample.

#### 3.2.4.1. Protein quantitation (study I)

In study I, protein quantification was used to identify liver proteins differentially expressed between sexes. The two most common approaches to protein quantitation are mass-labelling in which an isobaric mass tag is added to each sample prior to mass spectrometry analysis (e.g. iTRAQ, Wiese et al., 2007) and label free quantitation in which quantitation is based on counting the spectra observed for a given peptide (Zhu et al., 2010). In study I, quantitation was done with label free quantitation. Prior to quantitation, the proteins were first identified by searching the acquired spectra with Mascot Daemon v.2.2.2 (Matrix Science, London UK) against a custom prepared database with threespine stickleback proteins (Ensembl v. 57) as well as a list of common contaminants that could be introduced to the sample during processing (e.g. human keratin). For each protein in the database theoretical peptide mass spectra are created and the acquired mass spectra are compared against these to identify which peptides, and thus which proteins, are present in the sample. As the measurement of mass in the mass spectrometer is influenced by the instrument calibration and the chemical modifications introduced to the sample while processing it, these also need to be considered in the search algorithm. Label-free quantitation of the identified proteins between males and females was performed with normalized spectral abundance factor (NSAF) (Paoletti et al., 2006). In brief, the spectral abundance factor for a protein was determined by dividing the unique spectral count by the length of the protein and then normalizing this value by the sum of all spectral abundance factors of a sample. Protein length is incorporated as it has been shown to affect the number of peptides obtained after enzymatic digestion (Asara et al., 2008). Statistical testing for differential expression was performed with reproducibility optimized test statistic (ROTS, Elo et al. 2008, 2009), and linear modeling (LIMMA package, Smyth 2005). ROTS package has been proposed as statistical package for experiments with small number of biological replicates and to overcome the statistical power issues due to small

sample sizes significance is determined by optimizing the test statistic for the given the data by maximizing the reproducibility of the top identified proteins.

#### 3.2.4.2. Y-specific protein identification (study II)

In study II, protein identification was performed to identify which transcripts from the sex chromosome were translated in males and females. Special focus was on identifying peptides translated from the Y-copy alleles coming from the degenerating Y. To ensure detection of as many peptides as possible, each sample was fractionated into 10 fractions and proteins in the fractions were digested with two different enzymes.

For identifying Y-specific proteins in males, a custom database was created which contained, in addition to the Ensembl predictions from the reference genome (v. 73), the translated protein sequences of the Y-copy genes (with their associated mutations identified from the DNA capture inserted into the protein sequences, see section 3.2.). By comparing the peptides identified with mass spectrometer to the predicted male and female specific peptides from the database which are expected to be observed with the mass spectrometer given its detection range, I estimated the proportion of translated Y transcripts that were observable with mass spectrometry (see figure 2.); this comparison allows for estimating the amount of Y transcripts that are translated.

As Y-specific peptides were present in males, I further examined the effects of the Y-specific mutations on protein physiological properties, and thus potential function, by identifying the corresponding protein structures from the PDB database (Research Collaboratory for Structural Bioinformatics) and predicting the 2-dimensional structure for the Y proteins with Modeller v 9.12 (Eswar et al., 2006)and is usually facilitated by having an accurate three-dimensional (3-D. Effects of mutations for the protein were predicted with SNAP (Bromberg and Rost, 2007). Each mutation was further assigned to a specific structure (e.g α-sheet or backbone) with STRIDE (Heinig and Frishman, 2004) based on the structure from Modeller to estimate if the Y mutations are accumulating in certain parts of the protein structure.

## 3.3. Bioinformatics methods

### 3.3.1. Biomart and BLAST in gene functional annotation

Knowing the function of a gene, or a group of genes, in high throughput studies is worthwhile for gathering meaningful interpretations out of the flood of information. The most commonly used method to identify gene functions is functional annotation based on the Gene Ontology (GO) database. It is transferrable between species and thus allows annotating non-model species based on annotations from model species (The Reference Genome Group of the Gene Ontology Consortium, 2009). GO terms are divided into three categories which define them broadly; biological process, cellular compartment

and molecular function. Due to the annotation of the genome of threespine stickleback being incomplete, I used annotation based on orthologous human and zebra fish genes to improve the annotation of the threespine stickleback for downstream analysis in studies I, II and IV. For the annotation I first used orthologous gene information provided by Biomart (Ensembl, Kasprzyk, 2011). Genes for which no orthologs or annotation were found using Biomart, I queried the translated protein sequence against human (studies I and IV), zebra fish or the non-redundant database (study IV) provided by NCBI with Basic Local Alignment Search Tool (BLAST, Altschul et al., 1990).

The downside with BLAST-based annotation is the identification of false positives which can lead to a biased interpretation of gene functions if the false positive alignments are used for assigning GO annotation. False positives can be avoided by using more than one threshold or reciprocal searches (Primmer et al., 2013). To avoid false positive annotations I used an E-value cutoff of 0.001 in study I and manual curation. Because the numbers of genes were substantially higher in study IV than in study I (approximately 13 000 compared to only 342) and manual curation of all of them was not feasible, I used a combination of cutoffs: an orthologous gene was accepted with E-value <0.00001 and percentage of alignment >50. To support these criteria, assignment of orthologous genes from human and zebra fish were in agreement. Gene assignments from the non-redundant database were mainly sequences obtained from various non-model fish species and thus functional annotation was not possible for these.

### 3.3.2. Gene functional enrichment and pathway analysis

Functional annotation from gene ontology (GO) terms provides, depending on the gene, from one to several functions for a gene. With high throughput data, interpretation of each term alone is neither sufficient nor informative. Therefore, to overcome the problem with the amount of data, several methods to reduce the data to meaningful entities have arisen (Primmer et al., 2013). Functional enrichment is one of these methods, and it aims at finding similar functions, through the use of GO terms, that are overrepresented in a subset of genes (usually those differentially expressed) when compared to all the genes present in the sample (i.e. the background). Methods and tools for identifying enrichment of GO terms are common practice and provide insight into which biological processes, molecular mechanisms and cellular compartments are likely involved with the biological question at hand (Maere et al., 2005; Primmer et al., 2013). Selection for an appropriate entity of genes (i.e. background) against which enrichment in the identified subset of genes is tested, is crucial for the outcome of the enrichment test; erroneous interpretations can be made if the background does not match the study question at hand. Often, though, identifying only enrichment terms is not enough or the amount of data is still too much to comprehend; in these cases functional clustering based on term hierarchy and pathway analysis (Ramanan

et al., 2012) based on experimental evidence can yield more meaningful interpretations of the data.

In study I functional annotation and grouping of proteins was performed with Cytoscape (Cline et al., 2007) using the ClueGo (Bindea et al., 2009) plug-in. ClueGo groups and generates networks based on the functional relatedness of the enriched GO terms by forming functional GO clusters based on kappa statistics (i.e. term relatedness). As ClueGo bases functional annotation on the relatedness of the GO terms and not on the true interaction, I used Ingenuity Pathway Analysis (IPA, Ingenuity Systems, QIAGEN Redwood City, www.qiagen.com/ingenuity) to identify which curated pathways and networks are present among the identified proteins. IPA uses Ingenuity Knowledge Database which is a repository of biological interactions and functional annotations that is manually curated based on published research.

In study II enrichment of functional annotation terms was tested using DAVID (Huang et al., 2008). DAVID is similar to ClueGo in function; it identifies GO term enrichment in the data and also performs functional clustering based on kappa statistics. As the GO terms are redundant due to the hierarchical set up of the GO database, clustering was done with the redundancy reduced GO_fat terms available in DAVID (Huang et al., 2008).

In study IV Cytoscape plugin BiNGO (Cline et al., 2007; Maere et al., 2005) was used for testing for functional enrichment of GO terms as custom background gene lists can be easily used in the analysis. Similarly to study I, I used IPA for identifying which pathways are indicated by the differentially expressed genes in the breeding males and females. IPA also provides information on whether the gene expression patterns in the dataset are increasing or decreasing a specific biological function; the determination of the direction of change (i.e. increase or decrease) is based on the curated interactions of genes from previously published data in the Ingenuity Knowledge Database.

### 3.3.3. Gene variant effect prediction (II)

The severity of a given variant allele has on a gene function depends on whether the variant allele is located within a gene or in the intergenic area. To be able to predict the effects of the identified Y alleles from the capture data I used Variant Effect Predictor 2.8 (McLaren et al., 2010); it identifies all the possible effects of a mutation given the genome and annotation provided for the database (e.g. synonymous, frameshift, upstream or intergenic as determined by Sequence Ontology nomenclature, Eilbeck and Lewis, 2004). With the known inversion in the sex chromosomes of the threespine stickleback, I built a custom database using the custom reference genome (see section 3.5.1) and a custom gtf file where the reported inversion of the two super contigs was corrected (ensemble build version 71). The most severe effect that a variant had on a gene was retained for each gene when several effects for the gene were associated with the same variant (according to Ensembl's

categorization). A given variant can also have an effect on more than one gene at a time if the genes overlap the variant locus (e.g. due to the overlapping genes being on different strands).

### 3.4. Evolutionary analysis of the sex chromosome (study II)

I investigated the evolutionary rates of X and Y genes using dN/dS ratio by comparing the genes to an out-group species, the ninespine stickleback, *Pungitius pungitius* (Bell et al., 2009). Orthologous threespine and ninespine stickleback genes were determined by reciprocal searches with BLAST (Altschul et al. 1990). Sequences for alternative ninespine stickleback open reading frames and the threespine stickleback Y chromosome were manually built using the mutations from the ninespine transcriptomes and the DNA capture data. As codon alignment is crucial for accurate calculation of dN/dS, all the ninespine and threespine stickleback nucleotide sequences were translated to peptide sequence prior to the multiple alignment and then translated back to nucleotide sequence while retaining the codon alignment of the multiple alignment. In total 514 threespine genes from XIX and 3186 genes from autosomes with orthologous ninespine genes were identified. Firstly, I used pairwise comparison in PAML (Yang, 2007) to identify whether discrete gene categories (i.e. autosomal, PAR, X and Y) are exhibiting different selection pressures. Secondly, I investigated dN/dS ratios of specific genes in PAML to identify which ones are evolving faster in the Y than in the X chromosome.

# 4. MAIN RESULTS AND DISCUSSION

## 4.1. Sexual dimorphism at the protein level in the threespine stickleback (I)

In study I, sexual dimorphism was measured at the protein level in the threespine stickleback. I determined protein expression using label free quantitation of liver tissue for four males and four females from laboratory-reared adults which were not reproductively active. I identified 341 unique proteins with a 95% confidence level to be present in the samples. Functional categorization of the identified proteins matched the normal functions of liver; e.g. metabolism related to amino acids, carbohydrates and lipids, and also metabolism of xenobiotics.

Of the identified proteins 5.8% were differentially expressed between the sexes. This constituted qualitative and quantitative differences between the sexes. Qualitative comparison of the data revealed five proteins unique to one sex; two were observed only in males and three in females. As the qualitative analysis is affected by individual variability and small sample size, power to observe statistical significance is likely diminished. To avoid this, relaxed criteria were used; proteins were accepted as differentially expressed if they met the following criteria: an unadjusted p-value less than 0.05, fold change of at least 1.5, and they were identified by both ROTS and LIMMA. Altogether there were 15 proteins differentially expressed between the sexes; five female-biased and ten male-biased. Combining the quantitative and qualitative results, there were eight female-biased proteins and 12 male-biased proteins. Male-biased genes had functions related to amino acid metabolism and signaling whereas female-biased genes were related to protein synthesis.

