The genomic architecture of sex-biased gene expression in *Xenopus borealis*

The genomic architecture of sex-biased gene expression in *Xenopus borealis*

By Xue-Ying Song, HONOURS B.SC

A Thesis Submitted to the School of Graduate Studies in the Partial Fulfillment of the Requirements for the Degree Masters of Science

McMaster University
Masters of Science (2019)
Hamilton, Ontario (Department of Biology)

TITLE: The genomic architecture of sex-biased gene expression in *Xenopus borealis*
AUTHOR: Xue-Ying Song (McMaster University)
SUPERVISOR: Dr. Ben Evans
NUMBER OF PAGES: xiii, 85

Lay abstract

Sexual selection favours the evolution of distinctive traits in each sex in order to optimize the reproductive success of each one. However, because most of the genome is shared between the sexes, sexual selection may result in genomic conflict when mutations are beneficial to one sex but harmful to the other; this conflict is known as sexual antagonism. Genomic conflict associated with alleles with sexually antagonistic (SA) fitness effects can be resolved via the origin of sex-biased expression patterns and this may be catalyzed by genetic linkage to a sex-determining locus on a sex chromosome. Consequently, one might predict there to be an enrichment of genes with sex-biased expression patterns on the sex chromosome as compared to the autosomes. We tested this expectation in an African frog species *Xenopus borealis*, which has a relatively young sex chromosomes and a large region of recombination suppression on the female-specific W-chromosome. We found enrichment of sex-biased genes on the nonrecombining region of the sex chromosomes of this species in adult liver and gonad tissue and also tadpole mesenephros/gonads at two developmental stages. Additionally, we found that expression divergence of genes in the non-recombining region have a faster rate of evolution as compared to the rate of expression divergence of genes in other genomic regions. One possible explanation for these observations is that natural selection favours an expansion of recombination suppression (via unknown mechanisms) on sex chromosome such that polymorphic regulatory region become linked (or unlinked) to the sex determining locus in such a way as to resolve SA.

Abstract

Most vertebrates have separate sexes, and sex-specific traits that are regulated by genes with sex-biased expression patterns. In many species with genetic sex determination system, genetic recombination is suppressed in genomic regions linked to the master regulator of sex determination – the gene or set of linked loci that orchestrate sexual differentiation. Natural selection may favour alleles with sex-specific effects - including those with sexually antagonistic (SA) fitness effects (e.g., beneficial to females but harmful to males) – to become fixed in or be translocated to these non-recombining regions of sex chromosomes, because sex-specific or sex-biased modes of inheritance can resolve genomic conflict associated with SA. Sexually antagonism may also be resolve by sex-biased gene expression, and in theory these two mechanism (sex-linkage and sex-biased gene expression) could operate synergistically. However, there are relatively few empirical studies that test whether genes with sex biased expression patterns are indeed more abundant on sex chromosomes – and especially on newly evolved sex chromosomes. We explored this question with an African frog species *Xenopus borealis*, whose sex chromosome evolved within the last 25 million years (my) and have a large (50Mbp) region of suppressed recombination, making it a young sex chromosome system compared to many other intensively studied systems, such as the sex chromosomes of mammals. We tested the possibility that a higher proportion of genes with sex-biased expression would be located on the sex-linked region of the sex chromosome of this species. By examining gene expression in adult liver and gonad and also tadpole gonad/mesenephros at two developmental stages, we found that the sex-linked region of these sex chromosome do have a higher proportion of sex biased genes compared to the non-sex-linked region of the same sex chromosomes, compared to (i) a homeologous genomic region in the tetraploid genome of *X. borealis*, and also (ii) the autosomes of this species. We did not observe the same pattern in a closely related frog species, *Xenopus laevis*, which has sex chromosome that are not homologous to those of *X. borealis* and, unlike *X. borealis*, lacks a large region of suppressed recombination on its sex chromosome. Using Brownian Motion model, we found as well that expression divergence evolution of genes in the sex-linked region of *X. borealis* is faster compared to its non-sex-linked homeologs (within *X.borealis*), and also compared to orthologous regions that are also non-sex-linked. One possible explanation for these observations is that natural selection favoured an expansion of recombination suppression (via unknown mechanisms) on chromosome such that polymorphic regulatory regions became linked (or unlinked) to the sex determining locus in such a way to resolve SA. Alternatively, it is possible that these sex-biased expression pattern evolved rapidly after recombination suppression.

Acknowledgements

This project has been an incredible learning experience, one cannot be possible without all the supports.

To my family, thank you for your unconditional supports even though you might not understand exactly what I am doing in school, but you always have my back through out my entire journey. To my friends in Hamilton or not in Hamilton, you are the life part in my study/work-life balance. I really enjoyed the time that we spent together.

To my Evans lab members, Ben Furman, Caroline Cauret, and Martin Knytl, I learn a little bit from you everyday. I really enjoyed the passionate conversations between Ben and Caroline when I started grad school, those motivated me to learn more. Thank you, Ben, for walking me through the first year of project while I needed a lot of additional explanation during and after our Skype conversation with our advisor.

To my committee member, Dr. Ian Dworkin, thank you for all the advice and helps. I really appreciated the statistic explanations, which you probably always need to go back and start explaining from the fundamental level for me.

Finally, to my advisor, Dr. Ben Evans, I could not thank you enough for providing me the opportunity to learn from you and from doing projects starting from my undergraduate thesis project. You have been incredible in providing opportunities and supports (with those speedy email replies). Your cheerful altitude toward work made everything feels easier than it seems. ...

Contents

List of Figures

List of Tables

Acronyms

SA *sexually antagonistic*

X. borealis *Xenopus borealis*

X. laevis *Xenopus laevis*

Declaration of Authorship

I, Xue-Ying Song, declare that this thesis titled, "The genomic architecture of sex-biased gene expression in *Xenopus borealis*" and the work presented in it are my own.

I declare this thesis to be an original report of my research, except where indicated or cited. No part of this work has been submitted, in whole or in part, in any previous applications or publications for a degree at another institution.

This work represents original research conducted by myself in collaboration with Dr. Ben Furman and Dr. Martin Knytl under the supervision of Dr. Ben Evans and with advice input from Dr. Ian Dworkin. Ben Furman and Ben Evans contributed the design of the study, performed RNA extraction for adult tissue, and/or processed RNAseq data, assembled a transcriptome for adult tissue, and estimated a phylogenetic tree that was used for the Brownian motion analysis. Martin Knytl did the RNA extraction for tadpole gonads.

The research conducted in this thesis will be prepared as a manuscript for submission to a peer-reviewed journal.

Chapter 1

1 Introduction

1.1 Sexual selection

A myriad of phenotypes (e.g., external coloration, fin size, vocalization, horn size) can differ between males and females of the same species. Realizing that sex-specific fitness optima presented a challenge to natural selection, Darwin (1872) proposed the idea of sexual selections, wherein the evolution of distinctive traits in each sex optimizes reproductive successes of each one. Selection is thought to act through via intrasexual selection (e.g., male-male competition to gain mating access to female) and intersexual sexual selection (e.g., female choice on mating partners) (Darwin 1872). For example, combat weapons such as the enlarged antlers of male deer and horns of male beetles (e.g. scarabs – superfamily Scarabaeoidea) are likely consequence of intrasexual selection for excluding competitors and gaining better access to mates (Emlen 2008). Horn length of male horned dung beetles is positively correlated with success in male:male combat and with increased reproductive success (Emlen 2008). Conspicuous morphology, colouration and behaviour of peacocks and Manakins are thought to be examples of intersexual selection (Prum 1990; Petrie 1994) wherein female preference for exaggerated traits increases male reproductive success. Females might favour mates that provide fitness benefits, by for example, by offering access to resources or protection (Andersson 1994) . Female preference might also favour male traits that increase offspring fitness (Andersson 1994) by association with genetic variation that increases offspring survival (e.g., disease resistance gene in birds; Hamilton and Zuk 1982) or offspring reproductive success (Fisher 1930). The latter example could involve the fitness costs to one offspring sex while conferring benefits to the other under which scenario the trait would have sexually antagonistic (SA) fitness effect (Rice 1984; Seger and Trivers 1986; Rice and Chippindale 2001).

1.2 Sexual antagonism

Although female and male individuals may have distinctive fitness optima for a given trait, natural selection acts on phenotypes that are encoded by genomes, which are almost identical between the sexes. This creates an evolutionary tug-of-war at the genetic

level known as genomic conflict, wherein sexual selection exerts conflicting selective pressures on the two sexes (Parker 1979). Each sex are trying to reach their own fitness optima even at a cost to the other sexes (Parker 1979). SA mutations can occur on autosomes or sex chromosomes. Neutral, beneficial, and deleterious mutations all have the potential to go to fixation (Patwa and Wahl 2008) but the probability of fixation of a beneficial mutation depends on the mutation's net fitness effect across both sexes, its frequency, and the degree of dominance (Haldane 1927; Rice 1984; Orr 2003). Compared to a neutral mutation and assuming an equal variance in reproductive success of each sex, fixation of mutations with sexually antagonistic fitness effects is more likely on autosomes and on sex chromosome when the net fitness effect across the sexes is positive (Ellegren and Parsch 2007).

Rice (1984) developed expectations for the distribution of SA genes in the genome and concluded that sex chromosomes, especially the X or Z chromosome, should be hot spots for fixation of mutations with sexually antagonistic fitness effects because there is a broader range of scenarios where SA mutations have a net positive selection coefficient as compared to autosomal or pseudoautosomal genomic regions. These cases includes SA mutations that are 1) dominant and beneficial to the homogametic sex, or 2) recessive and beneficial to the heterogametic sex (Rice 1984; Rice and Chippindale 2001; Ellegren and Parsch 2007).

For a recessive, SA mutation that is beneficial to the heterogametic sex, fixation on autosomes is favoured by natural selection only when the benefit to one sex out weights the costs to the other. In contrast, when this same mutation occurs on a X or Z chromosome, natural selection will favour the mutation while it is rare enough to usually be heterozygous with the wildtype allele in the homogametic sex (Rice 1984; Ellegren and Parsch 2007). For SA mutations that benefit the heterogametic sex but harm the homogametic sex, natural selection favours fixation if it resides on the W or Y chromosome (Rice 1984). Consistent with the expectation that mutation with sexually antagonistic fitness effects are more likely to persist or fix on sex chromosome, artificial selection for two dominant eye color alleles in female Drosophila melanogaster was associated with a decrease in the male-female sex ratio, presumably due to an increased frequency of a male detriment allele (Rice 1992). Similarly, studies of the X chromosomes of D. melanogaster suggest that this chromosome contributed 97% of negative genetic covariance for adult fitness, which indicates that it harbours a large fraction of genes with SA fitness effects (Gibson et al. 2002). Additionally, genomic data from 10 human populations found that the X chromosome has higher proportion of single nucleotide polymorphism (SNP) with significantly different frequencies between sexes, which is a signature of SA selection on viability as compared to the autosomes (Lucotte et al. 2016).

1.3 Resolution of sexual antagonism

Sex chromosomes emerge from an ancestral pair of autosomal chromosomes when they acquire a genetic variation that drives sexual differentiation (Bull et al. 1983; Charlesworth et al. 2005). When a sex determination genes is in close genomic proximity to other genes that have sex-specific benefits or costs (i.e., SA fitness effects), natural selection is expected to favour the evolution of recombination suppression between them (Bull et al. 1983; Charlesworth et al. 2005). Recombination suppression between sex-chromosomes (X and Y, or Z and W) has the potential to cause mutation with SA fitness effects to be inherited in a sex-specific or sex-biased manner and thus resolved or reduced the selection pressure of the genomic conflict. Recombination suppression may arise via several non-exclusive mechanisms such as inversions (Charlesworth et al. 2005), mutation that increase divergence (insertion deletion events, point mutations, transposable elements) or mutations that affect motifs associated with recombination (Charlesworth et al. 2005). While in some species such as the African clawed frog, *Xenopus laevis* (*X. laevis*), the non-recombining region surrounding the sex-determining locus is small, in other species, such as the Kenyan clawed frog, *Xenopus borealis* (*X. borealis*), the non-recombining region is large, spanning much of the sex chromosomes (43%) (Furman and Evans 2018). Resolution of genomic conflict associated with mutations with SA fitness effects by sex-linkage could be an important factor that drives the expansion of nonrecombining region of sex chromosome (Rice 1987; Bergero and Charlesworth 2009).

Alternatively, genomic conflict associated with mutations with SA fitness effect may also be resolved via the evolution of sex-biased expression, wherein expression of the SA alleles is higher in the sex that it benefits and ultimately allow the evolution of sexual dimorphism in phenotype that influence fitness (Connallon and Knowles 2005; Ellegren and Parsch 2007; Otto et al. 2011). Sex-biased gene expression is thought to be the primary mechanism to resolve conflict associated with SA alleles on autosomes (Mank and Ellegren 2009) and this mechanism may also be important on sex chromosomes. Sex-biased expression could eliminate negative effects of SA alleles in the sex that it harms while preserving the positive effects to the sex that it benefits. On sex chromosomes, the origin of sex-biased expression of SA alleles could also reduce the selection pressure favouring expansion of recombination suppression. Several studies have used sex-biased expression as a proxy for the complete or partial resolution of SA associated genomic conflict (Connallon and Knowles 2005; Mank et al. 2008; Mank 2009; Mank and Ellegren 2009). In principle, linkage to sex chromosomes combined with sex-biased expression could resolve genomic conflict more quickly than sex-biased expression alone (which can resolve genomic conflict on autosomes or pseudoautosomal regions). For this reason, one might expect a higher proportion of genes with sex biased expression patterns on sex chromosomes as compared to autosomes and pseudoautosomal regions (Rice 1984). Empirical studies provide variable support for this expectation. An enrichment of genes with sex-biased expression, especially female-biased expression, was observed in the sex chromosomes of Drosophila and stickleback fish compared to their autosomes (Parisi et al. 2004; Mank and Ellegren 2009; Leder et al. 2010; Vicoso and Bachtrog 2015). However,

a recent study of the frog *Rana temporaria* found no enrichment of genes with se-biased expression on the sex chromosomes (Ma et al. 2018b). However, sex determination in this species is possible "leaky" in the sense that both sex chromosomes may segregate (to differing degrees) in both sexes. In emus, *Dromaius novaehollandiae*, a higher proportion of genes with sex-biased expression were detected on the sex chromosomes, including the pseudoautosomal (PAR) regions than the autosomes (Vicoso et al. 2013). This study also found a larger number of genes with male-biased than female-biased expression after gonadal differentiation began (Vicoso et al. 2013). However, theoretical work on the evolutionary dynamic of genes in the PAR indicated that SA mutations should evolve similarly for genes in the PAR region and autosomes unless there's strong selection or close linkage to the sex-determining region (Charlesworth et al. 2014). Depending on whether sex-biased regulatory divergence evolves ancestrally before recombination suppression or after, the age of sex chromosome might also influence the degree to which sex chromosomes resolve sexual conflict via sex-biased gene expression. The homomorphic sex chromosomes of *R. temporaria* may, for example, be relatively young compared to the heteromorphic sex chromosome of *Drosophila* fruit flies and stickleback.

1.4 Xenopus sex chromosomes

The sex chromosome of the Kenyan clawed frog, *X. borealis*, provide a compelling model with which to explore enrichment and evolution of sex-biased gene on a new sex chromosome. *Xenopus borealis* has homomorphic sex chromosomes (Tymowska 1991). The sex chromosomes of this species have a large nonrecombining region that spans almost half of the sex chromosomes (chromosome 8L; hereafter chr8L); the other half of chr8L undergoes recombination and thus a PAR (Furman and Evans 2018). This provide an opportunity to test whether a higher proportion of genes in a relatively large non-recombining portion of the sex chromosomes have sex-biased expression as compared to PAR and the autosomes. Because *X. borealis* is an allotetraploid (a tetraploid formed via fusion of two closely related diploid genomes), we can also compare the sex chromosomes of this species (chr8L) to its homeologous chromosome (chromosome 8S; hereafter chr8S) to gain further insights into the ancestral expression before chr8L become a sex chromosome. Additionally, a closely related allotetraploid species *X. laeivs* shares a most recent common ancestor with *X. borealis* about 18-34 Ma and also has homomorphic but non-homologous sex chromosomes (chromosome 2L; hereafter chr2L; Yoshimoto et al. 2008; Furman and Evans 2016). The non-sex-linked orthologs of *X. laeivs* chr8L thus provides another informative contrast to the sex-linked genes in *X. borealis* chr8L with which to better understand regulatory changes associated with recombination suppression in a new sex chromosome.

1.5 Goals

In this thesis, I tested whether the non-recombining portion of the sex chromosome of *X. borealis* has a higher proportion of genes with sex-biased expression as compared to other

portions of the genome that are not sex-linked (the PAR region and other autosomes, including autosomes that are homeologous to the sex chromosome). I considered whole transcriptome data from adult liver and gonad tissues and also mesenephros/gonad tissues from two tadpoles stages (46 and 48). A higher proportion of genes with sex-biased expression patterns in the sex-linked portion of the sex chromosomes were observed compared to other genomic regions. Comparison to homeologous and orthologous data (from *X. laeivs*) suggests sex biased expression arose recently – after speciation of *X. borealis* and *X. laeivs* – and possibly contemporaneously with the origin of the young sex chromosome of *X. borealis*.

2 Materials and Methods

2.1 RNAseq extraction and sequencing

We sequenced complete transcriptomes from multiple tissues from two species of African clawed frog – *X. borealis* and *X. laeivs*. Both were obtained directly or derived from individual that were initially purchased from XenopusExpress (Brooksville, Florida, USA, http://www.xenopus.com). Frogs were raised and kept in a recirculating water system with a 12/12 hour day/night cycles. Prior to tissue collection, animals were euthanized by transdermal overdose of tricaine methanesulfonate (hereafter MS222) (Sigma, Oakville). Liver and gonad tissues were sampled from adult individuals of a *X. borealis* family (father, mother, three sons, three daughters). Liver tissues were sampled from adult individuals from a *X. laeivs* family (3 sisters and 3 brothers). The phenotypic sex was determined by dissection. Tissues were homogenized using a hand-held homogenizer and RNA was extracted using the Qiagen RNeasy tissue extraction kit (Qiagen, Toronto) following the manufacturer's protocol. The quality of the RNA extractions was assessed using a Nanodrop spectrophotometer (Thermofisher Scientific, Burlington). Library preparation and transcriptome sequencing were performed at the Centre for Applied Genomics (Toronto, ON, Canada) using an Illumina HiSeq 2500 machine, and 150 base pair reads. Sequencing was performed over three lanes with each sample multiplexed across all three.

The mesenephros/gonad tissue complex were sampled from *X. borealis* tadpoles stage 46 and 48 from the same family. These tadpoles were lab bred and raised in the same condition as adults. Tadpoles were feed daily with powdered plankton food, which is mainly a mixture of Spirulina and krill. Sibling from the same family were euthanized via transdermal overdose of MS222, and the mesenephros/gonad tissue complex were dissected and stored in RNAlater (Qiagen). Developmental stages were identified based on external morphology as visualized with a microscope. The genetic sex of each tadpole was ascertained based on sex-linked genotypes at two genes for each individual: SOX3 and SF1, which are heterozygous in females but not males in our lab strain (Furman and Evans 2016). We sequenced transcriptome from four females and four males at stage 46, and eight females and five males at stage 48. RNA extraction, library preparation, and

sequencing were performed using the same method and facility as the adult tissues and multiplexed over one Illumina lane.

2.2 *De novo* Transcriptome

De novo transcriptome assembly

Sequence reads were trimmed to remove adapter sequences, low quality reads, and low quality bases using Trimmomatic 0.36 (Bolger et al. 2014) with following parameters: TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 MAXINFO:30:0.7 MINLEN:36. The quality of trimmed reads was assessed with FastQC v10.5 before and after trimming (Andrews 2010). We performed three separate *de novo* transcriptome assemblies using trimmed reads from *X. borealis* adults (liver and gonad tissues combined), *X. borealis* tadpoles (stage46 and 48 combined), and *X. laeivs* adults (liver tissue) respectively.*De novo* transcriptome assembly was performed using Trinity 2.8.2 (Grabherr et al. 2011; Haas et al. 2013) with minimum k-mer coverage of 2. Alternatively, we could have assembled one single *X. borealis* transcriptome with all its reads. With the latter approach, we would end up with a overly large transcriptome due to polymorphisms presents in different tissue types and developmental stages. Additionally, the *X. borealis* adults and tadpoles had different parents; hence, building separate transcriptome for adult and tadpoles would include little family dependent polymorphisms.

We additionally downloaded RNAseq data from *Xenopus tropicalis*(*X. tropicalis*) liver tissue from two males (SRR5412273 and SRR5412274) and two females (SRR5412275 and SRR5412276). Transcriptome assembly was performed for these *X. tropicalis* data using the same method as described above.

Transcriptome mapping

In order to assess genomic location of transcripts from each tetraploid species (*X. borealis* and *X. laeivs*), we used the *X. laeivs* genome assembly V9.2 (Session et al. 2016) as a reference assuming little lost in genes synteny between the two species. Assembled transcripts from *X. borealis* and *X. laeivs* were mapped to this genome using the splice-aware aligner GMAP (Wu and Watanabe 2005) with parameters: -A -B 5 -t 20 -f samse. Additionally, the parameter –cross-species was used when mapping *X. borealis* to *X. laeivs* genome to accommodate divergence between the assembled transcriptome and reference genome. The *X. laeivs* genome V9.2 and annotation gff file were obtained from Xenbase (Karimi et al. 2018).

2.3 Differential expression analysis

Binning of transcript counts based on genomic location

Transcript counts were quantified for each samples using Kallisto v.0.43.0 following the methods of (Bray et al. 2016) with default parameters for indexing (index –kmer-size 31) and quantification (quant -b 0 -t 1).

One goal of this study is to test whether the sex-linked region of the genome have a higher proportion of sex-biased genes than other regions of the genome. Challenges to this analysis are that 1) some genes might be present in *X. borealis* but not *X. laeivs*, or vice versa, 2) some genes might be presented in one or both species but are not annotated, 3) during transcriptome assembly, some genes might have more assembled isoforms than other genes either due to true biological reasons such as alternative splicing, or alternatively due to sequencing or assembly error or intraindividual polymorphism. In order to cope with these issues, we elected to combine expression counts of transcripts into a single locus based on their genomic location relative to the *X. laeivs* genome sequences. A disadvantage of combining expression values of transcripts is that we could lose the ability to identify sex-biased expression of splice variants. However, we felt that the benefit gained from combining information from partially expressed transcripts outweighed this limitation. Thus, expression counts were summed for overlapping transcripts that mapped to annotated genes or overlapped with annotated genes for at least 200bp on the *X. laevis* genome V9.2 gff file assuming an exon should be at least 200bp in length. Additionally, for genes that did not map to annotated *X. laeivs* annotated genes, we collapsed transcripts that mapped to overlapping region into one transcript bin. Hereafter, we refer to these transcript bins as genes, and the names of those genes are formatted as following: chr:start-end, where chr is the name of the chromosomes or scaffold, the start and end are the upstream and downstream coordinates in the *X. laeivs* reference genome, respectively.

Differential expression analysis

Differential expression analysis between sexes was performed separately for each of the four *X. borealis* tissue types/developmental stages. Two R packages, edgeR (Robinson et al. 2010) and DESeq2 (Love et al. 2014) were used because previous studies identified advantages and disadvantages and cons of each approach (Conesa et al. 2016; Schurch et al. 2016). Collapsed raw read counts from Kallisto (Bray et al. 2016) were used for each analysis. We discarded genes that had total summed read count across samples less than the total number of samples, meaning that each sample was required to have an average of at least 1 raw read. The EdgeR package (version 3.4) was used with vanilla pipeline (e.g. calcNormFactors, estimateCommonDisp, estimateTagwiseDisp, exactTest for comparison between males and females) following the edgeR vignette (Robinson et al. 2010). We defined significantly sex-biased genes after Benjamini-Hochberg correction for a false discovery rate (FDR) of 0.05. Differential expression using DESeq2 package was

performed using the default function (e.g., contrast between male and female) following DESeq2 vignette (Love et al. 2014). Following these DESeq2 recommendation, shrinkage was used with adaptive t prior shrinkage estimator from the package "apelm" (Zhu et al. 2019). This option reduces the mean squared error of expression levels of each gene relative to the classical estimator, especially for genes with low expression levels (Love et al. 2014; Zhu et al. 2019). We defined significantly sex-biased genes based on a p-value (adjusted with Benjamini-Hochberg correction) cutoff of 0.05 from the DESeq2 output. Differential expression results from DESeq2 were summarized in the Results section and were used for downstream analyses unless otherwise stated.

2.4 Identification of orthologs and paralogs

In order to examine the evolutionary rate of expression divergence in a phylogenetic context (section 2.6), we needed to identify orthologous and paralogous genes in the *X. borealis*, *X. laeivs* and *X. tropicalis* transcriptomes.

To this end, a transcriptome from *X. tropicalis* was assembled *de novo* using the method detailed above, and then mapped to *X. laeivs* genome assembly V9.2 using GMAP with the setting detailed above (with –cross-species). Transcripts from different species (*X. borealis* and X. troplicalis) that mapped to or were annotated in *X. laeivs* in the same genomic location (i.e. the same chromosome and genomic coordinates with a minimum overlap of 200bp) were considered to be orthologous. If a transcript mapped to multiple locations on the same chromosome, genomic location of this transcript was assumed to be the one with the highest mapping quality and the most upstream genomic coordinates. In cases in which multiple alignments had the same mapping quality, the transcript was be mapped to the genome again with BLASTN (e-value cutoff 1e-10) and the best match from this approach was defined as orthologous. In the cases where mapping from GMAP and BLASTN were inconsistent, we assumed no ortholog could be identified. In order to identify paralogs in the tetraploid species *X. borealis* and *X. laeivs*, using the X.laevis genome annotation gff file from Xenbase (Karimi et al. 2018), we generated an *X. laeivs* exome that included only transcribed sequences. We then extracted exonic sequences of annotated genes. Concatenated exons from each *X. laeivs* gene were then used as a query to interrogate all of the concatenated exon sequences using BLASTN(e-value cutoff 1e-10). We expected that each sequence would have itself as the top hit and its homeologous gene (if present) as the second best hit. If a gene did not have itself as the top hit, the gene was excluded from further paralog identification. Genes in which the second-best match did not map to the homeologous chromosomes or a scaffold (which could potentially be on a homeologous chromosome) were assumed to be singletons in which the homeolog was lost or not highly expressed in the transcriptome we studied. We additionally required that paralogous pairs be each others reciprocal best BLAST hit, i.e., both paralogs were required to be each other's second-best match. A caveat to this approach is that we only allow one paralogous match per gene and thus might miss some of paralogous pairs where a gene might have multiple paralogous matches (e.g., homeologs in which a segmental duplication of one homeolog occurred).

2.5 W chromosome degeneration and lack of dosage compensation as a possible cause of sex-biased expression

Under a scenario in which the portion of the W and Z chromosome do not recombine and diverge, it is possible that genes are lost in the non-recombining portion of the W. If this is the case, male-biased expression could reflect a difference in Z chromosomes allele numbers between the sexes (two in ZZ males, one in WZ females) rather than a difference in gene regulation. This could potentially inflate the proportion of sex-biased genes in the regions. To explore this possibility, we used whole genome sequencing data from (Furman and Evans 2018) from one male and one female individual that was mapped to *X. laeivs* reference genome version 9.2. We calculated the per base coverage of genes with male-biased expression patterns using Bedtool (Quinlan and Hall 2010). The average per base coverages of these male-biased gene were normalized by dividing by the genome-wide per base coverage to account for the difference in sequencing depth between the male and female individual. This approach allowed us to evaluate the possibility of differential expression due to gene deletion on the W but not due to pseudogenization without deletion (e.g. by a nonsense mutation or a change in the transcription factor binding sites on the W).

Under the assumption the ancestral gene expression of paralogs of sex-linked genes in the L subgenome can be approximated by expression of their paralogs in the S subgenome, we would expect degeneration of the W chromosome to lead to a lower L/S expression ration in females than in males. We used a linear model to test this possibility – specifically, whether the L/S ratio was affected by the interaction of predictor variables of sexes (female or male) and subgenomes (L or S).

$$L/S\,ratio \sim sex * subgenome$$

The L/S ratios were calculated using counts in TPM quantified by Kallisto (Bray et al. 2016). L/S ratios for sex-linked genes in each sampled *X. borealis* individual were considered excluding genes that had no expression in either the L or S paralogs.

2.6 Comparing the magnitudes of expression divergence in sex-linked and non-sex-linked genomic regions

We used a linear model to test whether a response variable, the absolute value of the log2 transformed male:female expression ratio, which represent magnitude of expression divergence between males and females, was affected by the interaction between the predictor variables of sex-linkage (yes or no) and subgenomes (L or S):

$$abs(log2FC) \sim sex\text{-}linkage * subgenome$$

We performed this analysis using ratios estimated from DESeq2. We had a one-sided expectation that the absolute value to this ratio would be higher on the sex-linked portion of chromosome 8L of *X. borealis*, which would be consistent with more rapid sex-biased expression divergence on the non-recombining portion of this sex chromosome. A significant p-value of the interaction term coefficient would indicated that the sex-linked portion of the L subgenome has a significant effect on the absolute value of the male:female expression ratio. With the R package "emmeans" (Estimated Marginal Means) (Lenth 2019) we can visualize the predicted means for genes in each categories (chr8L sex-linked region, chr8L non-sex-linked region, chr8S sex-linked region, chr8S non-sex-linked region region). Preliminary analyses indicated that the absolute values of the male:female expression ratios were not normally distributed based on the residual QQ-plots. To accommodate this, we fitted a linear model with randomized residual permutation procedure using lm.rrpp from the R package "RRPP" using 1000 permutations (Collyer and Adams 2018).

Another concern is that there could be co-variance in expression levels among genes in similar genomic locations due to co-regulation. This would violate the assumption of our linear model that the absolute value of the male:female expression ratio of each genes is independent of the value of other genes. To assess this possibility, we permutated the sexes of individuals within each tissue type and developmental stage and then performed differential expression analysis on the permuted data. For example, for adult liver samples, there were 4 males and 4 females, and there are thus 70 possible permutation of these sex assignments to each of the 8 transcriptomes (=8!/(4!*4!)).

Another possible concern with our analysis is that genes with significant sex-biased expression patterns might have more extreme magnitude of expression divergence as compared to genes without statistically significant bias, and that this could inflate the mean magnitude of expression divergence. To explore this issue, we fitted the linear model to sex-biased genes as defined by four corrected p-value cut-off levels (in order of decreasing stringency): 0.05, 0.1, 0.15 and 1. At a p-value cut-off of 1.00, all genes were included in the analysis.