As previous analysis of sex-biased liver transcriptome expression was available, I compared the protein level results with those results. Microarray analysis revealed over 11 000 genes expressed at the mRNA level from liver with 11.3% showing sex-biased expression (Leder et al., 2010). These numbers are quite different compared to the 342 proteins with 5.8% differentially expressed between sexes. Mainly the discrepancy between the results highlights some of the key differences between mass spectrometry-based proteomics and transcriptomics; while microarrays are targeting genes with the use of probes, shotgun proteomics is non-targeted with selection not based on certain peptides but rather on abundance of peptides (intensity). Also, at the time of this study, the difference between the sensitivity of microarrays and mass spectrometry instrumentation was responsible in part for the differences observed. In addition to technical effects, the biology will also contribute to the differences observed since the stability of mRNA and protein will have an effect on the probability of detecting a correlation between mRNA and protein abundance (Ning et al., 2012). These factors create inconsistencies between microarray-based mRNA measures and proteomic data (Cox and Mann, 2007; Ning et al., 2012).

Regardless of the limitations of the technique used in study I, I have established that sexual dimorphism previously identified at the liver transcriptome (Leder et al., 2010) is reflected to some extent in the proteome. A recent study by (Ning et al., 2012) showed that correlation between liver proteome (measured with NSAF) and transcriptome abundance was best when measured with RNA-Seq although differences between the other methods were not large. Although in study I, correlations between protein and mRNA abundance were on average around 0.5, the directions of change in expression between the two measurements were in agreement for 74% genes (i.e. if gene was up-regulated in mRNA it was also up-regulated in proteome). These types of comparisons will extend our understanding of how the phenotype is generated and comparisons between the transcriptome and the proteome will become more exact as development of instrumentation allows for better protein measurements (Cox et al., 2005; Tyers and Mann, 2003).

## 4.2. Effects of Y chromosome evolution on gene loss and persistence patterns (II)

In study II I investigated the level of Y chromosome degeneration and the effects of recombination suppression on the Y chromosome's gene content at three molecular levels: DNA (sequence capture), mRNA (RNA-Seq) and protein (mass spectrometry-based protein identification). The workflow of study II is briefly illustrated in Figure 2 (see section 3.2.2.2). Overall I found that the level of degeneration in the form of gene loss, transcript silencing and non-functionalization is relatively advanced in the threespine stickleback with 58% of genes either being lost or silenced from the Y; yet, accumulation of male-specific functions in the Y chromosome has not progressed as fast. I observed the previously identified regions of the threespine stickleback sex chromosome in my data (Ross and Peichel, 2008): a pseudoautosomal region where X and Y recombine (PAR), a region where recombination between X and Y is suppressed (stratum I) and a region with little homology between X and Y as it has largely been deleted from the Y (stratum II). Differences between the X and Y sequences (e.g. deletions) can be seen in the numbers of reads mapping in males and females for chromosome XIX. Thus, I examined the similarity of the male and female sex chromosomes by comparing the normalized average per base read coverage of males to that of the females for chromosome XIX (Figure 3.). When the ratio is close to one, the read depths of male and female are similar and suggests that males and females have a similar copy number (i.e. male has both X and Y sequences). Corresponding to the three described regions, normalized read coverage is quite similar between the sexes in the PAR and stratum I but is reduced in male in stratum II (Figure 3.). In stratum II, reduced male coverage coincides with deletions described for this region (Roesti et al., 2013; Ross and Peichel, 2008) although areas where the coverage between the male and female is quite similar were observed (Figure 3.).

As the focus in study II was on identifying novel male variation, alleles which may be fixed in the male but are still segregating in the female, were not considered. Following this conservative

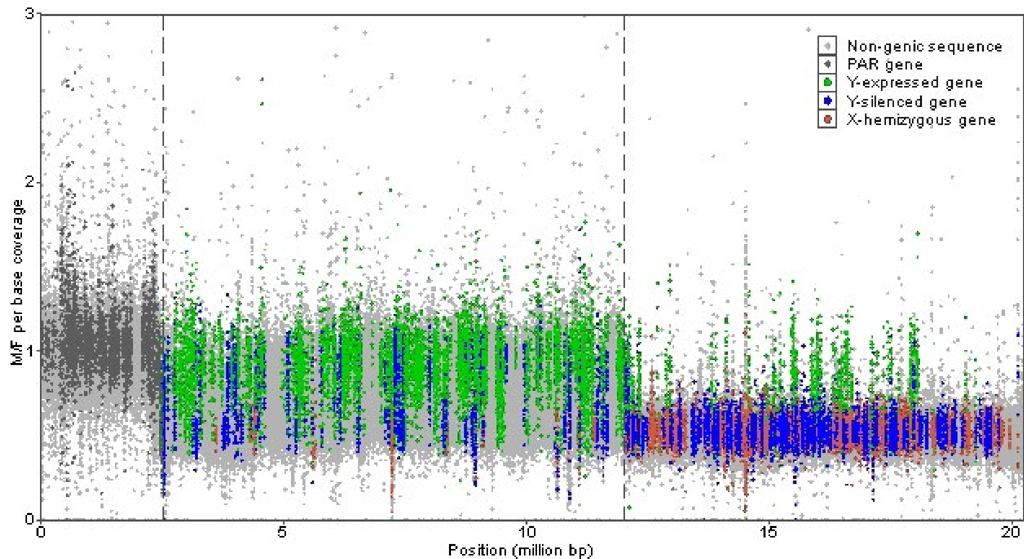
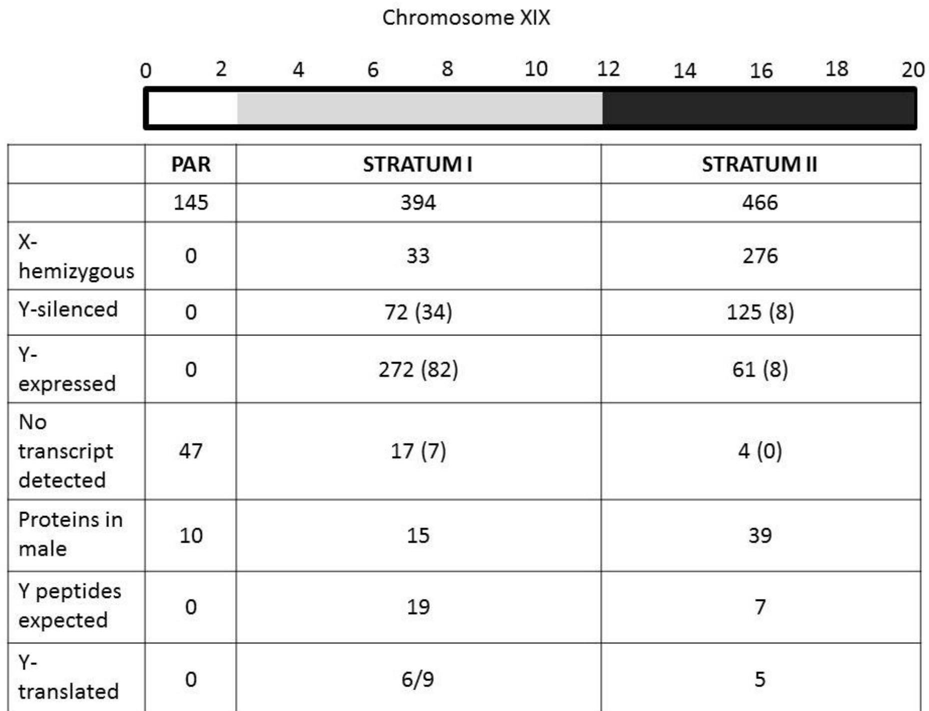



**Figure 3.** Normalized average per base read coverage for the captured males and females along chromosome XIX. A window of 100 bases was used to calculate average per base read coverage which was then normalized by the pseudoautosomal region of the corresponding sex. Horizontal axis represents the position along chromosome XIX. Dashed lines represent the borders between the PAR and stratum I, and the stratum I and stratum II, respectively. Light grey points represent non-genic sequence (intron or intergenic), dark grey points represent genes in the pseudoautosomal region, green represents genes that have a Y-copy which is expressed, blue represents genes that have a Y-copy which is not expressed and brown represents genes that have lost their Y-copy (i.e. hemizygous in males).

criterion, over 179 000 of the variable loci in chromosome XIX were determined to be variable due to Y-specific alleles. The three sex chromosome regions outlined above were also clear in the distribution of the Y alleles along sex chromosome (see Y-expressed genes in Figure 3.); the PAR extends to approximately 2.5 million bases after which the accumulation of Y alleles increases due to suppressed recombination in stratum I (Roesti et al., 2013). A clear drop in Y allele count around 12.0 million bases marks the boundary between stratum I and stratum II. Unexpectedly, Y alleles were identified also from stratum II. These were found primarily in gene regions and will be discussed in more detail later in the discussion.

With the Y allele information gathered from the DNA capture I was able to investigate the presence of Y-linked gene copies in the threespine stickleback using Ensembl gene predictions and RNA-Seq data for males and females. Of the genes in stratum I and stratum II 36% were regarded as X-hemizygous in males (i.e. Y-copy lost); 90 % of these hemizygous genes reside in stratum II which has suffered large deletions and thus provides an explanation for gene loss (Peichel et al., 2004; Ross and Peichel, 2008) (Figure 4.). When RNA-Seq data was examined, the Y transcript was silenced for 23% of genes in the two strata regardless of observing a Y-linked copy at the DNA level (i.e. Y-silenced) whereas for 39% of the genes the Y transcript




| | PAR | STRATUM I | STRATUM II |
|---|---|---|---|
| | 145 | 394 | 466 |
| X-hemizygous | 0 | 33 | 276 |
| Y-silenced | 0 | 72 (34) | 125 (8) |
| Y-expressed | 0 | 272 (82) | 61 (8) |
| No transcript detected | 47 | 17 (7) | 4 (0) |
| Proteins in male | 10 | 15 | 39 |
| Y peptides expected | 0 | 19 | 7 |
| Y-translated | 0 | 6/9 | 5 |

**Figure 4.** Patterns of gene loss and expression in the threespine stickleback sex chromosome (chromosome XIX). The three regions indicated are PAR; pseudoautosomal, stratum I; non-recombining and stratum II; non-homologous. Counts of Ensembl (v67) predicted protein coding genes in each region are given in first row, followed by numbers of X-hemizygous, Y-silenced, Y-expressed genes and genes for which no transcription was detected. In parenthesis is the number of genes with disrupted open reading frame due to Y specific variants in a given region. For proteins, counts of observed proteins in males, possible Y indicative peptides that are be detectable (based on mRNA and MS/MS instrumentation) and detected translated Y-specific peptides are given for each region to exemplify the discrepancy in numbers due to different methodologies (namely transcriptomics and proteomics). As the PAR is still recombining no Y-specific or X-hemizygous genes are detected for this region.

was transcribed (i.e. Y-expressed) (Figure 4.). As the transcription status of Y copies was determined with RNA-Seq, the probability of detecting the Y copy is dependent on whether the gene is expressed in a tissue in the first place; for only 2% of the predicted genes no expression, even in females, was observed. This indicates that the likelihood of not identifying the Y-copy gene, because the gene was not expressed in the first place, was low in this study.