2.7 The evolution of sex-biased gene expression in a comparative context

The analyses of differential expression detailed above examines expression patterns within *X. borealis*, including comparisons between the non-recombining and recombining portions of the sex chromosome, and between the non-recombining portion of the sex chromosome and the homeologous portion of the autosome. To further explore how these patterns vary among species and changed over time, we evaluate the evolution of sex-biased expression (represented by the log-transformed M/F expression ratio) in a phylogenetic context under the assumption that this ratio evolves under a Brownian motion model (Butler and King 2004; Revell 2012). A character evolving under Brownian motion accumulates variance at some rate (). In order to test the prediction that

the genes on the sex-linked region of *X. borealis* chr8L had a faster rate of evolution, we compared the fit of a model with one rate parameter over the entire phylogeny to a model with an independent rate parameter on the branch leading to *X. borealis* chromosome 8L (Fig. 1). Thus the one-rate model is a special case of the two rate model. We also fitted the same data set into a three-rate model and a four-rate model. In the three-rate model, we consider the one-rate parameter for S subgenomes branches of *X. borealis* and *X. laeivs*, another for the L subgenomes branches of *X. borealis* and *X. laeivs*, and a third rate for the diploid branches that subtend *X. tropicalis* combined with the diploid/tetraploid branches that subtend the L and S subgenomes of *X. borealis* and *X. laeivs*. The three-rate model was a special case of the four-rate model, which had separate rate parameters for the L subgenomes branches into the *X. borealis* 8L branch and *X. laeivs* 8L branch. We included in this analysis sex-linked genes and their orthologs and paralogs from one or both homeologs of *X. laeivs* and *X. borealis* (chr8L and chr8S) and from the orthologous gene from *X. tropicalis* (chr8). A phylogeny for this analysis was estimated from concatenated sequences from (Furman and Evans 2016). The log-likelihood of the data given this phylogeny was calculated using brownian.lite from R package "phytools" (Revell 2012). Because some homeologs and orthologs were not identified for some genes, we only considered genes that had data from at least four out of the five tips. The five possible case are 1) no missing data points; 2) missing *X. borealis* chr8S; 3) missing *X. laeivs* chr8L; 4) missing X.laevis chr8S; 5) missing *X. tropicalis* chr8 (Fig. 2).

We summed the log likelihoods across individual genes for each model and compared these sums between models using Akaiki Information Criterion (AICc) with small sample correction and pair-wise likelihood ratio tests with degrees of freedom equal to the number of genes (O'meara et al. 2006; Revell 2012). The p-value of the likelihood ratio test was computed against the X^2 distribution. Additionally, the average Brownian motion rate parameters(s) across genes was calculated with the one-sided expectation that the average rate for the sex-linked region of *X. borealis* chr8L from the two-parameter model would be the highest. To estimate a confidence interval for AICc, we performed permutation test by sampling the genes in the sex-linked region with replacement and repeated the permutation 1000 times. For each permutation, we fitted the log-transformed M/F expression ratio of genes from each sample set into all four Brownian motion models and obtain AICc for summed log-likelihood. Then we compared the observed AICc with the AICc from permutation.

Directionality of genes expressions regulation

To examine the directionality of expression regulation (e.g. whether a sex-biased gene is a consequence of upregulation in females or downregulation in males), we fitted a Brownian motion model on expression data from female and male separately. We normalized expression levels between samples with TMM (trimmed mean) using function from the Bioconductor package EdgeR and between species using a scale based normalization (SCBN) from the R package SCBN (Zhou et al. 2019) to allow cross species comparison.

Then we compared the rates of evolution for male and female expression in a two-rate model (with one rate on the lineage for chr8L in *X. borealis* and another rate for the rest of the phylogeny). Because multiple scenarios are possible for a given gene, we provide the interpretation and summary of the result in the Results section.

2.8 Documentation of pipeline

Scripts and pipelines for this study are available on github (github/evansb/IdeasforXue).

3 RESULTS

3.1 Transcriptome assembly and genomic localization

The transcriptomes assemblies had 972,785, 523,313, 391,563, and 207,184 transcripts, respectively, for *X. borealis* adults tissues (liver and gonad combind assembly), *X. borealis* tadpoles tissues (gonad/mesenephros from stage 46 and 48), *X. laeivs* adults livers, and *X. tropicalis* adults livers. These transcripts include assembly errors and biological complexity stemming from splice variants and heterozygosity. For these reasons, the number of assembled transcripts far exceeds the expected number of genes in these genomes (<100,000 for the tetraploids and half this for the diploid). To simplify the complexity of these transcriptomes and pin them down to specific genomic regions, we combined expression levels of transcripts that mapped to the same genomic region of a reference genome assembly of *X. laeivs* (for *X. borealis* and *X. laeivs* transcriptomes) or *X. tropicalis* (for *X. tropicalis* transcriptome). After collapsing, the *X. borealis* adult liver and gonad transcriptome consisted of 128,576 genes, in which 51,041 are orthologous to annotated X.laevis genes and 77,535 are orthologous to non-annotated loci. The *X. borealis* tadpole gonad/mesenephros transcriptome, the *X. laeivs* adult liver transcriptome, and *X. tropicalis* adult liver transcriptome consisted of 72, 531, 32,397, and 15,109 genes respectively. The lower number of genes in the transcriptomes of *X. laeivs* and *X. tropicalis* compared to *X. borealis* is expected because only one tissue type was analyzed from these species, whereas two tissue types/developmental stages combinations were assembled in each *X. borealis* transcriptome.

3.2 The genomic landscape of genes with sex-biased expression in *X. borealis*

We performed differential expression analysis with DEseq2 package for each tissue type (Love et al. 2014), and defined sex-biased gene as having false discovery rate (FDR) or adjusted p-value ≤ 0.05. Base on a previous study, the sex-linked region of the sex chromosome of *X. borealis* is chromosome 8L (chr8L) from positions 4,605,306 to 56,690,925 base pairs (Furman and Evans 2018). We divided the sex chromosome chr8L

into the sex-linked (from 4,605,306bp - 56,690,925bp) region and the non-sex-linked (all other regions of the sex chromosome) region; the latter is the PAR, which is similar to the autosomes and recombines during meiosis in males and females. For comparative purposes, we also demarcated the regions of chromosome chr8S that are homeologous to the sex-linked and non-sex-linked portions of chr8L. Both of these regions of chr8S are autosomal. Here we presented result of analyses done with DESeq2 package(Love et al. 2014); however, analysis of these data with edgeR package (Robinson et al. 2010) produced similar results ((Fig. S1 - S11, Table. S1 - S4).

We identified nine, 14, 20, and 151 sex-biased genes in chr8L sex-linked region of *X. borealis* tadpole gonad stage 46, tadpole gonad stage 48, adult liver, and adult gonad respectively (Fig. 7 - 10, Table. 1 - 4). Eight, 11, seven, and 10 of these had higher expression in females, respectively, and the others had higher expression in males (Fig. 7 - 10, Table. 1 - 4). In the non-sex-linked region of chr8L, we identified two, five, eight, and 184 sex-biased genes in *X. borealis* tadpole gonad stage 46, tadpole gonad stage 48, adult liver, and adult gonad, respectively. Of these, one, four, two, and 15 were female-biased respectively (Fig. 7 - 10, Table. 1 - 4). In the sex-linked region of chr8S, we identified two, four, five, and 108 sex-biased genes in *X. borealis* tadpole gonad stage 46, tadpole gonad stage 48, adult liver, and adult gonad, respectively. Two, four, zero, and eight of these, respectively were female-biased (Fig. 7 - 10, Table. 1 - 4).

In each of the four *X. borealis* tissues we examined (liver, tadpole mesonephros/gonad from stages 46 and 48, adult gonad) the sex-linked portion of chr8L had a significantly higher proportion of genes with significantly sex-biased expression than the non-sex-linked portion of this chromosome ($p \leq 0.001$ for all tests, binomial test, Fig. 3 - 6). This was also the case when the sex-linked portion of chr8L was compared to the rest of the genome (including the PAR of chr8L) and also when it compared to its homeologous portion of chromosome 8S for adult liver, adult gonad tissue, and gonad/mesenephros for both tadpole stages ($p \leq 0.001$ for all tests, binomial test, Fig. 3 - 6). The proportions of genes with sex-biased expression on the sex-linked region of chr8L had a large 95% confidence interval as a consequence of the relatively small number of genes in this region.

Under a scenario where genes on the W chromosome were lost due to degeneration associated with recombination suppression, significant male-biased expression could be a consequence of difference in allelic numbers (two in ZZ males and one in ZW females). To explore this possibility, we examined the per base genomic coverages of genes with significant male-biased expression in the sex-linked region. We found that 4% male-biased genes (7 genes) from all tissue types/developmental stages had 1.5X or greater per base coverages in WGS data from a male compared to data from a female (1/7 in liver, 6/7 in adult gonad, 0 in either tadpole stages). None of these had zero coverage in WGS data from the female. This suggests a modest role for gene deletion on the proportion of genes with male-biased in the non-recombining portion of the W chromosome.

Another possible expectation is that W chromosome degeneration could lead to down-regulation of gene expression in WZ females compared to ZZ males, and that this could contributed to the proportion of male-biased genes in this genomic region. If this were

the case, we expected that genes in the sex-linked region to have a lower L/S expression ratio in females than in males under the assumption that expression in the S subgenome is an approximation of the ancestral expression level. However, the L/S expression ratio of sex-linked genes and their homeologs was no significantly different between males and females for any tissue type or developmental stage (p= 0.8649, 0.2859, 0.3883, 0.4551, for adult liver, adult gonad, gonad of tadpoles/mesenephros at stage 46 and stage 48 respectively; Fig. 14 - 17). In the non-sex-linked regions, L/S expression ratios also were not significantly different between males and females (p= 0.9150, 0.4702, 0.8351, 0.0902 for adult liver, adult gonad, gonad of tadpoles/ mesenephros at stage 46 and stage 48 respectively; Fig. 14 - 17). When comparing L/S ratio within each sex between sex-linked and non-sex-linked genes on chr8L, it was significantly higher in the sex-linked genes in female adult gonad ($p = 0.0222$) but is not significant in male adult gonad($p = 0.0093$) or in the other tissues/developmental stages (p= 0.1361, 0.6981, 0.3395, in females and p= 0.2266, 0.1652, 0.2412, in males for adult liver, gonad of tadpoles/ mesenephros at stage 46 and stage 48 respectively; Fig. 14 - 17). Taken together, these results suggested a modest contribution of W chromosome degeneration (gene deletion and gene neofunctionalization) to the observed elevated proportions of genes with sex-biased on the non-recombining portion of the W.

3.3 Expression divergence between the sexes of genes in the sex-linked region and to non-sex-linked region

When considering all genes from *X. borealis* chr8L and chr8S (differential expression result without any filtering criteria, meaning FDR ≤ 1), the male/female expression ratios that were estimated from the linear model were larger in the chr8L sex-linked region compared to chr8L non-sex-linked region for all tissues; the differences were significant in all tissues except stage 46 gonad/mesenephros ($p = 0.0007$, 0.00002, 0.304, 0.0062 for adult liver, gonad of tadpoles/ mesenephros at stage 46 and stage 48 respectively; Fig. 18 - 21). The coefficient for the interaction between sex-linkage and subgenome was significant for adult gonad, but not for adult liver, or tadpole stage 46, or stage 48 gonad/mesenephros (Table. 5). A permutation test indicates that the observed interaction coefficient obtained form the model was unlikely to arise by chance ($p \leq 0.1$, Table. 6).

When we considered significantly sex biased genes by more stringent criteria (differential expression result with filtering criteria FDR-corrected p-value ≤ 0.05), we observed significant higher predicted mean expression ratio in adult gonad ($p = 0.0025$) and slightly higher though not significant in tadpole stage 48 gonad/mesenephros ($p = 0.815$), but not in the adult liver nor in tadpole stage 46 gonad/mesenephros (Fig. 22 - 25). The p-value of the interaction term was not significant for any of these tissues/developmental stages (Table. 5), and the permutation test indicated that the observed coefficient for the interaction between sex-linkage and subgenome did not depart from random expectations (Table. 6). These results suggest that genes with similar magnitude of expression divergences in the sex-linked region of sex-chromosomes and in the autosomes contributed to

the statistical signature of an enrichment of genes with sex-biased expression patterns on the sex-linked portion of chr8L.

3.4 Timing of sex-biased gene expression

To further explore whether these patterns of enriched sex-biased expression emerged before or after chr8L became a sex chromosome in *X. borealis*, we examined the distribution of sex-biased genes on chr8L in a closely related species (*X. laeivs*) where chr8L is not a sex chromosome. We found that the region of chr8L in *X. laeivs* that is orthologous to the sex-linked region of the *X. borealis* sex chromosomes was not enriched for sex-biased genes when compared to genes in regions of *X. laeivs* chr8L that are orthologous to the PAR of the sex chromosome of *X. borealis* (p= 0.67, binomial test; Fig. 26). There was also no significant difference when the proportion of *X. laeivs* genes in this genomic region with sex-biased expression genes were compared to proportion in the rest of *X. laeivs* genome (including region orthologous to the PAR of *X. borealis* chr8L; p=0.21, binomial test; Fig. 26). Fitting of X.laevis chr8L and chr8S expression male/female ratio with a linear model (with interaction terms between sex-linkage and subgenome) showed that the magnitude of expression divergence of the genes in the region orthologous to the *X. borealis* sex-linked region was not higher than that of genes in the region orthologous to *X. borealis* PAR, and the effect of the interaction terms were not significant for all level of filtering for sex-bias (p = 0.21, p = 0.082, p = 0.126, p = 0.242, for FDR corrections of 0.05, 0.1, 0.15, and 1; Fig. 27 - 28, Table. 7). These results suggested that patterns of enriched sex-biased expression in *X. borealis* sex-linked region emerged after chr8L became a sex chromosome in *X. borealis*.

3.5 The evolutionary rate of gene expression divergence between sexes in the chr8L sex-linked region

We explored how sex-linkage affects the evolutionary rate of gene expression divergence of *X. borealis* by fitting M/F expression ratio data to various Brownian motion models. This analysis was performed only on M/F expression ratio data from adult liver because we lacked orthologous data from *X. laeivs* and *X. tropicalis* for adult gonads and for tadpole mesonephros/gonads. After discarding genes with missing data (Method 2.6), a total of 143 genes were fitted. To explore model fitting across all loci, we compared the summed log-likelihoods of the one-parameter models and the two-parameters models. The summed of log likelihood of the one-parameter model and the two-parameters models were -1570.727 and -1458.383, respectively, and the more parameterized model was thus a significantly better fit to the data (likelihood ratio test, df =143, p-value < 0.001; AICc of 1-parameter > AICc of two-parameters; Table. 8 - 10). In the two-rate models, the average rate for the sex-linked branch (33.00) was higher that of non-sex-linked branches (10.14). We were concerned that the average rate might be inflated by outliners. However, with outliners removed, the average rate for the sex-linked branch (10.09) was still higher than that of non-sex-linked branches (5.96), indicating the evolution rate of

sex-biased expression divergent was faster for genes on the sex-linked region of the *X. borealis* sex chromosomes than for genes in other homeologous and orthologous genomic region (Table. 10). When we performed BM analysis for a subset of genes that were significantly sex-biased (FDR <= 0.05), only two of them had data on expression ratio for at least 4 branches. The 2-rate models for those two genes were not significantly better than the 1-rate model.