Of the expressed Y-copies, 27% had reading frame disrupting mutations whereas the remaining genes had accumulated mutations not affecting the reading frame (e.g. non-synonymous and inframe deletions) (Figure 4.). Genes, for which the Y-linked copy had a disrupted reading frame, were expressed at a significantly lower level than genes with intact reading frames (t-test, p-values: brain 0.06, testes 0.03 and 0.04, pooled organs 0.009). This has also been observed for genes in the neo-Y chromosome of *D. miranda* (Kaiser et al., 2011; Zhou and Bachtrog, 2012). This could suggest that in the threespine stickleback

these lowly expressed transcripts are likely not that important when it comes to losing gene function during degeneration (see also 4.3. for discussion of study III). Similarly, mutations in these lowly expressed genes are tolerated because a sufficient amount of transcript could be produced by the X, or it could be that they are in the process of being phased out of use by lowering the expression level while a compensating mechanism evolves.

To estimate expression level of Y-linked gene copies in different tissues, I examined allele-specific expression patterns in males; as expected, the amount of Y-biased loci in the testes was significantly higher than that of X-biased loci (t-test, p-value $6.1x10^{-11}$) but not in the somatic tissues. Y-specific expression is expected for reproduction related tissues but not for somatic tissues (Kaiser et al., 2011; Skaletsky et al., 2003; Wilson and Makova, 2009; Zhou and Bachtrog, 2012)2012. Testes-specific Y-biased expression was evident with 35 genes in testes having Y-biased expression in all Y-specific loci of the gene (Study II Additional file 5). Based on gene ontology enrichment analyses it appears that there is no specific male functions that are preferentially retained (or lost) from the threespine stickleback Y at its present evolutionary state of degeneration. In contrast, in *D. Miranda,* specific classes of genes related to male functions and regulatory or developmental processes seem to be preferentially retained in the neo-Y given the functionality of the genes (Kaiser et al., 2011; Zhou and Bachtrog, 2012). Similarly in mammals, the Y chromosome has become enriched for genes with testes-related functions during its degeneration (Vallender and Lahn, 2004). Thus, the threespine stickleback Y chromosome has acquired testes specific expression patterns, as expected, but has not been completely evolved to contain only male-specific functions regardless of the high degree of Y chromosome degeneration.

The occurrence of Y alleles, and expressed Y gene copies, in stratum II implies that these locations have most likely become associated with  the non-recombining region during the structural rearrangements that have taken place in the Y chromosome (Ross and Peichel, 2008) (supported by the observation that patterns of fixed mutations are similar to stratum I, Figure 3. in study II). The observation that these genes have been retained in the Y after recombination became suppressed is also supported by dN/dS analysis of the genes in the sex chromosome; dN/dS ratio for stratum II Y genes was lower than for stratum I Y genes regardless of the genes being expressed or silenced, (Wilcox test, p-value 0.04 and 0.01 respectively, Figure 5.). On the other hand, stratum II genes were not different from their X-homologs in regard to dN/dS ratio (Wilcox test p-value 0.39 and 0.74 for Y-expressed and Y-silenced respectively). Similar to what I found in the threespine stickleback, X-linked genes with functional Y homologs are evolving slower (i.e. have lower dN/dS) than those X-linked genes with non-functional or lost Y homologs in humans and in *D. miranda* (Wilson Sayres and Makova, 2012; Zhou and Bachtrog, 2012). Additionally, stronger purifying selection could be preserving these Y-linked copies from stratum II within the degenerating Y. This preservation could be driven by the Y-linked gene copies having an overall higher expression level than hemizygous genes (see 4.3. for discussion of study III) which is similar to what

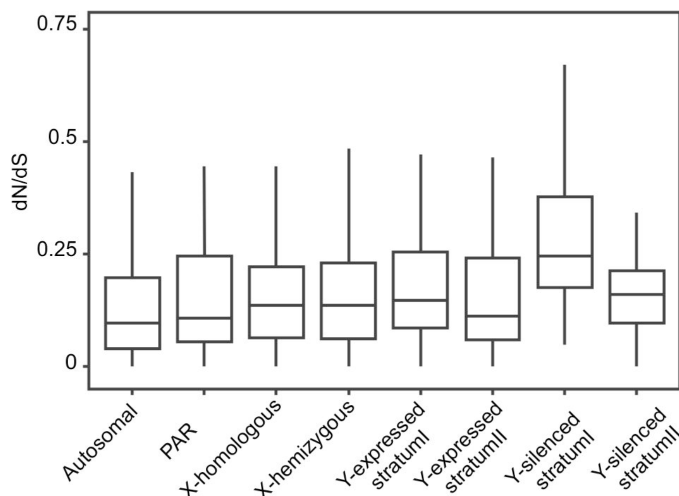



**Figure 5.** Pairwise dN/dS ratios obtained from comparison between orthologous threespine and ninespine stickleback genes. Threespine stickleback genes are divided according to location on X and the transcription status: from left to right the autosomal and pseudoautosomal (PAR) genes, X genes which have a homologous Y-copy and which do not have homologous Y-copy, followed by Y-expressed genes from stratum I and stratum II, and lastly Y-silenced genes from stratum I and stratum II. The lower and upper edge of the box represents the first and the third quartile with median indicated by the horizontal solid line. Vertical solid lines indicate the outer range of the distribution.

was recently shown in mammalian species. Genes with naturally high expression are harder to compensate for if one of the copies is lost, and thus these highly expressed genes are less likely to be lost during Y degeneration (Cortez et al., 2014).

To estimate expression level of Y-linked gene copies in different tissues, I examined allele-specific expression patterns in males; as expected, the amount of Y-biased loci in the testes was significantly higher than that of X-biased loci (t-test, p-value 6.1x10-11) but not in the somatic tissues. Y-specific expression is expected for reproduction related tissues but not for somatic tissues (Kaiser et al., 2011; Skaletsky et al., 2003; Wilson and Makova, 2009; Zhou and Bachtrog, 2012). Testes-specific Y-biased expression was evident with 35 genes in testes having Y-biased expression in all Y-specific loci of the gene (Study II Additional file 5). Based on gene ontology enrichment analyses it appears that there is no specific male functions that are preferentially retained (or lost) from the threespine stickleback Y at its present evolutionary state of degeneration. In contrast, in D. Miranda, specific classes of genes related to male functions and regulatory or developmental processes seem to be preferentially retained in the neo-Y given the functionality of the genes (Kaiser et al., 2011; Zhou and Bachtrog, 2012). Similarly in mammals, the Y chromosome has become enriched for genes with testes-related functions during its degeneration (Vallender and Lahn, 2004). Thus, the threespine stickleback Y chromosome has acquired testes specific expression patterns, as expected, but has not been completely evolved to contain only male-specific functions regardless of the high degree of Y chromosome degeneration.

Although cessation of recombination has produced new male alleles, many of these Y mutations do not seem to affect the physical properties of the protein; structural predictions showed that most of the mutations were neutral by effect and they were distributed in both the backbone and structural parts of the protein (α helix and β sheets) whereas non-neutral mutations were located mostly in the backbone of the protein when compared to α helicases or β sheets (study II Additional file 9). Although protein structure was unaffected, Y alleles could still affect the function of a protein (e.g. binding efficacy to substrate) which was not tested here. In humans both *TBL1* (Yan et al., 2005) and *DDX3* (Ditton, 2004) have acquired distinct mRNA and protein expression patterns across tissues of their Y copies compared to the X copy. Furthermore the Y copy of *DDX3* has acquired a novel function in spermatogenesis (Ditton, 2004). Given that we have encountered both coding and potential regulatory mutations, some Y genes can be expected to acquire their own function and expression pattern which are distinct from the X copy of the gene during the course of Y degeneration in the threespine stickleback. Male-specific proteins present in threespine stickleback are also potential targets for male-specific selection in threespine stickleback which can further drive the separation of the sex chromosomes or the sex-specific traits (Rice, 1987, 1992).

## 4.3. Dosage compensation in threespine stickleback (III)

As outlined in the introduction, sex chromosome evolution usually consists of three steps: suppression of XY (or ZW) recombination, degeneration of Y (or W) and dosage compensation (Charlesworth, 1996; Rice, 1996; Charlesworth et al., 2005). First two of these steps have been discussed already in the previous chapter (study II), namely effects of suppression of recombination and Y degeneration. In study III dosage compensation was investigated using RNA-Seq data obtained from brain tissue from four populations. Expression levels of sex chromosome and autosomes were compared in both sexes to identify dosage imbalances which are a result of the Y chromosome degeneration. The results of study III were twofold; although the overall ratio of median sex-linked and autosomal expression levels for the entire sex chromosome (1.03 for females and 0.87 for males, Figrue 6.) indicated incomplete dosage compensation in the male, region specific analysis of the sex chromosome revealed a lack of general dosage compensation within the male but local dosage compensation in stratum II.

The regions of the threespine stickleback sex chromosome, PAR, stratum I and stratum II, were investigated separately due to the different evolutionary backgrounds (as shown in study II). In PAR the male and female have nearly identical sex chromosome-autosome ratio (approximately 1.05 in both sexes). In stratum I the observed XXfemale:AA ratio is 0.89 and the XYmale:AA ratio is 0.86. After removing the effect of the Y chromosome (i.e. examining only the -genes expressed from the X chromosome) the resulting median

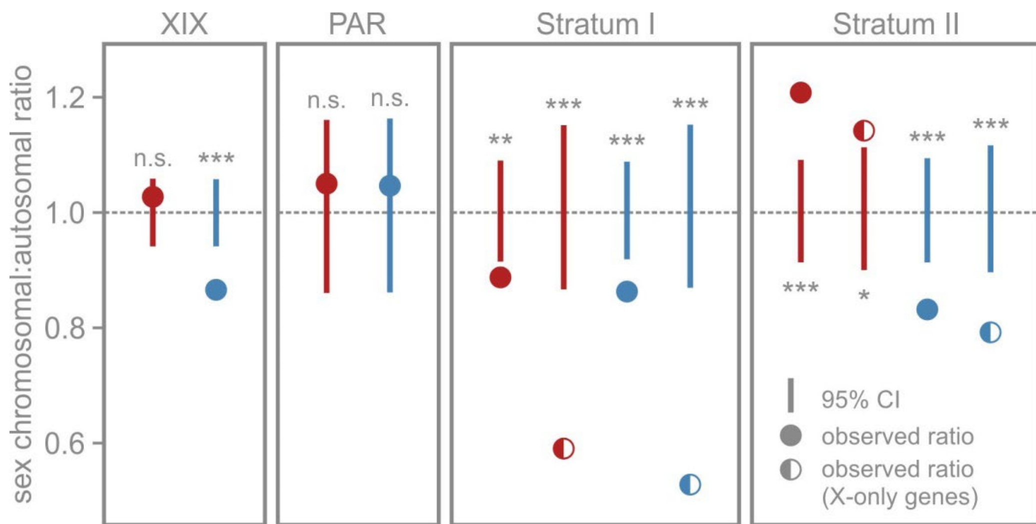



**Figure 6.** Results of the sex-specific differential expression analyses (XY:XX) for the whole chromosome XIX and for each of the three regions in the sex chromosome: the pseudoautosomal region (PAR), stratum I, and stratum II. Genes with male-biased expression levels are shown in blue and genes with female-biased expression levels in red. Filled circles indicate median expression for all chromosome XIX genes identified from the transcriptome (i.e. XX:AA in female and XY:AA in male) and half-filled circles indicate median expression for genes which are expressed only form the X in males (i.e. XX:AA in female and X:AA in male). Genes in the recombining PAR are present in two copies for both sexes.

Xmale:AA ratio is 0.53 which is slightly lower than the 0.59 median of the XXfemale:AA for the same set of genes X-only expressed genes . Removal of the Y-copy genes also shows that their expression has a substantial contribution to overall male expression level in stratum I (Figure 6.). In stratum II the observed median XXfemale:AA ratio was 1.21 while the median XYmale:AA ratio was 0.83, and after removing the effect of Y, the median Xmale:AA ratio was decreased to 0.79 whereas the XXfemale:AA decreased only to 1.14. The difference in Xmale:AA ratios between the two strata indicate local up-regulation for hemizygous genes in stratum II. As selection will favor up-regulation over no compensation to restore dosage imbalance created by sex chromosome degeneration (Hall and Wayne, 2013), the restriction of male X up-regulation in stratum II in threespine stickleback coincides with the observed large deletions in this stratum. As the threespine stickleback sex chromosome is relatively young (16-13 My) and compensation seems to be restricted to stratum II only, general male compensation can evolve over time in the Y to account for the expanding degeneration in the Y chromosome (Charlesworth, 1996, study II). The regional up-regulation of the male X chromosome in stratum II is likely a response to reduced gene expression due to the large, and likely sudden, deletions of Y chromosome sequence during the earlier steps of the Y chromosome evolution (Ross and Peichel, 2008).

In addition, substantial preservation of expressed Y-linked genes was observed in both strata (similar to study II). Examining the expression levels indicated that genes that have their Y

copy expressed have overall high expression level; log2 baseline expression in stratum I ranges from 4-9 and in stratum II from 5-9. Assuming that transcription rates are limited, males cannot compensate for the loss of a Y-copy gene which is initially highly expressed by further up-regulating their remaining X-copy, e.g. via modifier genes (Rice, 1984; Vicoso and Charlesworth, 2009). Thus highly expressed genes are sheltered from degeneration because loss of their expression cannot be easily compensated. Similar findings regarding the preservation of Y copies with high original expression level are found from neo-Y chromosome of Drosophila miranda (Kaiser et al., 2011), the human Y chromosome (Wilson Sayres and Makova, 2012), and across mammalian species (Cortez et al., 2014).

## 4.4. Hypothalamic transcriptome expression during reproduction in males and females (IV)

In study IV the aim was to quantitatively measure, and to functionally annotate, whole genome level transcriptome expression in the hypothalamus in relation to 1) reproduction and sex (i.e. dominant male-female and subordinate male-female) and 2) social status during breeding season (i.e. dominant male-subordinate male). Overall I found more differential expression related to reproduction and sex than for social status. Among the sex-specific genes, neuron cell structure and changes in the structures were found to be functionally enriched, and there was an indication for these functions to also be important for determining social status. Furthermore, functional annotation indicated that in threespine stickleback several genes with reproduction related functions were located in the autosomes whereas no specific enrichment was observed for genes in the sex chromosome.

Significant sex-biased expression was observed in the genome and a substantial proportion of it was female-biased and concentrated on the sex chromosome. Male-biased genes, however, were more evenly distributed along the genome (Figure 7.). This pattern of female-bias localized in the sex chromosome in hypothalamus is similar to the pattern observed for liver tissue between the sexes from a previous microarray experiment (Leder et al., 2010) and in the RNA-Seq data in study III. These findings highlight the fact that the female-biased expression in the sex chromosome is a general feature which is not restricted to liver tissue alone. In the dominant male-female comparison overrepresentation of male-biased genes was observed also for chromosomes IV, IX and XII ($\chi2$ test, unadjusted p= 0.046, 0.008 and 0.46, respectively) and for the subordinate male-female comparison for chromosomes XII, XVII and XXI ($\chi2$ test, unadjusted p = 0.017, 0.005 and 0.012, respectively).

Considering the dominant-female comparison, IX and XII are interesting in light of what is known of X1X2Y sex chromosome systems within Gasterosteus (Ross et al., 2009; Natri et al., 2013; Kitano et al., 2009). In G. wheatlandi the neo-Y is a result of fusion between chromosomes XIX and XII, and similarly in the Japan Sea G. aculeatus, a fusion has occurred between chromosomes XIX and IX. In the Japan Sea lineage, quantitative trait loci related

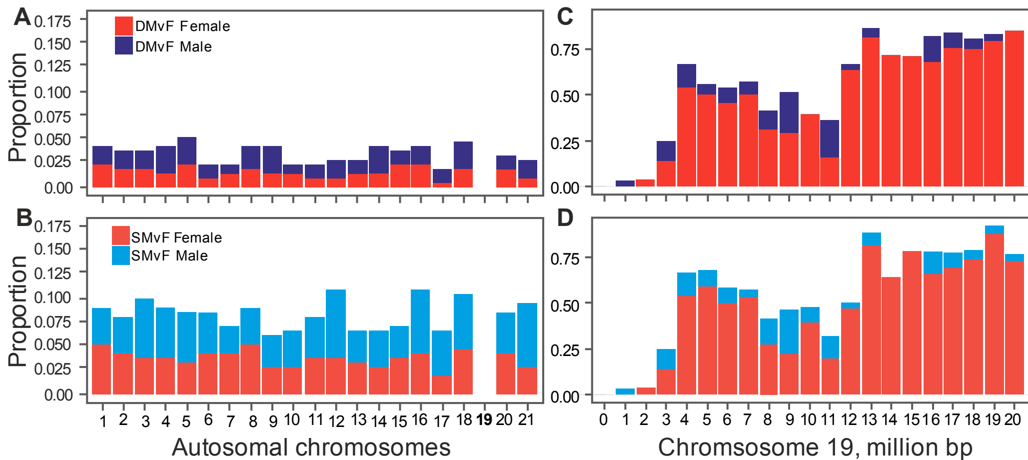



**Figure 7.** Accumulation of significantly differentially expressed genes within the autosomal chromosomes (A and B) and the sex chromosome (C and D) of the threespine stickleback in the dominant male-female (DMvF) and the subordinate male-female (SMvF) comparisons. For the dominant male-female A) and subordinate male-female B) autosomal chromosomes are indicated on the horizontal axis and proportion of differentially expressed male and female-biased genes of all genes expressed on the given chromosome is presented on the vertical axis. Sex chromosome (19) is omitted from A) and B) and is presented in detail in C) for the DMvF comparison and in D) for the SMvF comparison. For the sex chromosome, location along chromosome 19 is indicated on the horizontal axis in 1 million base bins. Proportions are presented on the vertical axes and are calculated as the number of up-regulated male and female genes compared to the total number of expressed genes within that bin. Female up-regulated genes from all comparisons are shown in red, dominant male up-regulated genes in DMvF comparisons are dark blue and subordinate male up-regulated genes in the SMvF comparisons are light blue.

to behavioral reproductive isolation (e.g. dorsal pricking behavior) have been shown to localize in chromosome IX (Kitano et al., 2009; Yoshida et al., 2014). Even though the enrichments on these chromosomes for male-biased genes are only marginally significant, the observation that two of these regions have independently fused to the Y chromosome in other stickleback systems (i.e. neo-Y) suggests that these regions are truly male-specific and are thus following the theory that Y chromosomes will accumulate male-specific functions. Furthermore, underrepresentation of male-biased genes in the X chromosome have been observed as a result of a female-biased mode of inheritance, Y chromosome degeneration and the accompanying dosage imbalances, especially in somatic tissues (not in gonads) (Reinius et al., 2012; Vicoso and Charlesworth, 2009). The lack of male-biased mRNA expression in the sex chromosome was observed in this study for hypothalamus and also for whole brain in study III.

Overall, hypothalamic gene expression in stickleback reflects similar functions as identified in other species. These results corroborate previous studies which have shown that hypothalamic gene expression is characterized by high energy metabolism, cell structure and communication, and transportation (Nishida et al., 2005a, 2005b; St-Amand et al.,

2012). An interesting pattern emerging from functional enrichment analysis of the sex-specific genes was neuronal growth and change in neuronal morphology which have been associated with reproductive maturation and social behaviors in mammals (Ferri and Flanagan-Cato, 2012; Sisk and Zehr, 2005). In teleost, this could serve as a potential future target of research related to reproduction. Although expression differences between males with different social status were not large, the result likely suggests plasticity of the social behavior in threespine stickleback as well as individual variation. Plasticity in behavior is likely driven by subtle changes in the brain structure (e.g. axon connections and dendritic growth) as indicated by the observed gene expression patterns for the dominant male-subordinate male comparison (Ferri and Flanagan-Cato, 2012; Pasterkamp and Giger, 2009). Interpretation of expression patterns of single genes was not as straightforward as the interpretation of general patterns of expression since genes are usually part of a regulatory cascade; understanding the change in one part alone can only give a partial understanding of the whole cascade. Thus for understanding the molecular background of breeding behaviors identifying gene expression patterns and the pathways to which they belong, as seen with my data, is recommended.

Although several genes were found to be differentially expressed between the sexes, no specific enrichment of GO terms was found on the sex chromosome even though the hemizygous genes, which could bias the result due to incomplete dosage compensation (study III), were removed. In autosomes, however, terms related to reproduction were observed in the functional enrichment analysis (study IV supplementary tables S4 and S5). Finding no specific functional enrichment in the sex chromosome of threespine stickleback is in line with the previous microarray study in liver tissue (Leder et al., 2010) and with the results of study II. However, it is in contrast to mammals where Y chromosomes harbor genes related to male-specific reproductive functions although it has been predicted that the original gene content of the autosome pair that evolved into the sex chromosome pair was likely quite different (Vallender and Lahn, 2004). As the sex chromosome of the stickleback is young, lack of enrichment of reproduction and gonad related functions from XIX (other than sex determination) likely reflect the original gene composition of chromosome XIX and translocation of sexually antagonistic male beneficial genes to Y might not have taken place yet. Alternatively, the observed pattern of autosomal presence of reproductive functions and male-biased expression (Figure 7.) could result from gene regulation imposed by the sex chromosome over the autosomes (i.e. modifier mechanism proposed by Rice, 1984). No support for enrichment of transcription factors in the sex chromosome, or increased expression of transcription factors, was observed in the data. This result could, nevertheless, be influenced by the female-bias in the sex chromosome as only 4 male-biased transcription regulators were observed compared to 17 female-biased (after removing the genes previously determined as X-hemizygous from the data). A recent study by Bellott et al. (2014) found that gene regulation is an important function

of X-Y genes that have been retained the human sex chromosomes. To properly identify transcription regulators, and their role for molecular sexual dimorphism in the threespine stickleback, requires an experimental setup which would allow for testing presence of *cis*- and *trans*-regulation in the genome.