We also fitted the same full data set to models with three and four rates parameters. Pairwise comparison of all models' summed log-likehood and AICc indicated that the four-parameters models fit the data significantly better than other models (likelihood ratio test, df=143, p-value <0.001; AICc of 3-rate models > AICc of 4-rate models; Table. 8 - 9). The bootstrapping result indicated that the observed AICc values were found within 95% confidence interval (Table. 9). In three-rate models, the average rate of L subgenome branches (21.66) was higher than the S subgenome branches (13.298) and the background (0.26) (Table. 10; with outliners removed, the mean rate was 11.84, 7.03, and 0.17 for L branches, S branches, and background respectively). In the four-parameters models, the average rate of *X. laeivs* 8L branches (25.61) was higher than the *X. borealis* 8L branches (24.5), S subgenome branches (14.88), and the background (0.14) (Table. 10; with outliners removed, the mean rates were 8.88, 7.83, 6.85, and 0.099 for *X. borealis* 8L, S subgenome, *X. laeivs* 8L, and background respectively). The three-rate models and four-rate models suggested that the faster gene expression divergence of the *X. borealis* 8L branches in the two-rate models could be largely driven by the L subgenome evolution rather than sex-linkage.

If the evolution of L subgenome was the main driver of gene expression divergence, we would expect to observe the same patterns when we repeat the same four BM models for genes in the recombining region of *X. borealis*. Its two-rate models suggested the sex-linked branches was evolving faster than the non-sex-linked branches; however, the three-rate models suggested the S subgenome branches were evolve faster than the L subgenome branches and the background. Additionally, the four-rate models showed that *X. laeivs* 8L branches was evolving faster than S subgenome branches, *X. borealis* 8L branches, and background in the order of decreasing rates. Those patterns remained the same after outliners removed.

3.6 Directionality of gene expression divergence between sexes in the chr8L sex-linked region

We investigated the directionality of expression divergence by examining evolution rate (with BM model) of female expression and male expression of genes in the non-recombining region of *X. borealis* sex chromosome (the sex-linked branch) compared to their homeologs and orthologs (X.borealis 8S, *X. laeivs* 8L, and *X. laeivs* 8S; also are non-sex-linked branches). We found that the evolutionary rate of female expression and male expression were both faster in the sex-linked branch than the non-sex-linked branches (Table. 11 - 13). Additionally, the mean and median of the female expression and male expression are

higher for genes in the *X. borealis* non-recombining region compared to their homeologs and orthologs (Table. 13). These results suggested that gene expression in females and males were both upregulated compared to their orthologs and paralogs. We repeated the same analyses to female-biased genes and male-biased genes separately and found the same patterns - a faster evolution rates, higher means and medians for the sex-linked branches for females expressions data and for males expressions data compared to their homeologs and orthologs in the non-sex-linked branches.

4 DISCUSSION

Owing to their seminal roles in sexual differentiation and unique modes of inheritance, sex chromosomes offer unique insights into the drivers of genome evolution. Mutations with SA fitness effects – wherein the same allele confers benefits to one sex but detriments to the other – offer a prime example where distinctive evolutionary fates are expected depending on whether the mutation is present on one or the other sex chromosome, or on an autosome. Sex-linkage on sex chromosomes has the potential to resolve genomic conflicts associated with alleles with SA fitness effects under various circumstances (e.g. the degree of recessivity of the allele benefiting the homogametic sex, or the degree to which alleles are beneficial or harmful to each sex). Instead of or in addition to linkage to sex chromosome, genomic conflict associated with SA alleles may be resolved via the evolution of sex-biased expression, wherein alleles are expressed most highly in the phenotypic sex they benefit. Because these mechanisms are nonexclusive, in principle one might expect a higher proportion of genes with sex-biased expression on sex-linked portions of the sex chromosomes, as compared to non-sex-linked portions of the sex chromosomes (the PAR) and/or autosomes.

In this study, we explored this expectation in the newly evolved sex chromosomes of the Kenyan clawed frog *Xenopus borealis*. In this allotetraploid species, one of two homeologous chromosomes – chr8L – acquired a female-specific trigger for sex determination, thereby establishing a W chromosome and new sex chromosome with female (WZ) heterogamy and male (ZZ) homogamy (Furman and Evans 2016). Using RNAseq data from multiple tissue types and developmental stages, we recovered a modest enrichment of genes with sex-biased expression patterns in the sex-linked portion of these sex chromosomes. We found that the magnitude of expression divergence was higher for genes in this region as compared to genes in the non-sex-linked portion of the sex chromosomes of *X. borealis*.

4.1 The proportion of sex-linked genes with sex-biased expression increases through development

In *X. laeivs*, a gene called *dm-w* is a female-specific trigger for sex determination and its expression is detectable by tadpole stage 50, just prior to the development of morphologically differentiated gonads(Yoshimoto et al. 2008). Although the trigger for sex determination has not been identified in *X. borealis*, we expected gonadal differentiation to occur around the same stage as *X. laeivs* owing to their relatively close phylogenetic affinity. Consistent with this speculation, we found that the gonad/mesonephros transcriptome of the earlier development stage, tadpole stage 46 had fewer significantly sex-biased genes (0.06% at FDR corrected p-value of 0.05) as compared to the transcriptome from the same tissues in tadpole stage 48 (0.09% at FDR corrected p-value of 0.05). Perhaps not surprisingly, the transcriptome of the adult gonad tissue had a higher proportion of genes with significantly sex-biased expression patterns (8% at FDR corrected p-value of 0.05). This finding were consistent with expectation that the intensity of sexual selection should be similar between sexes in undifferentiated juveniles and higher in adults (Chippindale et al. 2001; Rice and Chippindale 2001; Gibson et al. 2002; Mank et al. 2010). There were a few genes that was significantly sex-biased across multiple stages of gonad development (Fig. 11 - 13). This was consistent with founding in birds (Mank et al. 2010); one explanation is that natural selection favoured sex-biased expression after sexual differentiation, and this regulatory pattern also affected expression prior to differentiation. We also found that the proportions of genes with significantly sex-biased expression was much smaller in liver tissue from *X. borealis* and *X. laeivs* (0.3% and 2% at FDR corrected p-value of 0.05) compared to in *X. borealis* gonad. The proportion of sex-biased genes in the transcriptome varies depends on species and tissue types; in gonads of the fruit fly Drosophila melanogaster, zebrafish, and mice, 91%, 38% and 71% of transcriptome were sex biased, respectively; in the brain tissue of zebrafish and Passerines bird, about 2% of the transcriptome were sex biased (Clayton et al. 2009; Assis et al. 2012; Naurin et al. 2012; Harrison et al. 2015; Grath and Parsch 2016; Ingleby et al. 2015). As compared to gonad, liver functions are presumably quite similar in each sex. Sex-biased gene expression in liver could therefore stem from pleiotropic functions of these genes in other tissue types (including gonads).

4.2 The non-recombining region of *X. borealis* sex chromosome is a hotspot for genes with sex-biased expression

In all four tissue types/developmental stage we studied, we detected an enrichment in the proportion of genes with sex-biased expression on the non-recombining region of the *X. borealis* sex chromosome chr8L compared to the rest of the genome (including the recombining portion of the sex chromosomes). To explore the possibility that the nonrecombining region had a higher proportion of genes with sex-biased expression before it become the sex chromosomes, we examined the proportions of sex-biased genes on the recombining region of the *X. borealis* sex chromosome, and on the portion of chr8S that

is homeologous to the non-recombining region of the *X. borealis* chr8L sex chromosome. In both comparisons, we found the non-recombining region of sex chromosome of *X. borealis* had a greater proportion of sex-biased genes compared to these other genomic regions. We performed the same three set of comparisons with orthologous regions in *X. laeivs* and found the proportions of sex-biased genes in the *X. laeivs* genomic regions that is orthologous to the non-recombining region of *X. borealis* chr8L was not significantly higher than other regions in all three comparisons. This indicates that the enrichment of sex-biased genes on the sex chromosome in *X. borealis* was not an ancestral condition and that instead this pattern evolved after divergence of the ancestor of *X. borealis* and *X. laeivs*, and around the time that this genomic region acquired a trigger for sex determination. Strong but incomplete linkage to a sex-determining locus can also favor the accumulation of sexual antagonistic genes (Rice 1987; Charlesworth et al. 2014) but this effect was not apparent in the PAR of *X. borealis* genome, at least in genes expressed in liver or gonad ($p<0.001$ for comparison between non-recombining portion of *X. borealis* 8L and the homeologous portion of *X. borealis* 8S), probably because all or almost all of the PAR is in linkage equilibrium with the non-recombining region of this chromosome.

These observations are consistent with observed enrichment of sex-biased genes on the X chromosome in D. melanogaster (Parisi et al. 2004), in mouse (Khil et al. 2004) and in threespine stickleback (Leder et al. 2010). However, a recent study of the European common frog, *(*Rana temporaria), did not detect an enrichment of genes with sex-biased expression on the sex chromosome (Ma et al. 2018b). The discrepancy between their finding and previous works could be due to the stringent filtering of the expression data (FDR <=0.05, |log2FC| > 1, average Log2CPM > 0, and CPM > 1 in at least half of the samples per sex) and transcriptional noise (Ma et al. 2018b). With less stringent filtering (Log2CPM > 2 or average Log2CPM > 0), the proportion of female-biased genes on the sex chromosome was higher than autosomes in adult gonad tissue and also in froglet (whole bodies with digestive tracts removed), in which gonadal tissue of males and females are morphologically distinct (Ma et al. 2018b). In the dioecious plant Salix viminalis, which has recently evolved ZW sex chromosomes, 44% (31 genes) of the genes in its small region of recombination suppression (3.5-8.8Mbp) have sex-biased expression (Darolti et al. 2018). Similarly, a microarray study in chicken found a greater proportion of Z-linked genes with sex biased expression pattern compared to autosomal genes (Mank et al. 2008a). They also found a higher proportion of Z-linked sex-biased genes to be male-biased and hypothesized that this could be due to a lack of dosage compensation.

Dosage compensation can affect genes expression between sexes (Vicoso and Charlesworth 2006; Vicoso and Charlesworth 2009). Several species with male heterogamy, such as Drosophila fruit flies, equilibrate expression of sex-linked gene through dosage compensation (Straub and Becker 2007; Gelbart and Kuroda 2009). However, several species with female heterogamy lack dosage compensation including chicken (Ellegren et al. 2007), birds (Itoh et al. 2007), and silkworms (Zha et al. 2009) and consequently have an over-representation of genes with male-biased expression on the Z chromosome (Storchová and Divina 2006; Arunkumar and Nagaraju 2009). Thus, it is possible that male-biased

genes on *X. borealis* nonrecombining region stem from loss of alleles on the W chromosome coupled with lack of dosage compensation. To investigate this possibility, we examined the genomic coverage of genes with male-biased expression in whole genome sequencing data of one male and one female. Across all four tissue types/developmental stages, we found a small fraction of genes with male-biased expression had a genomic coverage of 1.5X or higher in male compared to female. Additionally, using a linear model we compared the L/S expression ratio in both sexes, with the expectation that a lower L/S expression ratio in females would be consistent with allelic loss on the W chromosome, which is in the L subgenome. However, no significant difference between female L/S expression ratio and male L/S expression ratio was recovered for genes in the sex-linked region of the *X. borealis* sex chromosomes. These findings suggest that the effect of an allelic difference due to W chromosome degeneration coupled with a lack of dosage compensation on the numbers of male-biased genes on the nonrecombining region of the sex chromosome but its effect are quite modest. Moreover, when these genes are excluded from our analysis, we still see a significant enrichment of sex biased genes for all tissue types/developmental stages ($P<0.001$ for adult liver and adult gonad, and gonad/mesenephros for tadpole stages 46 and 48, respectively). Hence, in *X. borealis*, it is less likely that sex-biased genes on the non-recombining region were evolved as a result of lack of dosage compensation and is likely that it evolved as the resolution of sexual antagonistic conflict. The higher proportion of sex-biased genes observed in *X. borealis* is in line with Rice's prediction regarding sex chromosome is a hotspot for SA alleles. As discussed earlier, genomic conflict that arises from mutations with sexually antagonistic fitness effects can be resolved by linkage to a sex-determining locus and/or by the origin of a sex-biased expression pattern. The first mechanism, if widespread, could lead to an expansion of the region of recombination suppression, and over time this could lead to divergence between sex chromosome and the origin of sex chromosome heteromorphy. Alternatively, if genomic conflict is frequently resolved via the second mechanism, homomorphic sex chromosomes may persist for an extended period of evolutionary time. The sex chromosomes of *X. borealis* offer a combination of these scenarios - they are cytologically homomorphic and enriched for sex-biased genes compared to autosomes (which match the second mechanism) but nonetheless have a large region of recombination suppression (which match the first mechanism) (Furman and Evans 2018). One possible explanation for this is that these sex chromosomes evolved so recently that their consequences of recombination suppression (i.e. divergence) are not yet apparent. Another possible explanation is that recombination suppression could be driven by meiotic drive, or heterozygote advantage, and genetic drift rather than sexual antagonism (Ponnikas et al. 2018). The accumulation of genes with sexually antagonistic effect might have happened after recombination suppression and favour sex-biased gene expression as a general mechanism to resolve sexual antagonism. To some degree, this could account for persistent sex chromosome homomorphy over evolutionary time.

4.3 A higher magnitude of expression divergence between sexes for genes on sex-linked versus non-sex-linked portions of the sex chromosome

When we included all the genes (FDR = 1) in the *X. borealis* sex-linked region, we found a higher magnitude of expression divergence for genes on the sex-linked region than the non-sex-linked portion in all tissue types and the differences were significant for all tissues other than in mesonephros/gonads at tadpole stage 46. However, with only significantly sex-biased genes (FDR <= 0.05), this comparison was not significantly different in adult liver and mesonephros/gonads at stage 46. However, the magnitude of expression divergence was significantly higher on the sex-linked region compared to the non-sex-linked region in adult gonad and was higher but not significantly so in mesonephros/gonads at tadpole stage 48. This suggested that there's no strong signal that significantly sex-biased genes in the sex-linked regions of the sex chromosome evolved to have a higher magnitude of expression divergence compared to the non-sex-linked region. The results for significantly sex-biased genes are consistent with findings from Rana temporaria, which has a pro-Y chromosome (Ma et al. 2018a). This suggests that even though sex biased gene expression may resolve sexual antagonism; it might not be necessary to evolve a greater magnitude of expression divergence to resolve the sexual conflict.

A possibility is that we underestimated the magnitude of expression divergence because we pooled splice variants, which could have different or even opposing patterns of sex-biased expression. In D. melanogaster, the male sex biased genes from the sex chromosome tends to have greater expression divergence than female biased genes (Ranz et al. 2003; Ellegren and Parsch 2007). In plant Salix viminalis, the magnitude of sex bias in male-biased genes from the genome was greater than that of female-biased genes (Darolti et al. 2018). Another possibility is that significantly sex-biased genes represented on going resolution of genomic conflict. Cheng and Kirkpatrick Cheng and Kirkpatrick (2016) found that the strength of sex-specific selection, measured by average genetic divergence (FST) between males and females for SNPs in humans or in flies, were highest for autosomal loci with expression that is moderately female-biased or moderately male-biased. FST was lower for autosomal loci with highly female-biased, highly male-biased, or unbiased expression. Expression divergence of highly expressed genes are likely to be constrained by pleiotropy, which could act against the evolution of highly sex-biased expression (Drummond et al. 2005; Darolti et al. 2018). It is possible that we included genes with less stringent filters (FDR = 1), we are including more genes that is lowly expressed which could have higher magnitude of sex-biased due to less pleiotropy constrain or as an artifact of low expression level. When we included genes with more stringent filtering (FDR <= 0.15, or 0.1, or 0.15), we were including genes with mostly moderate or higher expression level which might not have extreme magnitude of sex-biased due to higher level of pleiotropy constrained and is experiencing on-going sex-specific selection. Difference in pleiotropic constraints thus could account for the different statistical signatures we recovered ad different stringency levels (i.e., significant enrichment of high expression divergence between sexes with FDR =1, but

non-significant higher when FDR <1).