# 5. IMPLICATIONS AND FUTURE DIRECTIONS

The four studies I have presented have provided examples of how one can exploit the range of molecular methods available to investigate evolutionary questions. My results have shown that sex-specific differential expression is present in the transcriptome and proteome of the threespine stickleback and it reflects differences in physiology between the sexes as determined by the functions of differentially expressed genes in liver and the hypothalamus. Regardless of tissue, the sex-specific expression patterns in the transcriptome (study III and IV) were affected by the young sex chromosome which is still in the early stages of evolution. This effect of the sex chromosome on sex-biased expression was not observed in the liver proteome in study I. Nevertheless, the male proteome has Y-specific proteins which are transcribed and translated while containing novel mutations; these novel alleles could serve as material for becoming male-specific genes if they develop male-beneficial features or expression patterns. There was evidence for the Y chromosome accumulating mutations in the gene regulatory areas which can further affect the expression patterns of genes, but examining this further was beyond the scope of this thesis.

In order to understand sex-specific selection and accumulation of male-specific functions in the Y chromosome, detailed research on the expressed Y proteins and their functional properties are also worthwhile since the mRNA expression alone might not represent the final functional stage. This is evident in study IV regarding isotocin and vasotocin; mRNA expression of these neuropeptides in my data does not reflect measurements at the protein level because the functional protein needs to be cleaved from the pro-peptide (Kleszczyńska et al. 2012). My thesis also shows that research on sex chromosome evolution and sexual dimorphism will benefit from broad-minded use of molecular techniques as they often complement each other and can provide insight into regulatory steps that happen in the transitions from DNA to mRNA and to protein as shown in study II for Y-specific gene copies. Furthermore, according to theory, the sex chromosomes in male would act as a regulators of autosomal expression (Rice, 1984) and this has been shown with the preservation of regulatory controllers in mammals during Y chromosome evolution (Bellott et al., 2014). Although I could not verify this with my data, experimental approaches to test for gene regulatory relationships do exist. One such approach is expression quantitative trait mapping (eQTL); it combines transcriptome expression with traditional QTL mapping. eQTL has been successfully applied in several taxa (Derome et al., 2008; Holloway et al., 2011; Morley et al., 2004; Whiteley et al., 2008), and in threespine stickleback it could provide the data to test for the role of chromosome XIX as a regulatory controller over autosomal genes to produce sexually dimorphic expression patterns.

For examining sex-specific expression, timing and tissue of choice are important. I observed that there are varying levels of sex-specific differential expression in the tissues I examined and it can be partly due to the functional properties of the tissue: a metabolic tissue, such as liver, versus a tissue which exerts control over other tissues like the hypothalamus. Metabolism is more continuously sex-specific whereas sex-specific control of reproductive tissues can be restricted to breeding season. As expression patterns between tissues vary, obtaining data from different tissues at different time points will complement the general picture of sex-specific expression. In species where sex is facultative (i.e. cannot be changed during life), it is expected that majority of the sex-specific differential expression happens at early development and thus differences which are observed later in life are primarily to maintain the sex-specific traits of an individual. In my studies I did not investigate gene expression during development and this should be an important area of study in the future for the threespine stickleback. Although gene expression can be used to identify potential sexual antagonism (or intralocus sexual conflict) in the genome (Innocenti and Morrow, 2010), it does not tell a complete story as expression is dependent on time (e.g. breeding season), place (e.g. tissue), the genetic architecture of the genome (e.g. sex chromosomes) and most importantly sexual selection (Bonduriansky and Chenoweth, 2009; Connallon and Clark, 2010; Meisel et al., 2012; Parsch and Ellegren, 2013). I observed dimorphic expression between males and females but to address the antagonistic nature of this dimorphic expression, one would need to experimentally test how the sex-specific expression of genes affects the fitness of males and females (Fairbairn and Roff, 2006). Proper experimental designs to elucidate the sexually antagonistic nature of sex-biased expression would be a worthwhile pursuit in the future. As many of the sexually dimorphic traits are well-characterized in the threespine stickleback, understanding the molecular background in even greater depth could help to resolve the interplay between sex chromosome evolution and development of sexual antagonism.

As several genes are retained for extended periods of time in the degenerating Y chromosome, investigating the functional properties of the accumulating mutations on the Y specific genes could provide insight on how and why dosage compensation develops, and how the gene content of the Y chromosome is modified over time for male benefit in XY sex chromosome systems. Furthermore, as the results of studies II and III show, there is a need to understand the dynamic nature of sex chromosome evolution when assessing degeneration and development of dosage compensation. Thus prospective future research could incorporate data from *G. wheatlandi* and the Japan Sea lineage of the threespine stickleback (Yoshida et al., 2014) to compare and contrast chromosome XIX and their respective neo-sex chromosomes with our results. This line of research could provide further empirical data on the dynamic nature of sex chromosome evolution and its potential effects on sexual dimorphism.

# 6. ACKNOWLEDGEMENTS

So many people have contributed to the completion of my thesis as well as to my growth as a researcher over the years I've spent in the Genetics lab; I hope I succeed in conveying my appreciation to all of you.

I guess you would not have read my thesis this far unless my supervisor, Erica Leder, hadn't given me an opportunity to "learn lab stuff" back in 2007. First you became my guide to the "lab work", then a supervisor of my master's thesis, and finally the supervisor of my thesis. I haven't been the easiest to supervise but you have managed to do that. You've given me space and patience, and tolerated my stubbornness. Although our conversations have been heated towards the end of my PhD, I've always appreciated you as a researcher and I cannot express my gratitude towards how much you've thought me, not just whether to hyphenate words or not. I think I might never have taken my skies from Tampere (where you have snow) to Turku (where you never know if you have snow) if it hadn't been for our cross-country skiing trips.

I would also like thank The Leder Group; although small yet very lively assembly of people. Roland Schultheiß and Fernando Ayillon, thank you for choosing Finland and Erica, and for your tricky questions about science, but also about life and dogs. You've also taught me to be more critical toward what I read and write. I also want to thank our research assistants: Sanna Pausio for the enormous pipetting efforts (although that data never ended up being used in my thesis) and Tiina Sävilammi for guiding me to the world of command lines from MS Excel. You girls have also brought so much laughter to my days, that I think people in the Genetic lab will be glad for the reduction in giggle levels. I would also like to thank Susan Balenger, Johanna Lehto and Krishna Parek who have spent a short time in the Leder Group.

I think that without the big umbrella group of Craig Primmer, the PnP, the atmosphere in the Genetics lab would not be the same. Craig has managed to create a very inspiring surrounding where a young scientist can grow up and be taken seriously, feeling as an important part of the group. An integral part of my PhD has been my officemates Veronika Laine, Ksenia Zueva and Pavel Matos. No matter what the topic, science or life, you were always there. Thank you for your inspiring company! I also want to thank Ville Aukee and Meri Lindqvist for having the patience to answer my questions whether they concerned ordering, equipment's or my defense. For inspiring "coffee room" conversations I would like to thank several past and present PnP members: Paula Lehtonen, Spiros Papakostas, Hannu Mäkinen, Olaf Thalman, Matthiue Bruneaux, Susan Johnston, Anti Vasemägi. For my fellow PhD students Megha Dash and Siim Kahar, thank you for sharing the agony of doing a PhD and stay strong for the final push of your thesis. One thing I definitely will be missing is laskiainen and sledding in Hirvensalo with PnPs but without an aching tailbone.

From the Genetics lab I would like to thank Raija Rouhiainen, Irma Saloniemi, Tatjana Saarinen, Christina and Seppo Nokkala, and Johanna Iljamo. I want to say thank you also to our lovely lunch group, Sanna, Elsi Pulkkinen and Katja Salminen. Our lunch conversations have made the last two years of PhD bearable by providing a moment to take my mind off of matters related to my PhD and the thesis, often to something more cute and furry. A special thanks to my fellow PhD Elsi; I hope I have managed to be as good of a support to you as you have been to me during the years. I can never forget that we spent a month working face to face at the post-pcr lab without mentioning our dogs until we saw each other at agility competition.

I have never been the most social person but I have managed to keep with me a few true friends from Tampere. So, thank you Jenni Lahtinen, for being my personal trainer and taking my mind off of work and for pushing me to be social from time to time by asking how I'm doing. Tiina Järvenpää and Anu Puistovaara; remember still Ally McBeal and those girls who at the age of 15 thought that 30 is the end of the world? Well it is not, and for sure it is not even half way. Although our lives have taken us to separate directions you have always been there for me and I appreciate it alot. When I moved to Turku people told me that getting to know people is not going to be easy; I disagree or then it does not apply to people with dogs. Atleast Liina Salonen, Sanna Mattila (although we officially met during my first year as a biology student in Jyväskylä, you both ended up in Turku as I did), Hanna-Mari Helama and Jenni Ripatti have shown that it is possible to find friends in Turku. So thank you for sharing the furry and the not-so-furry side of life with me. Fridays would not be anything without agility and I want to thank my trainees as well as my own training group.

Although a cliche I would not be here without my family. Äiti ja iskä, kiitos, että olette antaneet jokaisen meistä itse vapaasti valita, yrittää, erehtyä ja onnistua elämässä. Home in Ikuri has also provided a place to get away from science towards the end of my PhD. Eveliina and Mikael, thank you for taking my mind off of science during the marathon long board game nights in Ikuri.

This was the hardest part to write in the whole thesis as it required a lot of handkerchiefs as so much emotion was involved. Jouni, you have been my bedrock and stood by me through all the struggles I've had while pursuing my dream of becoming a biologist. You have never questioned my choice but have pushed me forward during the desperate moments. Kiitos, että olet vierelläni.

Lastly I would like to thank the Finnish Population Genetics Graduate School, Turku University Foundation, Emil Aaltonen Foundation and the Finnish Academy for funding during my PhD.

# 7. REFERENCES


Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198–207.

Aguirre, W.E., and Akinpelu, O. (2010). Sexual dimorphism of head morphology in three-spined stickleback *Gasterosteus aculeatus*. J. Fish Biol. *77*, 802–821.

Aguirre, W.E., Ellis, K.E., Kusenda, M., and Bell, M.A. (2008). Phenotypic variation and sexual dimorphism in anadromous threespine stickleback: implications for postglacial adaptive radiation. Biol. J. Linn. Soc. *95*, 465–478.

Ahluwalia, A. (2003). Sexual Dimorphism of Rat Liver Gene Expression: Regulatory Role of Growth Hormone Revealed by Deoxyribonucleic Acid Microarray Analysis. Mol. Endocrinol. *18*, 747–760.

Albert, A.Y.K., Sawaya, S., Vines, T.H., Knecht, A.K., Miller, C.T., Summers, B.R., Balabhadra, S., Kingsley, D.M., and Schluter, D. (2007). THE GENETICS OF ADAPTIVE SHAPE SHIFT IN STICKLEBACK: PLEIOTROPY AND EFFECT SIZE. Evolution *0*, 76–85.

Almeida, O., Gozdowska, M., Kulczykowska, E., and Oliveira, R.F. (2012). Brain levels of arginine–vasotocin and isotocin in dominant and subordinate males of a cichlid fish. Horm. Behav. *61*, 212–217.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Anders, S., and Huber, W. (2012). Differential expression of RNA-Seq data at the gene level–the DESeq package.

Anderson, L., and Seilhamer, J. (1997). A comparison of selected mRNA and protein abundances in human liver. Electrophoresis *18*, 533–537.

Asara, J.M., Christofk, H.R., Freimark, L.M., and Cantley, L.C. (2008). A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. PROTEOMICS *8*, 994–999.

Bachtrog, D. (2011). Plant Sex Chromosomes- A Non-Degenerated Y. Curr. Biol. *21*, R684–R685.

Barnett, A., and Bellwood, D.R. (2005). Sexual dimorphism in the buccal cavity of paternal mouthbrooding cardinalfishes (Pisces: Apogonidae). Mar. Biol. *148*, 205–212.

Bell, M.A., and Foster, S.A. (1994). The Evolutionary Biology of the Threespine Stickleback (New York: Oxford University Press).

Bell, M.A., Stewart, J.D., and Park, P.J. (2009). The World's Oldest Fossil Threespine Stickleback Fish. Copeia *2009*, 256–265.

Bellott, D.W., Hughes, J.F., Skaletsky, H., Brown, L.G., Pyntikova, T., Cho, T.-J., Koutseva, N., Zaghlul, S., Graves, T., Rock, S., et al. (2014). Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. Nature *508*, 494–499.

Bergero, R., and Charlesworth, D. (2011). Preservation of the Y Transcriptome in a 10-Million-Year-Old Plant Sex Chromosome System. Curr. Biol. *21*, 1470–1474.

Berlin, S., and Ellegren, H. (2005). Fast Accumulation of Nonsynonymous Mutations on the Female-Specific W Chromosome in Birds. J. Mol. Evol. *62*, 66–72.

Bernardo Carvalho, A., Koerich, L.B., and Clark, A.G. (2009). Origin and evolution of Y chromosomes: Drosophila tales. Trends Genet. *25*, 270–277.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H., Pages, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics *25*, 1091–1093.

Bonduriansky, R., and Chenoweth, S.F. (2009). Intralocus sexual conflict. Trends Ecol. Evol. *24*, 280–288.

Bromberg, Y., and Rost, B. (2007). SNAP: predict effect of non-synonymous polymorphisms on function. Nucleic Acids Res. *35*, 3823–3835.

Buchanan, K.L., and Catchpole, C.K. (2000). Song as an indicator of male parental effort in the sedge warbler. Proc. R. Soc. B Biol. Sci. *267*, 321–326.

Buggiotti, L., Primmer, C.R., Kouvonen, P., Bureš, S., Corthals, G.L., and Leder, E.H. (2008). Identification of differentially expressed proteins in Ficedula flycatchers. PROTEOMICS *8*, 2150–2153.

Carroll, S.B. (2008). Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. Cell *134*, 25–36.

Carroll, S.B. (2009). Endless forms: the evolution of gene regulation and morphological diversity. Philos. Biol. Anthol. 193.

Chapman, T., Arnqvist, G., Bangham, J., and Rowe, L. (2003). Sexual conflict. Trends Ecol. Evol. *18*, 41–47.

Charlesworth, B. (1991). The Evolution Of Sex Chromosomes. Science *251*, 1030–1033.

Charlesworth, B. (1996). The evolution of chromosomal sex determination and dosage compensation. Curr. Biol. *6*, 149–162.

Charlesworth, B., and Charlesworth, D. (2000). The degeneration of Y chromosomes. Philos. Trans. R. Soc. B Biol. Sci. *355*, 1563–1572.

Charlesworth, B., Coyne, J.A., and Barton, N.A. (1987). The relative rates of evolution of sex chromosomes and autosomes. Am. Nat. *130*, 113–146.

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

Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., Christmas, R., Avila-Campilo, I., Creech, M., Gross, B., et al. (2007). Integration of biological networks and gene expression data using Cytoscape. Nat. Protoc. *2*, 2366–2382.

Connallon, T., and Clark, A.G. (2010). SEX LINKAGE, SEX-SPECIFIC SELECTION, AND THE ROLE OF RECOMBINATION IN THE EVOLUTION OF SEXUALLY DIMORPHIC GENE EXPRESSION. Evolution *64*, 3417–3442.

Cortez, D., Marin, R., Toledo-Flores, D., Froidevaux, L., Liechti, A., Waters, P.D., Grützner, F., and Kaessmann, H. (2014). Origins and functional evolution of Y chromosomes across mammals. Nature *508*, 488–493.

Corthals, G.L., Wasinger, V.C., Hochstrasser, D.F., and Sanchez, J.-C. (2000). The dynamic range of protein expression: A challenge for proteomic research. Electrophoresis *21*, 1104–1115.

Cox, J., and Mann, M. (2007). Is Proteomics the New Genomics? Cell *130*, 395–398.

Cox, B., Kislinger, T., and Emili, A. (2005). Integrating gene and protein expression data: pattern analysis and profile mining. Methods *35*, 303–314.

Darwin, C.R. (1859). On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life (London: John Murray).

Dean, R., and Mank, J.E. (2014). The role of sex chromosomes in sexual dimorphism: discordance between molecular and phenotypic data. J. Evol. Biol. *27*, 1443–1453.

Derome, N., Bougas, B., Rogers, S.M., Whiteley, A.R., Labbe, A., Laroche, J., and Bernatchez, L. (2008). Pervasive Sex-Linked Effects on Transcription Regulation As Revealed by Expression Quantitative Trait Loci Mapping in Lake Whitefish Species Pairs (Coregonus sp., Salmonidae). Genetics *179*, 1903–1917.

Disteche, C.M. (2012). Dosage Compensation of the Sex Chromosomes. Annu. Rev. Genet. *46*, 537–560.

Ditton, H.J. (2004). The AZFa gene DBY (DDX3Y) is widely transcribed but the protein is limited to the male germ cells by translation control. Hum. Mol. Genet. *13*, 2333–2341.

Eilbeck, K., and Lewis, S.E. (2004). Sequence Ontology Annotation Guide. Comp. Funct. Genomics *5*, 642–647.

Ellegren, H., and Parsch, J. (2007). The evolution of sex-biased genes and sex-biased gene expression. Nat. Rev. Genet. *8*, 689–698.

Elo, L.L., Filen, S., Lahesmaa, R., and Aittokallio, T. (2008). Reproducibility-Optimized Test Statistic for Ranking Genes in Microarray Studies. IEEE/ACM Trans. Comput. Biol. Bioinform. *5*, 423–431.

Elo, L.L., Hiissa, J., Tuimala, J., Kallio, A., Korpelainen, E., and Aittokallio, T. (2009). Optimized detection of differential expression in global profiling experiments: case studies in clinical transcriptomic and quantitative proteomic datasets. Brief. Bioinform. *10*, 547–555.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., and Mitchell, S.E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS One *6*, e19379.

Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eremian, D., Shen, M., Pieper, U., and Sali, A. (2006). Comparative Protein Structure Modeling With MODELLER. Curr. Protoc. Bioinforma. *Supplement 15*.

Fairbairn, D.J., and Roff, D.A. (2006). The quantitative genetics of sexual dimorphism: assessing the importance of sex-linkage. Heredity *97*, 319–328.

Ferri, S.L., and Flanagan-Cato, L.M. (2012). Oxytocin and dendrite remodeling in the hypothalamus. Horm. Behav. *61*, 251–258.

FitzGerald, G.J., and Wootton, R.J. (1993). Behavioural ecology of sticklebacks. In Behavior of Teleost Fishes, (Chapman & Hall),.

Foster, S.A., Shaw, K.A., Robert, K.L., and Baker, J.A. (2008). Benthic, limnetic and oceanic threespine stickleback: profiles of reproductive behaviour. Behaviour *145*, 4–5.

Frayer, D.W., and Wolpoff, M.H. (1985). Sexual dimorphism. Annu. Rev. Anthropol. 429–473.

Genner, M.J., and Turner, G.F. (2005). The mbuna cichlids of Lake Malawi: a model for rapid speciation and adaptive radiation. Fish Fish. *6*, 1–34.

Gregg, C., Zhang, J., Butler, J.E., Haig, D., and Dulac, C. (2010). Sex-Specific Parent-of-Origin Allelic Expression in the Mouse Brain. Science *329*, 682–685.

Griffin, R.M., Dean, R., Grace, J.L., Ryden, P., and Friberg, U. (2013). The Shared Genome Is a Pervasive Constraint on the Evolution of Sex-Biased Gene Expression. Mol. Biol. Evol. *30*, 2168–2176.

Hall, D.W., and Wayne, M.L. (2013). Ohno's "Peril of Hemizygosity" Revisited: Gene Loss, Dosage Compensation, and Mutation. Genome Biol. Evol. *5*, 1–15.

Hedrick, A.V., and Temeles, E.J. (1989). The evolution of sexual dimorphism in animals: hypotheses and tests. Trends Ecol. Evol. *4*, 136–138.

Heinig, M., and Frishman, D. (2004). STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. Nucleic Acids Res. *32*, W500–W502.

Hellqvist, A., Schmitz, M., and Borg, B. (2008). Effects of castration and androgen-treatment on the expression of FSH-β and LH-β in the three-spine stickleback, *Gasterosteus aculeatus*—Feedback differences mediating the photoperiodic maturation response? Gen. Comp. Endocrinol. *158*, 178–182.

Higuchi, M., and Goto, A. (1996). Genetic evidence supporting the existence of two distinct species in the genus Gasterosteus around Japan. Environ. Biol. Fishes *47*, 1–16.

Holloway, B., Luck, S., Beatty, M., Rafalski, J.-A., and Li, B. (2011). Genome-wide expression quantitative trait loci (eQTL) analysis in maize. BMC Genomics *12*, 336.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. *4*, 44–57.

Innocenti, P., and Morrow, E.H. (2010). The Sexually Antagonistic Genes of Drosophila melanogaster. PLoS Biol. *8*, e1000335.

Isensee, J., and Noppinger, P.R. (2007). Sexually dimorphic gene expression in mammalian somatic tissue. Gend. Med. *4*, S75–S95.

Jakobsson, S., Borg, B., Haux, C., and Hyllner, S.J. (1999). An 11-ketotestosterone induced kidney-secreted protein: the nest building glue from male three-spined stickleback, *Gasterosteus aculeatus*. Fish Physiol. Biochem. *20*, 79–85.

Johnston, S.E., McEWAN, J.C., Pickering, N.K., Kijas, J.W., Beraldi, D., Pilkington, J.G., Pemberton, J.M., and Slate, J. (2011). Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. Mol. Ecol. *20*, 2555–2566.

Jones, I. (2001). Molecular Cloning and Characterization of Spiggin. AN ANDROGEN-REGULATED EXTRAORGANISMAL ADHESIVE WITH STRUCTURAL SIMILARITIES TO Von Willebrand FACTOR-RELATED PROTEINS. J. Biol. Chem. *276*, 17857–17863.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S., et al. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. Nature *484*, 55–61.

Joyce, A.R., and Palsson, B.Ø. (2006). The model organism as a system: integrating “omics” data sets. Nat. Rev. Mol. Cell Biol. *7*, 198–210.

Kaiser, V.B., and Bachtrog, D. (2010). Evolution of Sex Chromosomes in Insects. Annu. Rev. Genet. *44*, 91–112.

Kaiser, V.B., Zhou, Q., and Bachtrog, D. (2011). Nonrandom Gene Loss from the *Drosophila miranda* Neo-Y Chromosome. Genome Biol. Evol. *3*, 1329–1337.

Kasprzyk, A. (2011). BioMart: driving a paradigm change in biological data management. Database *2011*, bar049–bar049.

Kejnovsky, E., and Vyskot, B. (2010). *Silene latifolia*: The Classical Model to Study Heteromorphic Sex Chromosomes. Cytogenet. Genome Res. *129*, 250–262.