4.4 Expression divergence of genes in the nonrecombining region evolve at a faster rate

The X (or Z) chromosome is hemizygous in the heterogametic sex and genes with recessive male-beneficial mutations could be more efficiently targeted for retention by natural selection if they are on the X chromosome compared to if they are on the autosome. In this way, natural selection could contribute to a faster evolutionary rate of X-linked genes (Charlesworth et al. 1987; Parsch and Ellegren 2013) and this has been reported in chimpanzees (Hvilsom et al. 2012) and Drosophila fruit flies (Baines et al. 2008) based on protein sequence divergences. Expression divergence of genes on the sex chromosome also appears to evolve faster compared to genes on the autosome in Drosophila fruit flies (Kayserili et al. 2012; Meisel et al. 2012). With a Brownian Motion (BM) model, we investigated the evolution rate of expression divergence (between males and females) for genes in the *X. borealis* chr8L nonrecombining region. We found that the gene expression ratio of genes in the *X. borealis* chr8L nonrecombining region evolved at a faster rate. Three-parameters BM model and four-parameters BM model suggested the faster evolution rates of the sex-linked branches seen in the two-parameter models were driven by L subgenome evolution rather than recombination suppression.

If the evolution of L subgenome was the main driver of gene expression divergence, we would expect to observe the same when we repeat the same four BM analyses for genes in the recombining region of *X. borealis*. However, contrary to this expectation, the three-rate models suggested the S subgenome branches were evolving faster than the L subgenome branches. Additionally, the four-rate models, expression divergence between sexes in *X. laeivs* 8L branches evolved faster than S subgenome branches, *X. borealis* 8L branches, and the background (in the order of decreasing evolutionary rates respectively). Taken together, these BM results highlighted the complex dynamic of expression divergence evolution of genes in the sex-linked regions. This complexity is a result of both subgenome effects and effects of sex-linkage. A study of D. melanogaster also used a BM model to investigate the evolutionary rates of gene expression divergence in oocyte and embryonic transcriptomes (Paris et al. 2015). They found that expression divergence of genes on the sex chromosomes to evolved more quickly than genes on the autosomes in egg but not in embryo (Paris et al. 2015). In species where the XY chromosome does not recombine, such as D. melanogaster, there's enrichment of sex-linked genes with higher expression in the homogametic sex; however, in species such as silkworm, where the recombination rates are similar between sexes, there's enrichment of sex biased genes in the sex-linked region compared to autosomes with higher expression in the heterogametic sex (Ranz et al. 2003; Khil et al. 2004; Arunkumar and Nagaraju 2009). Taken together, these studies suggested that recombination rate among sex chromosome affects the expression divergence of genes (Connallon and Clark 2010).

Our efforts to resolve the directionality of expression regulation (i.e. whether expression divergence resulted from upregulation in females or males) was unsuccessful. This analysis was based on the mean (and median) of expression data that was normalized across different species. The mean normalized expression was higher in *X. borealis* than *X. laeivs*; although the median was similar between them. If the lack of significant result was an artifact of cross-species normalization, we should see variations in mean (or median) of gene expression from *X. borealis* 8L (sex-linked) and from *X. borealis* 8S (non-sex-linked) since they are from the same set expression data from *X. borealis* transcriptome and no normalization was needed. In Drosophila melanogaster, male-biased expression was inferred to be a consequence of upregulated expression in males whereas female-biased upregulated expression was inferred to be a consequence of upregulated expression in females (Vicoso and Charlesworth 2009). We will need more data and study to gain insight into the directionality of expression divergence in *X. borealis*.

5 CONCLUSION

In the Kenyan clawed frog *X. borealis*, a higher proportion of genes on a large and newly evolved non-recombining region of the sex chromosome have sex-biased expression as compared to recombining portions of the sex chromosome or the autosomes. One explanation for this finding is that natural selection favoured the evolution of sex-biased expression in order to resolve genomic conflict stemming from mutations with SA fitness effects. This phenomenon is likely evolved after chr8L become the sex chromosome. Our analyses of WGS data suggest that sex-biased expression patterns generally do not stem from W chromosome degeneration. We also found that expression divergence of genes in the non-recombining region evolved faster than its homeologous genes. Additionally, the magnitude of expression divergence was higher in sex-linked genes compared to non-sex-linked genes on the sex chromosomes. However, we didn't a detect signal of significantly sex-biased genes in the nonrecombining region having significantly higher magnitude of expression divergence, nor faster rate of expression divergence. However, we have a very small sample size of significantly sex-biased genes and a lack of significance could be a consequence of low statistical power. Overall, we do see evidence that lack of recombination could influence expression of sex-linked genes. This intersexual divergence in gene expression could play a fundamental role in sexually dimorphic phenotypes, including morphology and behaviour.

Figures

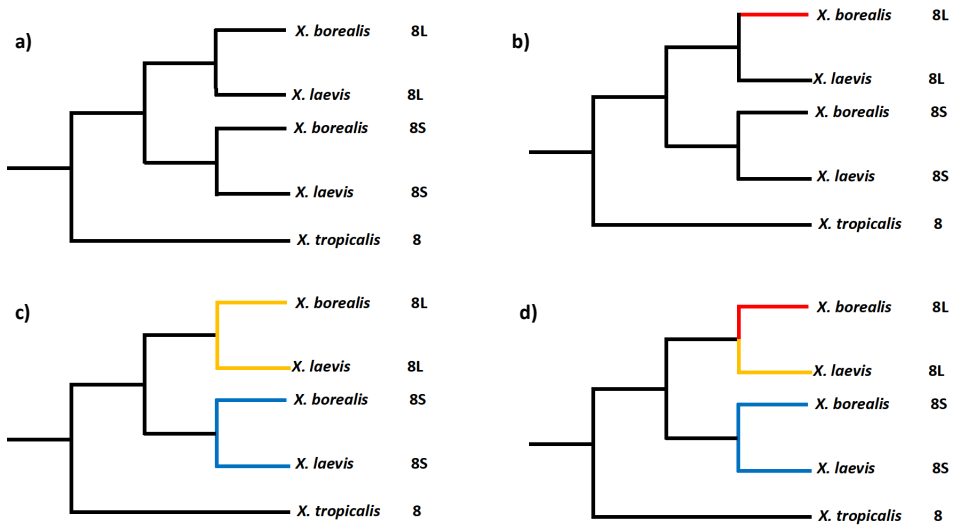

Figure 1: Phylogenetic tree and models considered for Brownian motion models. a) One-rate model; all the branches evolve under the same rate. b) Two-rate model; the red branch, on which sex-linked genes on *X. borealis* chr8L evolve, has a different rate from other orthologous or homeologous branches in black. c) Three-rate model; the red, blue, and black branches allow separate rates of expression divergence of genes in the L L subgenome, S subgenome, and other contexts, respectively. d) Four-rate model; an additional parameter represented by colour allows for a different rate of evolution between sex-linked and non-sex-linked orthologous genes within subgenome L.

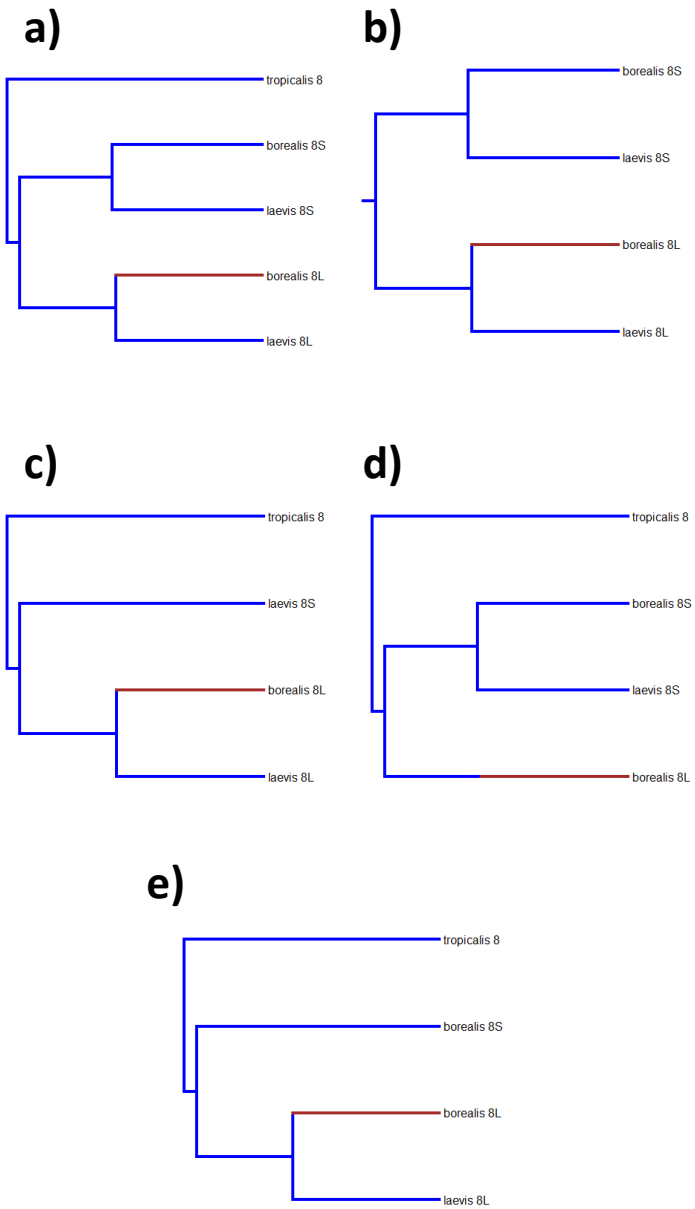

Figure 2: Possible type of data set used in the Brownian motion analyses. Using 2-rate model as an example, we considered data with a) no missing data points, b) missing data for X. tropicalis chr8, c) missing data for *X. laevis* chr8S, d) missing data for *X. borealis* chr8L, and e) missing data for X.laevis chr8S

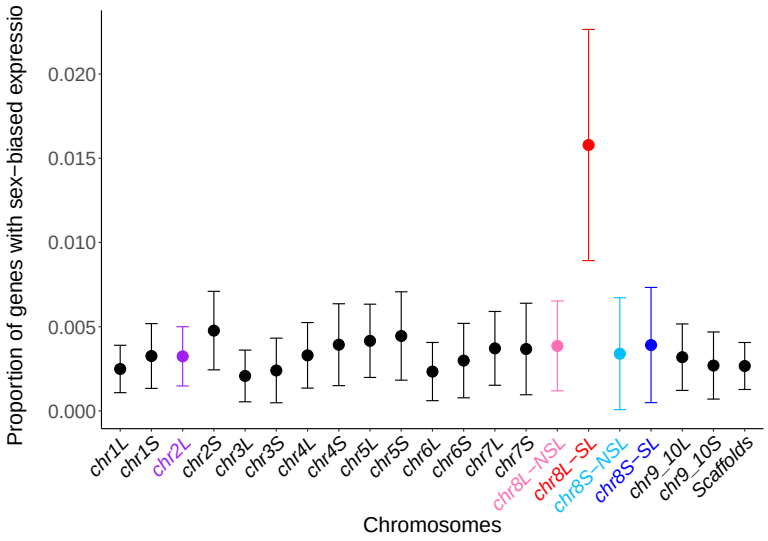

Figure 3: Proportions by chromosome and sex-linkage of sex-biased genes in adult liver (DESeq2).

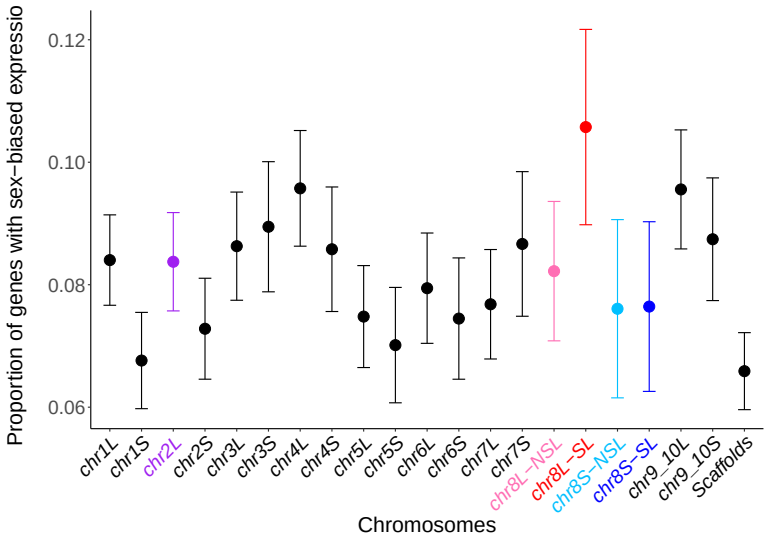

Figure 4: Proportions by chromosome and sex-linkage of sex-biased genes in adult gonad (DESeq2).

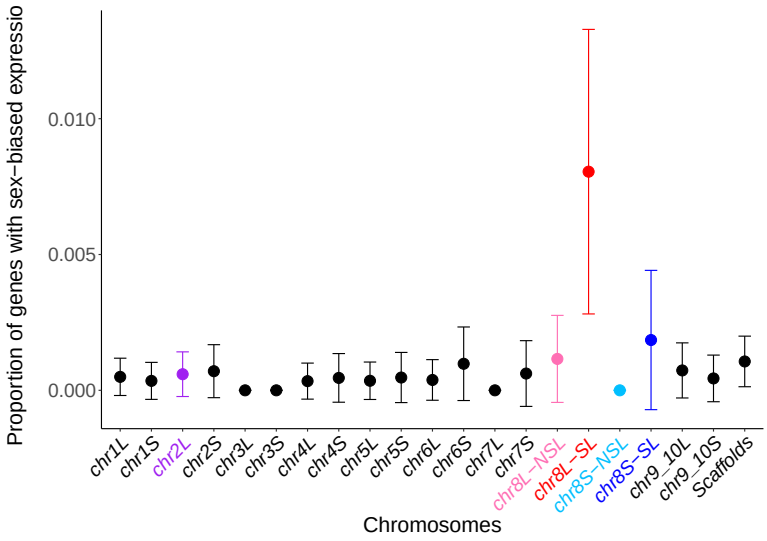

Figure 5: Proportions by chromosome and sex-linkage of sex-biased genes in tadpole stage46 gonad/mesenephros (DESeq2).

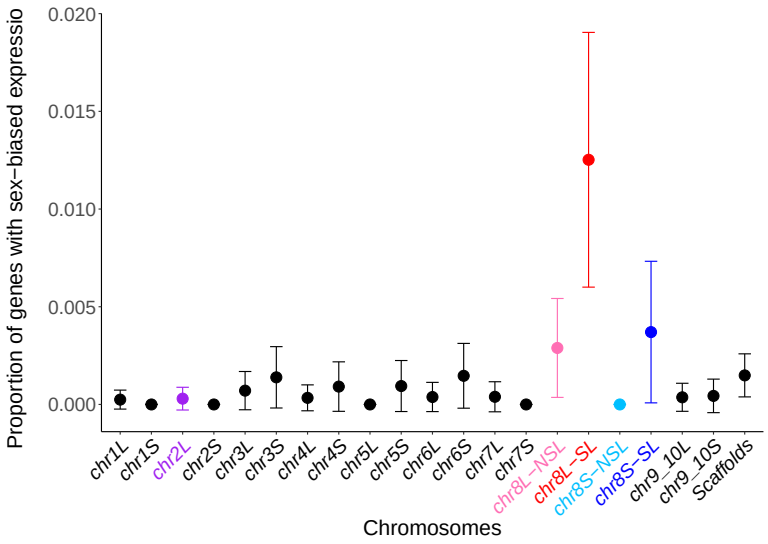

Figure 6: Proportions by chromosome and sex-linkage of sex-biased genes in tadpole stage48 gonad/mesenephros (DESeq2).