Kitano, J., Mori, S., and Peichel, C.L. (2007). Sexual Dimorphism in the External Morphology of the Threespine Stickleback (*Gasterosteus aculeatus*). Copeia *2007*, 336–349.

Kitano, J., Ross, J.A., Mori, S., Kume, M., Jones, F.C., Chan, Y.F., Absher, D.M., Grimwood, J., Schmutz, J., Myers, R.M., et al. (2009). A role for a neo-sex chromosome in stickleback speciation. Nature *461*, 1079–1083.

Kleszczyńska, A., Sokołowska, E., and Kulczykowska, E. (2012). Variation in brain arginine vasotocin (AVT) and isotocin (IT) levels with reproductive stage and social status in males of three-spined stickleback (*Gasterosteus aculeatus*). Gen. Comp. Endocrinol. *175*, 290–296.

Kondo, M. (2006). Genomic organization of the sex-determining and adjacent regions of the sex chromosomes of medaka. Genome Res. *16*, 815–826.

Kozak, M. (1992). Regulation of translation in eukaryotic systems. Annu. Rev. Cell Biol. *8*, 197–225.

Kwekel, J.C., Desai, V.G., Moland, C.L., Branham, W.S., and Fuscoe, J.C. (2010). Age and sex dependent changes in liver gene expression during the life cycle of the rat. BMC Genomics *11*, 675.

Lahn, B.T. (1999). Four Evolutionary Strata on the Human X Chromosome. Science *286*, 964–967.

Lande, R. (1980). Sexual dimorphism, sexual selection, and adaptation in polygenic characters. Evolution *34*, 292–305.

Leder, E.H., Cano, J.M., Leinonen, T., O’Hara, R.B., Nikinmaa, M., Primmer, C.R., and Merila, J. (2010). Female-Biased Expression on the X Chromosome as a Key Step in Sex Chromosome Evolution in Threespine Sticklebacks. Mol. Biol. Evol. *27*, 1495–1503.

Leinonen, T., Cano, J.M., MäKinen, H., and Merilä, J. (2006). Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. J. Evol. Biol. *19*, 1803–1812.

Leinonen, T., Cano, J.M., and Merilä, J. (2010). Genetic basis of sexual dimorphism in the threespine stickleback *Gasterosteus aculeatus*. Heredity *106*, 218–227.

Li, G., Bahn, J.H., Lee, J.-H., Peng, G., Chen, Z., Nelson, S.F., and Xiao, X. (2012). Identification of allele-specific alternative mRNA processing via transcriptome sequencing. Nucleic Acids Res. *40*, e104–e104.

Lin, Z., Wu, W.-S., Liang, H., Woo, Y., and Li, W.-H. (2010). The spatial distribution of cis regulatory elements in yeast promoters and its implications for transcriptional regulation. BMC Genomics *11*, 581.

Lindholm, A., and Breden, F. (2002). Sex chromosomes and sexual selection in poeciliid fishes. Am. Nat. *160*, S214–S224.

Livernois, A.M., Graves, J.A.M., and Waters, P.D. (2011). The origin and evolution of vertebrate sex

chromosomes and dosage compensation. Heredity *108*, 50–58.

Maan, M.E., and Seehausen, O. (2010). Mechanisms of species divergence through visual adaptation and sexual selection: Perspectives from a cichlid model system. Curr. Zool. *56*, 285–299.

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. Bioinformatics *21*, 3448–3449.

Maher, C.A., Palanisamy, N., Brenner, J.C., Cao, X., Kalyana-Sundaram, S., Luo, S., Khrebtukova, I., Barrette, T.R., Grasso, C., and Yu, J. (2009). Chimeric transcript discovery by paired-end transcriptome sequencing. Proc. Natl. Acad. Sci. *106*, 12353–12358.

Mäkinen, H.S., Cano, J.M., and Merilä, J. (2006). Genetic relationships among marine and freshwater populations of the European three-spined stickleback (*Gasterosteus aculeatus*) revealed by microsatellites. Mol. Ecol. *15*, 1519–1534.

Mäkinen, H.S., Cano, J.M., and Merilä, J. (2008). Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. Mol. Ecol. *17*, 3565–3582.

Malone, J.H., and Oliver, B. (2011). Microarrays, deep sequencing and the true measure of the transcriptome. BMC Biol. *9*, 34.

Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A., Howard, E., Shendure, J., and Turner, D.J. (2010). Target-enrichment strategies for next-generation sequencing. Nat. Methods *7*, 111–118.

Mank, J.E. (2013). Sex chromosome dosage compensation: definitely not for everyone. Trends Genet. *29*, 677–683.

Mank, J.E., Hall, D.W., Kirkpatrick, M., and Avise, J.C. (2006). Sex chromosomes and male ornaments: a comparative evaluation in ray-finned fishes. Proc. R. Soc. B Biol. Sci. *273*, 233–236.

Martínez-Fernández, M., Rodríguez-Piñeiro, A.M., Oliveira, E., Páez de la Cadena, M., and Rolán-Alvarez, E. (2008). Proteomic Comparison between Two Marine Snail Ecotypes Reveals Details about the Biochemistry of Adaptation. J. Proteome Res. *7*, 4926–4934.

Martyniuk, C.J., and Denslow, N.D. (2009). Towards functional genomics in fish using quantitative proteomics. Gen. Comp. Endocrinol. *164*, 135–141.

McGregor, P.K., Krebs, J.R., and Perrins, C.M. (1981). Song Repertoires and Lifetime Reproductive Success in the Great Tit (*Parus major*). Am. Nat. *118*, 149–159.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

McKinnon, J.S. (1996). Red coloration and male parental behaviour in the threespine stickleback. J. Fish Biol. *49*, 1030–1033.

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., and Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics *26*, 2069–2070.

McLean, T.I. (2013). “Eco-omics”: A Review of the Application of Genomics, Transcriptomics, and Proteomics for the Study of the Ecology of Harmful Algae. Microb. Ecol. *65*, 901–915.

McPherson, F.J., and Chenoweth, P.J. (2012). Mammalian sexual dimorphism. Anim. Reprod. Sci. *131*, 109–122.

Meiklejohn, C.D., Coolon, J.D., Hartl, D.L., and Wittkopp, P.J. (2014). The roles of cis-and trans-regulation in the evolution of regulatory incompatibilities and sexually dimorphic gene expression. Genome Res. *24*, 84–95.

Meisel, R.P., Malone, J.H., and Clark, A.G. (2012). Disentangling the relationship between sex-biased gene expression and X-linkage. Genome Res. *22*, 1255–1265.

Metzker, M.L. (2009). Sequencing technologies — the next generation. Nat. Rev. Genet. *11*, 31–46.

Miyagi, R., and Terai, Y. (2013). The diversity of male nuptial coloration leads to species diversity in Lake Victoria cichlids. Genes Genet. Syst. *88*, 145–153.

Mode, A., and Gustafsson, J.-Å. (2006). Sex and the Liver – A Journey Through Five Decades. Drug Metab. Rev. *38*, 197–207.

Moller, A.P., and Cuervo, J.J. (1998). Speciation and Feather Ornamentation in Birds. Evolution *52*, 859.

Morley, M., Molony, C.M., Weber, T.M., Devlin, J.L., Ewens, K.G., Spielman, R.S., and Cheung, V.G. (2004). Genetic analysis of genome-wide variation in human gene expression. Nature *430*, 743–747.

Natri, H.M., Shikano, T., and Merila, J. (2013). Progressive Recombination Suppression and Differentiation in Recently Evolved Neo-sex Chromosomes. Mol. Biol. Evol. *30*, 1131–1144.

Ning, K., Fermin, D., and Nesvizhskii, A.I. (2012). Comparative Analysis of Different Label-Free Mass Spectrometry Based Protein Abundance Estimates and Their Correlation with RNA-Seq Gene Expression Data. J. Proteome Res. *11*, 2261–2271.

Nishida, Y., Yoshioka, M., and St-Amand, J. (2005a). Sexually dimorphic gene expression in the hypothalamus, pituitary gland, and cortex. Genomics *85*, 679–687.

Nishida, Y., Yoshioka, M., and St-Amand, J. (2005b). The top 10 most abundant transcripts are sufficient to characterize the organs functional specificity:

evidences from the cortex, hypothalamus and pituitary gland. Gene *344*, 133–141.

Ong, S.-E., Foster, L.J., and Mann, M. (2003). Mass spectrometric-based approaches in quantitative proteomics. Methods *29*, 124–130.

Ozsolak, F., and Milos, P.M. (2010). RNA sequencing: advances, challenges and opportunities. Nat. Rev. Genet. *12*, 87–98.

Packer, C. (1983). Sexual dimorphism: the horns of African antelopes. Science *221*, 1191–1193.

Páll, M.K., Mayer, I., and Borg, B. (2002a). Androgen and Behavior in the Male Three-Spined Stickleback, *Gasterosteus aculeatus* I.—Changes in 11-ketotestosterone Levels during the Nesting Cycle. Horm. Behav. *41*, 377–383.

Páll, M.K., Mayer, I., and Borg, B. (2002b). Androgen and Behavior in the Male Three-Spined Stickleback, *Gasterosteus aculeatus* II.—Castration and 11-Ketoandrostenedione Effects on Courtship and Parental Care during the Nesting Cycle. Horm. Behav. *42*, 337–344.

Páll, M.K., Hellqvist, A., Schmitz, M., Olsson, P.-E., Mayer, I., and Borg, B. (2005). Changes in reproductive physiology and behaviour over the nesting cycle in male three-spined sticklebacks. J. Fish Biol. *66*, 1400–1410.

Paoletti, A.C., Parmely, T.J., Tomomori-Sato, C., Sato, S., Zhu, D., Conaway, R.C., Conaway, J.W., Florens, L., and Washburn, M.P. (2006). Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. Proc. Natl. Acad. Sci. *103*, 18928–18933.

Parker, G., and Partridge, L. (1998). Sexual conflict and speciation. Philos. Trans. R. Soc. B Biol. Sci. *353*, 261–271.

Parnell, N.F., and Streelman, J.T. (2013). Genetic interactions controlling sex and color establish the potential for sexual conflict in Lake Malawi cichlid fishes. Heredity *110*, 239–246.

Parsch, J., and Ellegren, H. (2013). The evolutionary causes and consequences of sex-biased gene expression. Nat. Rev. Genet. *14*, 83–87.

Pasterkamp, R.J., and Giger, R.J. (2009). Semaphorin function in neural plasticity and disease. Curr. Opin. Neurobiol. *19*, 263–274.

Peichel, C.L., Ross, J.A., Matson, C.K., Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Mori, S., Schluter, D., and Kingsley, D.M. (2004). The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. Curr. Biol. *14*, 1416–1424.

Pennell, T.M., and Morrow, E.H. (2013). Two sexes, one genome: the evolutionary dynamics of intralocus sexual conflict. Ecol. Evol. *3*, 1819–1834.

Petrie, M., Tim, H., and Carolyn, S. (1991). Peahens prefer peacocks with elaborate trains. Anim. Behav. *41*, 323–331.

Pike, T.W., Blount, J.D., Lindstrom, J., and Metcalfe, N.B. (2007). Dietary carotenoid availability influences a male's ability to provide parental care. Behav. Ecol. *18*, 1100–1105.

Pointer, M.A., Harrison, P.W., Wright, A.E., and Mank, J.E. (2013). Masculinization of Gene Expression Is Associated with Exaggeration of Male Sexual Dimorphism. PLoS Genet. *9*, e1003697.