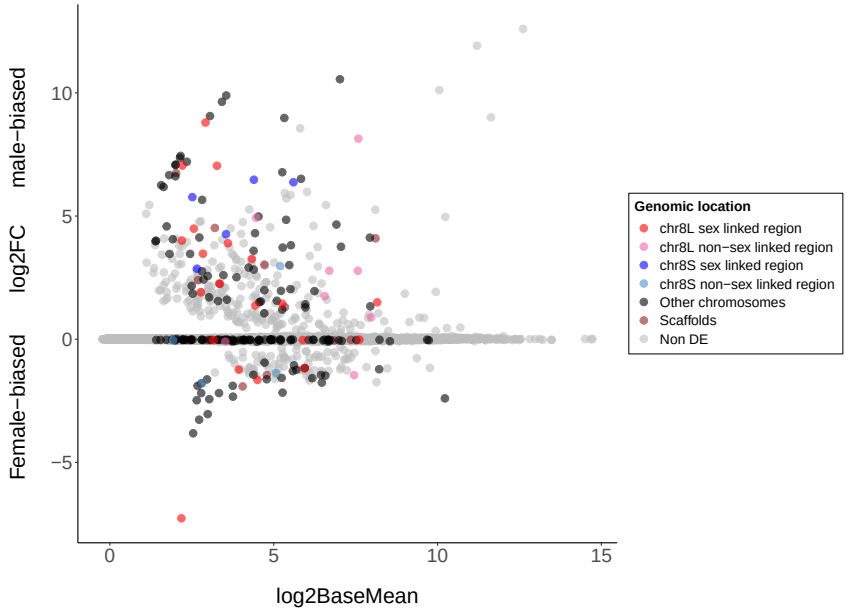

Figure 7: M-A plot of differentially expressed genes in adult liver (DE-Seq2, FDR $\leq$0.05).

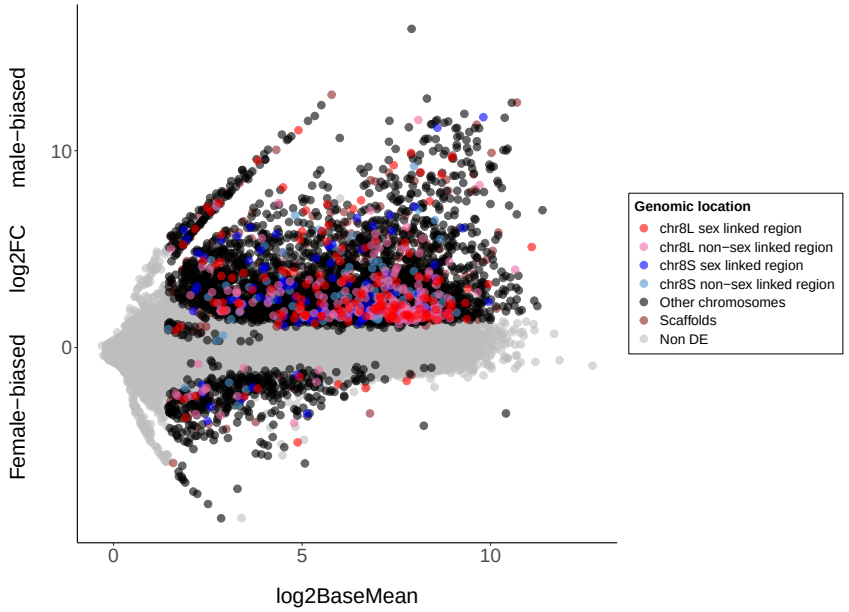

Figure 8: M-A plot of differentially expressed genes in adult gonad (DESeq2, FDR $\leq$0.05).

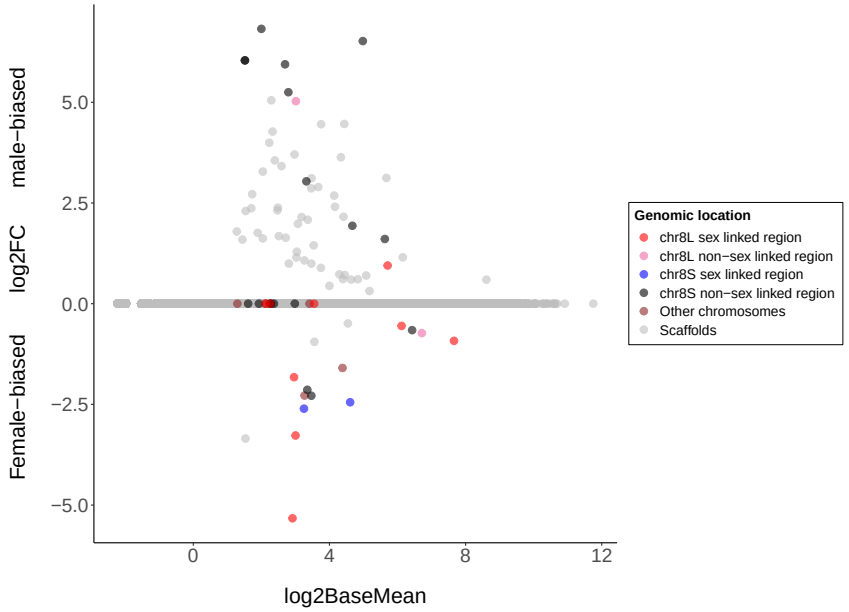

Figure 9: M-A plot of differentially expressed genes in tadpole stage46 gonad/mesenephros (DESeq2, FDR $\leq$0.05).

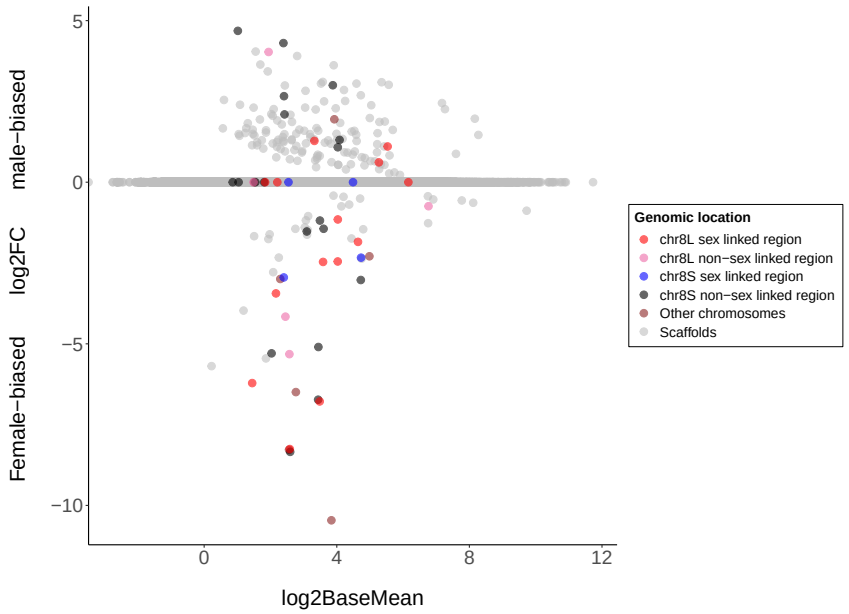

Figure 10: M-A plot of differentially expressed genes in tadpole stage48 gonad/mesenephros (DESeq2, FDR $\leq$0.05).

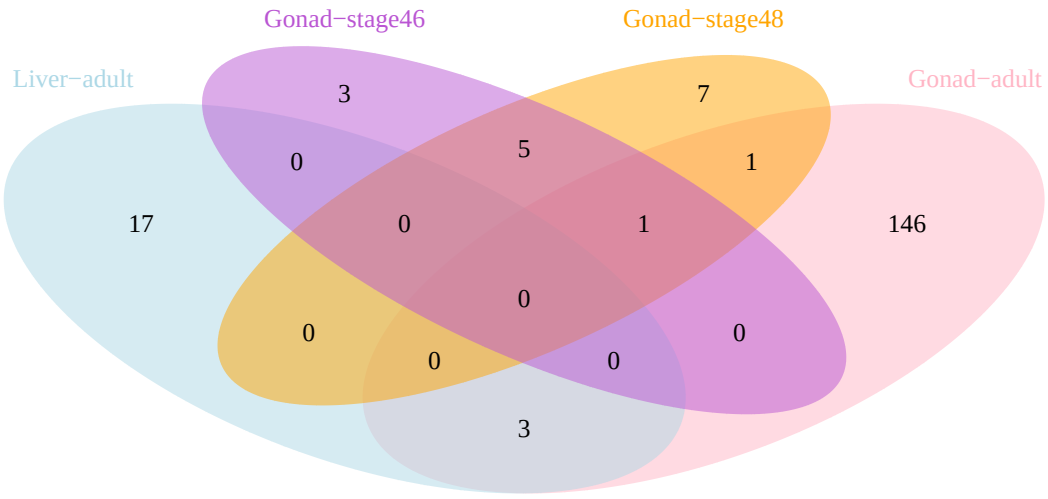

Figure 11: Venn Diagram of the number of sex-biased genes in each tissue type and developmental stage (DESeq2, FDR $\leq$0.05).

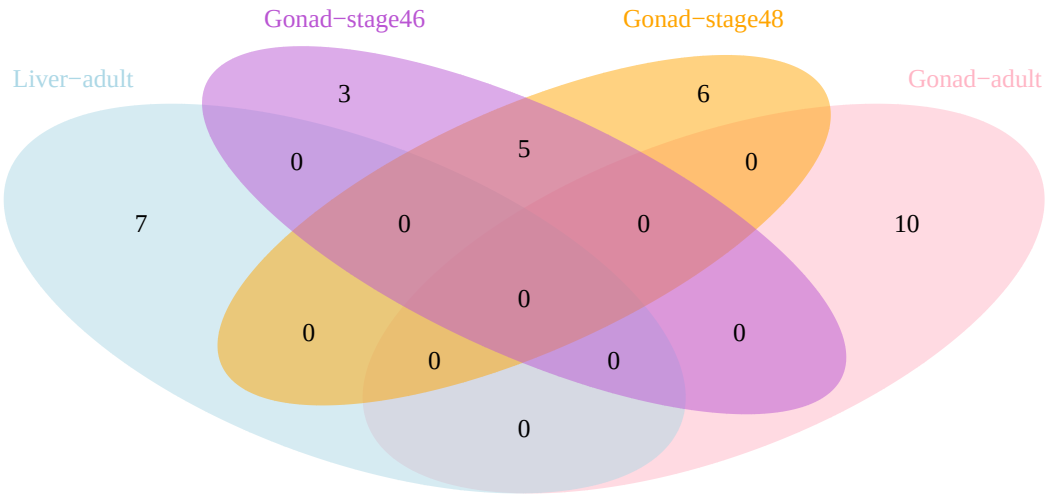

Figure 12: Venn Diagram of the number of female-biased genes in each tissue type and developmental stage (DESeq2, FDR $\leq$0.05).

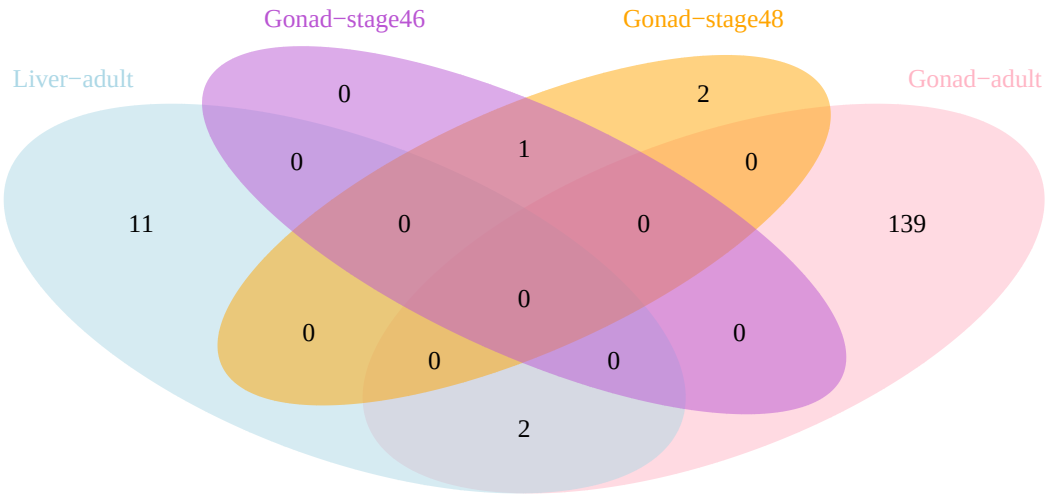

Figure 13: Venn Diagram of the number of male-biased genes in each tissue type and developmental stage (DESeq2, FDR $\leq$0.05).

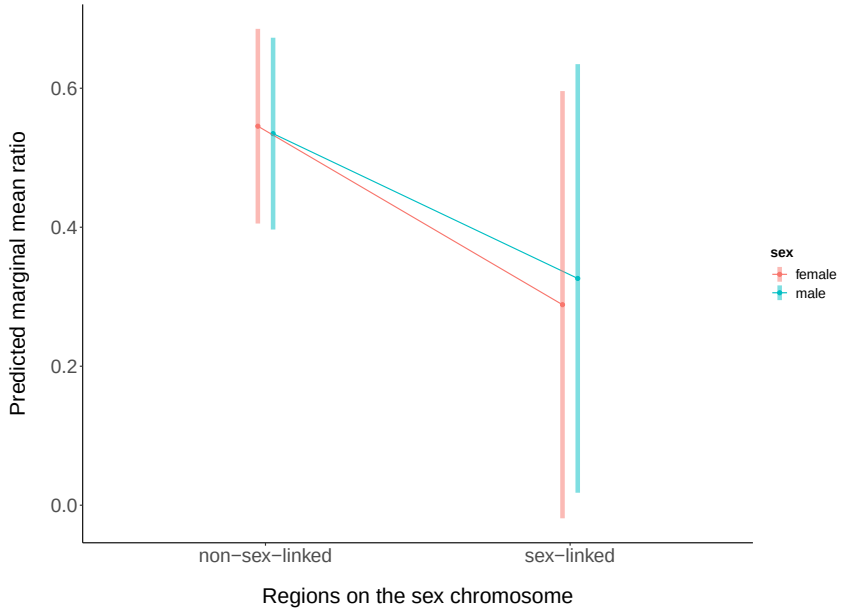

Figure 14: Plot of interaction between the L/S ratio and whether or not genes was sex-linked in *X. borealis* adult liver.

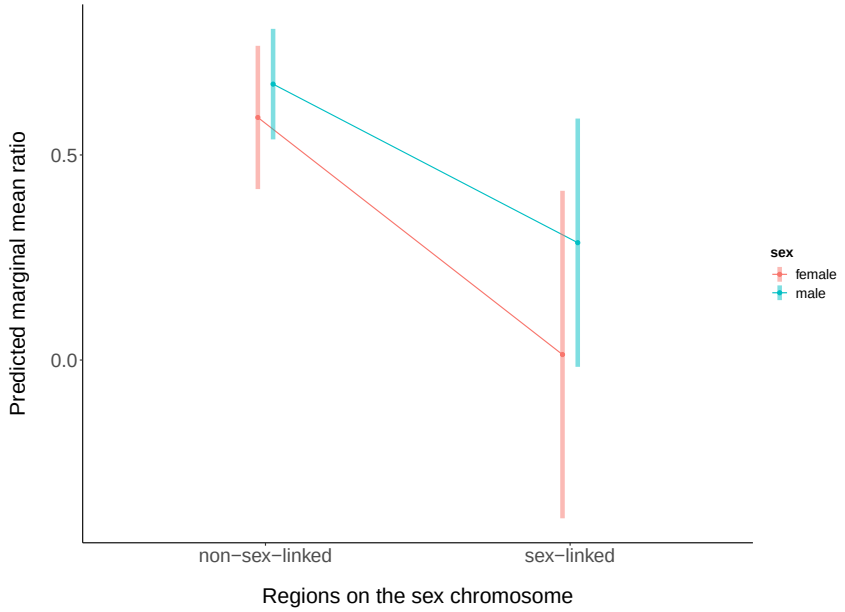

Figure 15: Plot of interaction between the L/S ratio and whether or not genes was sex-linked in *X. borealis* adult gonad.