Pressley, P.H. (1981). Parental Effort and the Evolution of Nest-Guarding Tactics in the Threespine Stickleback, *Gasterosteus aculeatus* L. Evolution *35*, 282.

Primmer, C.R., Papakostas, S., Leder, E.H., Davis, M.J., and Ragan, M.A. (2013). Annotated genes and nonannotated genomes: cross-species use of Gene Ontology in ecology and evolution research. Mol. Ecol. *22*, 3216–3241.

Ramanan, V.K., Shen, L., Moore, J.H., and Saykin, A.J. (2012). Pathway analysis of genomic data: concepts, methods, and prospects for future development. Trends Genet. *28*, 323–332.

Reinius, B., Johansson, M.M., Radomska, K.J., Morrow, E.H., Pandey, G.K., Kanduri, C., Sandberg, R., Williams, R.W., and Jazin, E. (2012). Abundance of female-biased and paucity of male-biased somatically expressed genes on the mouse X-chromosome. BMC Genomics *13*, 607.

Rice, W.R. (1984). Sex Chromosomes and the Evolution of Sexual Dimorphism. Evolution *38*, 735–742.

Rice, W.R. (1987). The Accumulation of Sexually Antagonistic Genes as a Selective Agent Promoting the Evolution of Reduced Recombination between Primitive Sex Chromosomes. Evolution *41*, 911–914.

Rice, W.R. (1992). Sexually Antagonistic Genes: Experimental Evidence. Science *256*, 1436–1439.

Rice, W.R. (1996). Evolution of the Y sex chromosome in animals. Bioscience *46*, 331–343.

Rice, W.R., and Chippindale, A.K. (2001). Intersexual ontogenetic conflict. J. Evol. Biol. *14*, 685–693.

Roberts, R.B., Ser, J.R., and Kocher, T.D. (2009). Sexual Conflict Resolved by Invasion of a Novel Sex Determiner in Lake Malawi Cichlid Fishes. Science *326*, 998–1001.

Roesti, M., Moser, D., and Berner, D. (2013). Recombination in the threespine stickleback genome-patterns and consequences. Mol. Ecol. *22*, 3014–3027.

Rosenzweig, B.A., Pine, P.S., Domon, O.E., Morris, S.M., Chen, J.J., and Sistare, F.D. (2004). Dye-Bias Correction in Dual-Labeled cDNA Microarray Gene Expression Measurements. Environ. Health Perspect. *112*, 480–487.

Ross, J.A., and Peichel, C.L. (2008). Molecular Cytogenetic Evidence of Rearrangements on the Y Chromosome of the Threespine Stickleback Fish. Genetics *179*, 2173–2182.

Ross, J.A., Urton, J.R., Boland, J., Shapiro, M.D., and Peichel, C.L. (2009). Turnover of Sex Chromosomes in the Stickleback Fishes (Gasterosteidae). PLoS Genet. *5*, e1000391.

Roy, A.K., and Chatterjee, B. (1983). Sexual dimorphism in the liver. Annu. Rev. Physiol. *45*, 37–50.

Schlinger, B.A., Greco, C., and Bass, A.H. (1999). Aromatase activity in hindbrain vocal control region of a teleost fish: divergence among males with alternative reproductive tactics. Proc. R. Soc. B Biol. Sci. *266*, 131–136.

Schulz, K.M., and Sisk, C.L. (2006). Pubertal hormones, the adolescent brain, and the maturation of social behaviors: Lessons from the Syrian hamster. Mol. Cell. Endocrinol. *254-255*, 120–126.

Schulze, W.X., and Usadel, B. (2010). Quantitation in Mass-Spectrometry-Based Proteomics. Annu. Rev. Plant Biol. *61*, 491–516.

Sisk, C.L., and Zehr, J.L. (2005). Pubertal hormones organize the adolescent brain and behavior. Front. Neuroendocrinol. *26*, 163–174.

Skaletsky, H., Kuroda-Kawaguchi, T., Minx, P.J., Cordum, H.S., Hillier, L., Brown, L.G., Repping, S., Pyntikova, T., Ali, J., and Bieri, T. (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature *423*, 825–837.

Slijkhuis, H., De Ruiter, A.J.H., Baggerman, B., and Wendelaar Bonga, S.E. (1984). Parental fanning behavior and prolactin cell activity in the male three-spined stickleback *Gasterosteus aculeatus* L. Gen. Comp. Endocrinol. *54*, 297–307.

Smyth, G.K. (2005). Limma: linear models for microarray data. In Bioinformatics and Computational Biology Solutions Using R and Bioconductor, (New York: Springer), pp. 397–420.

Sparkes, T.C., Rush, V., and Foster, S.A. (2008). Reproductive costs, condition and carotenoid-based colour in natural populations of threespine stickleback (*Gasterosteus aculeatus*). Ecol. Freshw. Fish *17*, 292–302.

St-Amand, J., Yoshioka, M., Tanaka, K., and Nishida, Y. (2012). Transcriptome-Wide Identification of Preferentially Expressed Genes in the Hypothalamus and Pituitary Gland. Front. Endocrinol. *2*, 1–16.

Stewart, A.D., Pischedda, A., and Rice, W.R. (2010). Resolving Intralocus Sexual Conflict: Genetic Mechanisms and Time Frame. J. Hered. *101*, S94–S99.

Sun, N., Pan, C., Nickell, S., Mann, M., Baumeister, W., and Nagy, I. (2010). Quantitative Proteome and Transcriptome Analysis of the Archaeon *Thermoplasma acidophilum* Cultured under Aerobic and Anaerobic Conditions. J. Proteome Res. *9*, 4839–4850.

Takahashi, T., Sota, T., and Hori, M. (2013). Genetic basis of male colour dimorphism in a Lake Tanganyika cichlid fish. Mol. Ecol. *22*, 3049–3060.

The Reference Genome Group of the Gene Ontology Consortium (2009). The Gene Ontology's Reference Genome Project: A Unified Framework for Functional Annotation across Species. PLoS Comput. Biol. *5*, e1000431.

Tinbergen, N. (1951). The study of instinct (Oxford University Press).

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. *7*, 562–578.

Tyers, M., and Mann, M. (2003). From genomics to proteomics. Nature *422*, 193–197.

Valencia-Sanchez, M.A. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev. *20*, 515–524.

Valenzuela, N. (2008). Sexual Development and the Evolution of Sex Determination. Sex. Dev. *2*, 64–72.

Valle, A., Guevara, R., Garcia-Palmer, F.J., Roca, P., and Oliver, J. (2007). Sexual dimorphism in liver mitochondrial oxidative capacity is conserved under caloric restriction conditions. AJP Cell Physiol. *293*, C1302–C1308.

Vallender, E.J., and Lahn, B.T. (2004). How mammalian sex chromosomes acquired their peculiar gene content. BioEssays *26*, 159–169.

Vicoso, B., and Charlesworth, B. (2009). The Deficit of Male-Biased Genes on the D. melanogaster X Chromosome Is Expression-Dependent: A Consequence of Dosage Compensation? J. Mol. Evol. *68*, 576–583.

Wang, E., and Purisima, E. (2005). Network motifs are enriched with transcription factors whose transcripts have short half-lives. Trends Genet. *21*, 492–495.

Wang, J., Na, J.-K., Yu, Q., Gschwend, A.R., Han, J., Zeng, F., Aryal, R., VanBuren, R., Murray, J.E., and Zhang, W. (2012). Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. Proc. Natl. Acad. Sci. *109*, 13710–13715.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. *10*, 57–63.

Waxman, D.J., and Holloway, M.G. (2009). Sex Differences in the Expression of Hepatic Drug Metabolizing Enzymes. Mol. Pharmacol. *76*, 215–228.

Whiteley, A.R., Derome, N., Rogers, S.M., St-Cyr, J., Laroche, J., Labbe, A., Nolte, A., Renaut, S., Jeukens, J., and Bernatchez, L. (2008). The Phenomics and

Expression Quantitative Trait Locus Mapping of Brain Transcriptomes Regulating Adaptive Divergence in Lake Whitefish Species Pairs (Coregonus sp.). Genetics *180*, 147–164.

Wiese, S., Reidegeld, K.A., Meyer, H.E., and Warscheid, B. (2007). Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. PROTEOMICS *7*, 340–350.

Williams, M.B. (1978). Sexual selection, adaptation, and ornamental traits: The advantage of seeming fitter. J. Theor. Biol. *72*, 377–383.

Williams, T.M., Selegue, J.E., Werner, T., Gompel, N., Kopp, A., and Carroll, S.B. (2008). The Regulation and Evolution of a Genetic Switch Controlling Sexually Dimorphic Traits in Drosophila. Cell *134*, 610–623.

Wilson, M.A., and Makova, K.D. (2009). Evolution and Survival on Eutherian Sex Chromosomes. PLoS Genet. *5*, e1000568.

Wilson Sayres, M.A., and Makova, K.D. (2012). Gene Survival and Death on the Human Y Chromosome. Mol. Biol. Evol. *30*, 781–787.

Wittkopp, P.J., Haerum, B.K., and Clark, A.G. (2004). Evolutionary changes in cis and trans gene regulation. Nature *430*, 85–88.

Woodgate, J.L., Mariette, M.M., Bennett, A.T.D., Griffith, S.C., and Buchanan, K.L. (2012). Male song structure predicts reproductive success in a wild zebra finch population. Anim. Behav. *83*, 773–781.

Wright, A.E., and Mank, J.E. (2013). The scope and strength of sex-specific selection in genome evolution. J. Evol. Biol. *26*, 1841–1853.

Yan, H.-T., Shinka, T., Kinoshita, K., Sato, Y., Umeno, M., Chen, G., Tsuji, K., Unemi, Y., Yang, X.-J., Iwamoto, T., et al. (2005). Molecular analysis of TBL1Y, a Y-linked homologue of TBL1X related with X-linked late-onset sensorineural deafness. J. Hum. Genet. *50*, 175–181.

Yang, X. (2006). Tissue-specific expression and regulation of sexually dimorphic genes in mice. Genome Res. *16*, 995–1004.

Yang, Z. (2007). PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol. Biol. Evol. *24*, 1586–1591.

Yang, Y.H., Dupoit, S., Luu, P., Lin, D.M., Vivian, P., Ngai, J., and Speed, T.P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res. *30*, 10.

Yoshida, K., Makino, T., Yamaguchi, K., Shigenobu, S., Hasebe, M., Kawata, M., Kume, M., Mori, S., Peichel, C.L., Toyoda, A., et al. (2014). Sex Chromosome Turnover Contributes to Genomic Divergence between Incipient Stickleback Species. PLoS Genet. *10*, e1004223.

Zhou, Q., and Bachtrog, D. (2012). Sex-Specific Adaptation Drives Early Sex Chromosome Evolution in Drosophila. Science *337*, 341–345.

Zhou, Q., Wang, J., Huang, L., Nie, W., Wang, J., Liu, Y., Zhao, X., Yang, F., and Wang, W. (2008). Neo-sex chromosomes in the black muntjac recapitulate incipient evolution of mammalian sex chromosomes. Genome Biol *9*, R98.

Zhu, W., Smith, J.W., and Huang, C.-M. (2010). Mass Spectrometry-Based Label-Free Quantitative Proteomics. J. Biomed. Biotechnol. *2010*, 1–6.