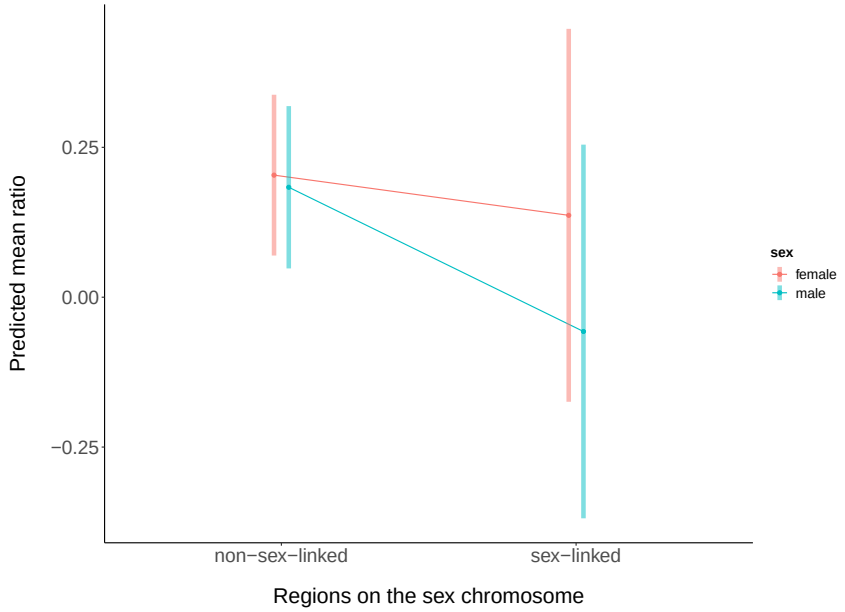

Figure 16: Plot of interaction between the L/S ratio and whether or not genes was sex-linked in *X. borealis* tadpole at stage 46 gonad/mesenephros.

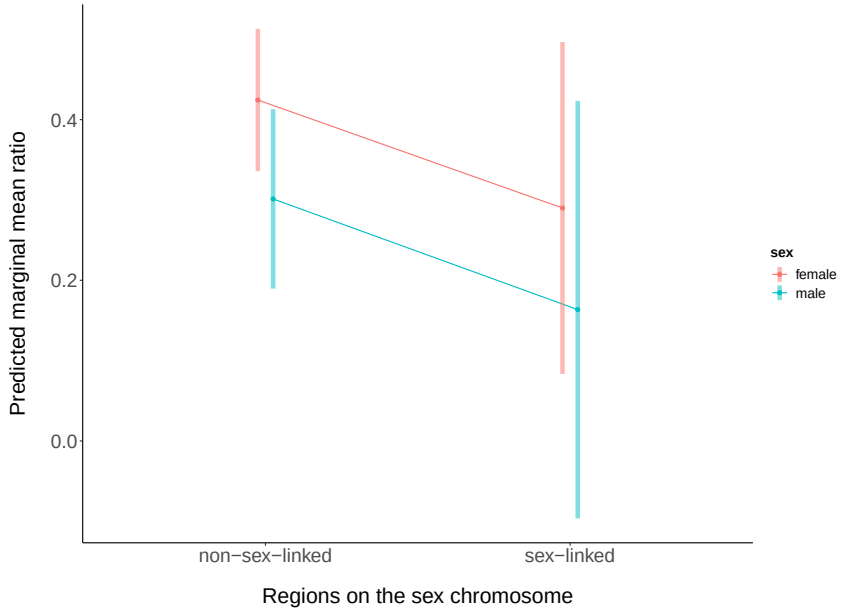

Figure 17: Plot of interaction between the L/S ratio and whether or not genes was sex-linked in *X. borealis* tadpole at stage 48 gonad/mesenephros.

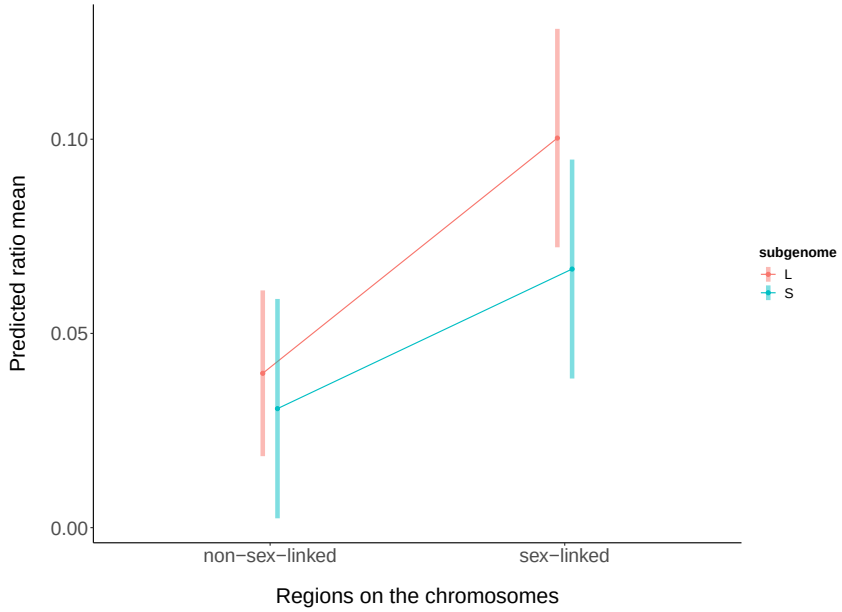

Figure 18: Plot of the interaction between magnitude of expression divergence and whether or not a gene was sex-linked in *X. borealis* adult liver (filtering FDR =1; DESeq2).

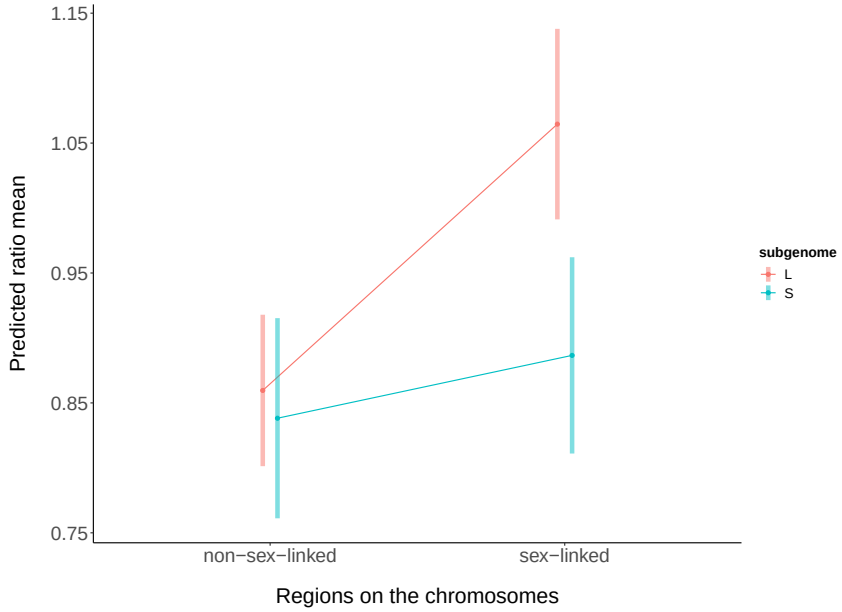

Figure 19: Plot of the interaction between magnitude of expression divergence and whether or not a gene was sex-linked in *X. borealis* adult gonad (filtering FDR =1).

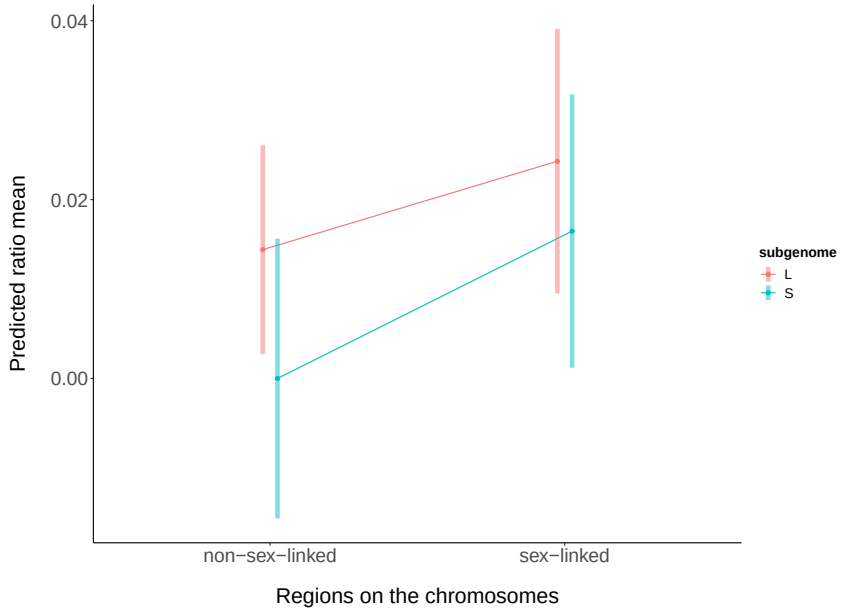

Figure 20: Plot of the interaction between magnitude of expression divergence and whether or not a gene was sex-linked in *X. borealis* tadpole stage 46 gonad/mesenephros (filtering FDR =1).

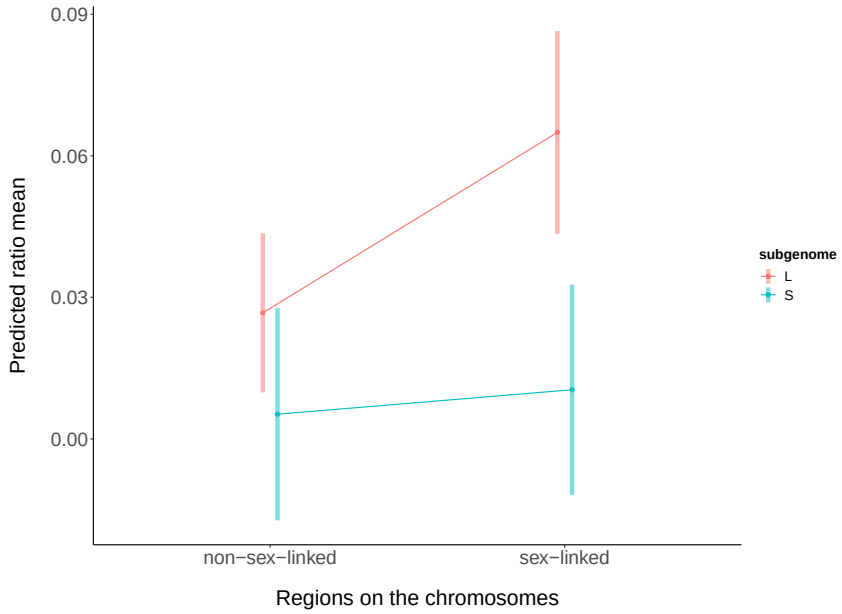

Figure 21: Plot of the interaction between magnitude of expression divergence and whether or not a gene was sex-linked in *X. borealis* tadpole stage 48 gonad/mesenephros (filtering FDR =1).

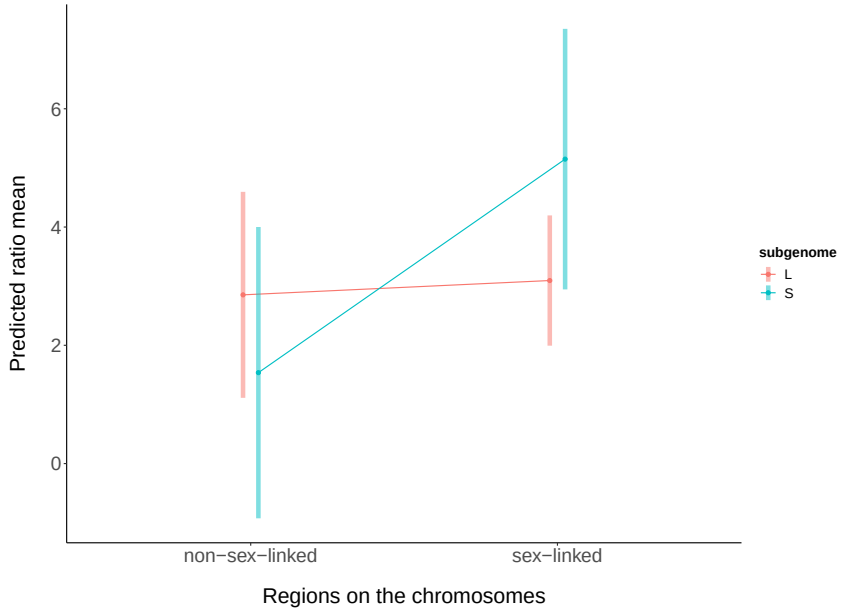

Figure 22: Plot of the interaction between magnitude of expression divergence and whether or not a gene was sex-linked in *X. borealis* adult liver (filtering FDR ≤ 0.05).

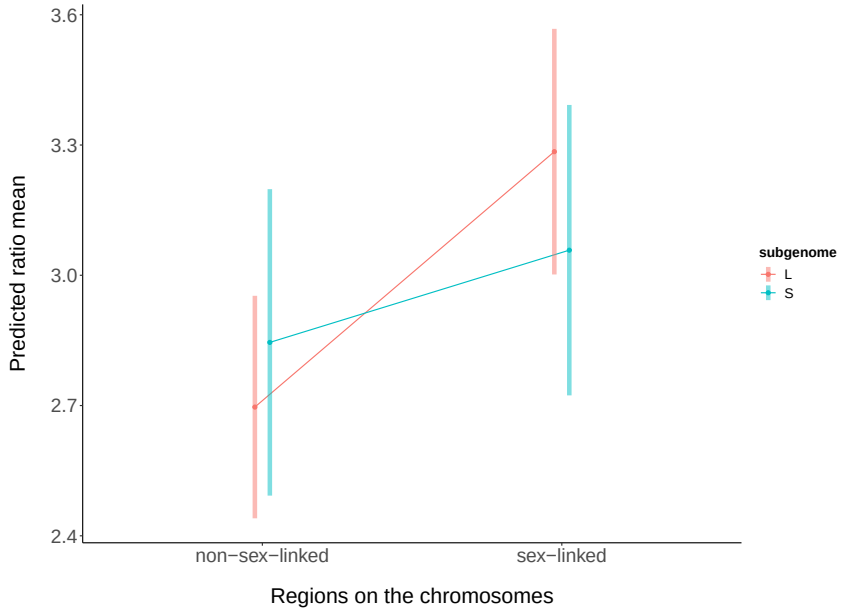

Figure 23: Plot of the interaction between magnitude of expression divergence and whether or not a gene was sex-linked in *X. borealis* adult gonad (filtering FDR ≤ 0.05).

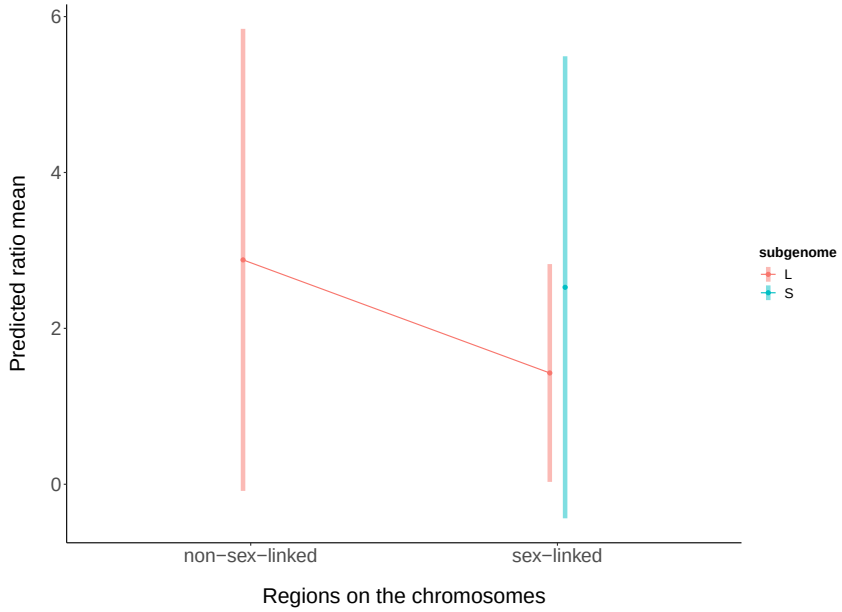

Figure 24: Plot of the interaction between magnitude of expression divergence and whether or not a gene was sex-linked in *X. borealis* tadpole stage46 gonad/mesenephros (filtering FDR ≤ 0.05).

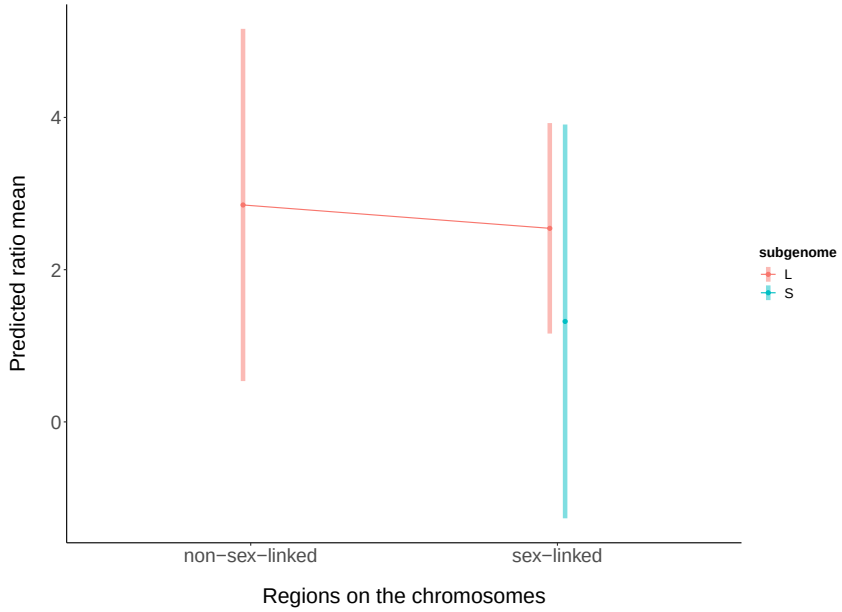

Figure 25: Plot of the interaction between magnitude of expression divergence and whether or not a gene was sex-linked in *X. borealis* tadpole stage48 gonad/mesenephros (filtering FDR ≤ 0.05).

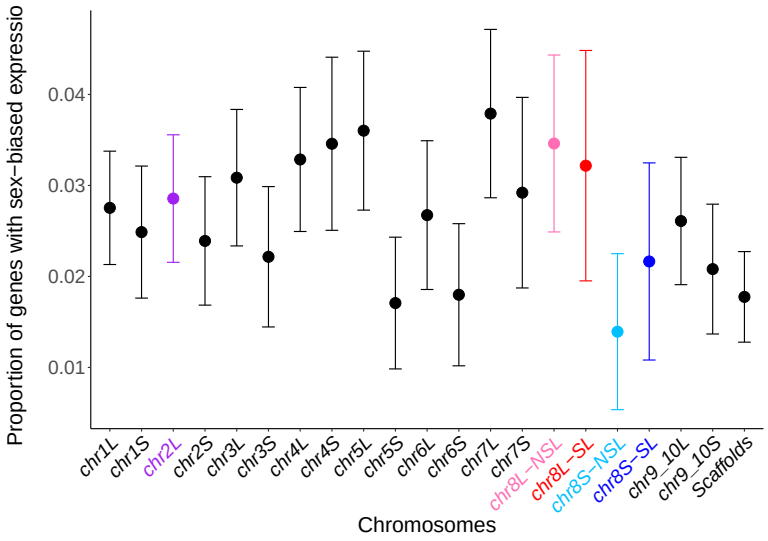

Figure 26: Proportions by chromosome and sex-linkage of sex-biased genes in *X. laevis* adult liver (DESeq2, FDR $\leq$0.05).

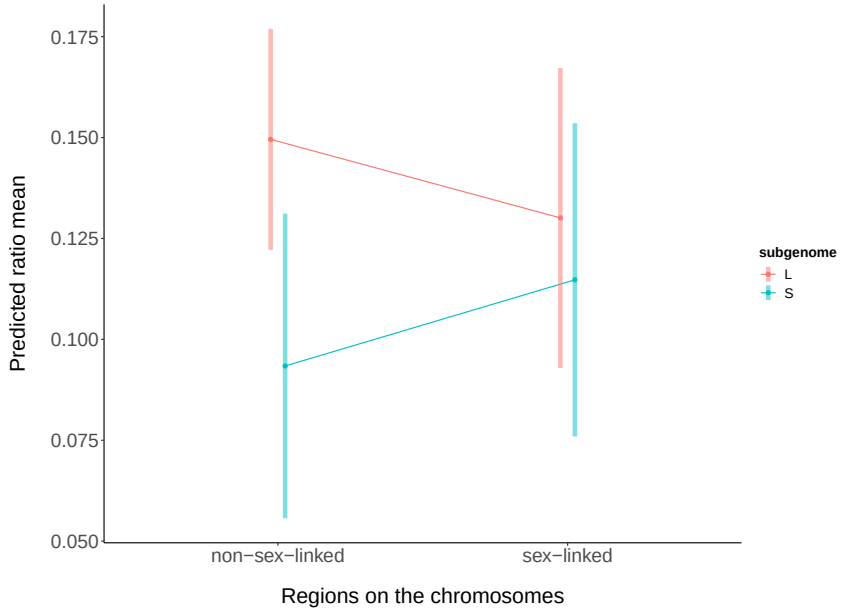

Figure 27: Interaction plot for analyses of magnitude of expression divergence in *X. laevis* adult liver (filtering FDR = 1).

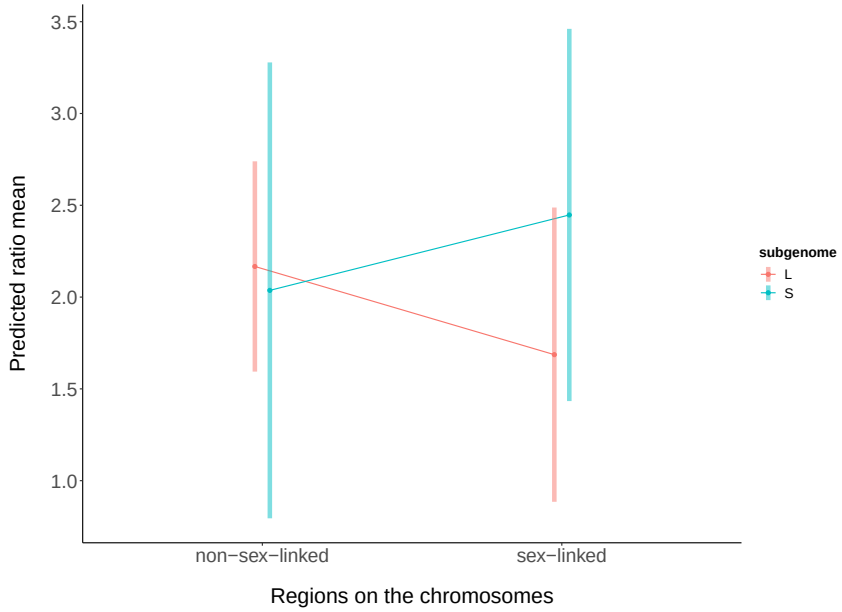

Figure 28: Plot of the interaction between magnitude of expression divergence and whether or not a gene was sex-linked in *X. laevis* adult livers (filtering FDR ≤ 0.05).

Tables

Table 1: Differentially expressed genes in *X. borealis* adult liver (DESeq2). Chr8L was divided into the sex-linked (SL) region and non-sex-linked (NSL) region; the same division was done for its homeologous region on chr8S.

Regions	Total gene	DE count	DE proportion	Male biased count	Female biased count
chr1L	4,821	12	$2.49 \cdot 10^{-3}$	4	8
chr1S	3,376	11	$3.26 \cdot 10^{-3}$	4	7
chr2L	4,013	13	$3.24 \cdot 10^{-3}$	5	8
chr2S	3,356	16	$4.77 \cdot 10^{-3}$	4	12
chr3L	3,375	7	$2.07 \cdot 10^{-3}$	3	4
chr3S	2,500	6	$2.4 \cdot 10^{-3}$	4	2
chr4L	3,335	11	$3.3 \cdot 10^{-3}$	5	6
chr4S	2,545	10	$3.93 \cdot 10^{-3}$	6	4
chr5L	3,365	14	$4.16 \cdot 10^{-3}$	7	7
chr5S	2,473	11	$4.45 \cdot 10^{-3}$	5	6
chr6L	2,998	7	$2.33 \cdot 10^{-3}$	2	5
chr6S	2,344	7	$2.99 \cdot 10^{-3}$	3	4
chr7L	2,963	11	$3.71 \cdot 10^{-3}$	1	10
chr7S	1,905	7	$3.67 \cdot 10^{-3}$	4	3
chr8L-NSL	2,074	8	$3.86 \cdot 10^{-3}$	6	2
chr8L-SL	1,267	20	$1.58 \cdot 10^{-2}$	13	7
chr8S-NSL	1,178	4	$3.4 \cdot 10^{-3}$	1	3
chr8S-SL	1,279	5	$3.91 \cdot 10^{-3}$	5	0
chr9_10L	3,136	10	$3.19 \cdot 10^{-3}$	1	9
chr9_10S	2,599	7	$2.69 \cdot 10^{-3}$	1	6
Scaffolds	5,251	14	$2.67 \cdot 10^{-3}$	7	7

Table 2: Differentially expressed genes in *X. borealis* adult gonad (DE-Seq2). Labels correspond to Table.1

Regions	Total gene	DE count	DE proportion	Male biased count	Female biased count
chr1L	5,427	456	$8.4 \cdot 10^{-2}$	416	40
chr1S	3,919	265	$6.76 \cdot 10^{-2}$	246	19
chr2L	4,573	383	$8.38 \cdot 10^{-2}$	361	22
chr2S	3,818	278	$7.28 \cdot 10^{-2}$	264	14
chr3L	3,882	335	$8.63 \cdot 10^{-2}$	319	16
chr3S	2,772	248	$8.95 \cdot 10^{-2}$	226	22
chr4L	3,729	357	$9.57 \cdot 10^{-2}$	329	28
chr4S	2,914	250	$8.58 \cdot 10^{-2}$	225	25
chr5L	3,837	287	$7.48 \cdot 10^{-2}$	268	19
chr5S	2,823	198	$7.01 \cdot 10^{-2}$	185	13
chr6L	3,462	275	$7.94 \cdot 10^{-2}$	248	27
chr6S	2,699	201	$7.45 \cdot 10^{-2}$	185	16
chr7L	3,411	262	$7.68 \cdot 10^{-2}$	235	27
chr7S	2,181	189	$8.67 \cdot 10^{-2}$	171	18
chr8L-NSL	2,238	184	$8.22 \cdot 10^{-2}$	169	15
chr8L-SL	1,428	151	0.11	141	10
chr8S-NSL	1,275	97	$7.61 \cdot 10^{-2}$	93	4
chr8S-SL	1,413	108	$7.64 \cdot 10^{-2}$	100	8
chr9_10L	3,516	336	$9.56 \cdot 10^{-2}$	308	28
chr9_10S	3,054	267	$8.74 \cdot 10^{-2}$	244	23
Scaffolds	5,980	394	$6.59 \cdot 10^{-2}$	366	28

Table 3: Differentially expressed genes in *X. borealis* tadpole stage 46 gonad/mesenephros (DESeq2). Labels correspond to Table.1

Regions	Total gene	DE count	DE proportion	Male biased count	Female biased count
chr1L	4,036	2	$5 \cdot 10^{-4}$	1	1
chr1S	2,877	1	$3.5 \cdot 10^{-4}$	1	0
chr2L	3,370	2	$5.9 \cdot 10^{-4}$	1	1
chr2S	2,840	2	$7 \cdot 10^{-4}$	1	1
chr3L	2,830	0	0	0	0
chr3S	2,161	0	0	0	0
chr4L	2,954	1	$3.4 \cdot 10^{-4}$	0	1
chr4S	2,188	1	$4.6 \cdot 10^{-4}$	0	1
chr5L	2,844	1	$3.5 \cdot 10^{-4}$	1	0
chr5S	2,122	1	$4.7 \cdot 10^{-4}$	1	0
chr6L	2,620	1	$3.8 \cdot 10^{-4}$	1	0
chr6S	2,046	2	$9.8 \cdot 10^{-4}$	0	2
chr7L	2,548	0	0	0	0
chr7S	1,619	1	$6.2 \cdot 10^{-4}$	1	0
chr8L-NSL	1,729	2	$1.16 \cdot 10^{-3}$	1	1
chr8L-SL	1,118	9	$8.05 \cdot 10^{-3}$	1	8
chr8S-NSL	1,035	0	0	0	0
chr8S-SL	1,080	2	$1.85 \cdot 10^{-3}$	0	2
chr9_10L	2,731	2	$7.3 \cdot 10^{-4}$	0	2
chr9_10S	2,283	1	$4.4 \cdot 10^{-4}$	1	0
Scaffolds	4,704	5	$1.06 \cdot 10^{-3}$	0	5

Table 4: Differentially expressed genes in X. borealis tadpole stage 48 gonad/mesenephros (DESeq2). Labels correspond to Table.1

Regions	Total gene	DE count	DE proportion	Male biased count	Female biased count
chr1L	4,036	1	$2.5 \cdot 10^{-4}$	1	0
chr1S	2,877	0	0	0	0
chr2L	3,370	1	$3 \cdot 10^{-4}$	0	1
chr2S	2,840	0	0	0	0
chr3L	2,830	2	$7.1 \cdot 10^{-4}$	1	1
chr3S	2,161	3	$1.39 \cdot 10^{-3}$	1	2
chr4L	2,954	1	$3.4 \cdot 10^{-4}$	1	0
chr4S	2,188	2	$9.1 \cdot 10^{-4}$	0	2
chr5L	2,844	0	0	0	0
chr5S	2,122	2	$9.4 \cdot 10^{-4}$	0	2
chr6L	2,620	1	$3.8 \cdot 10^{-4}$	1	0
chr6S	2,046	3	$1.47 \cdot 10^{-3}$	1	2
chr7L	2,548	1	$3.9 \cdot 10^{-4}$	0	1
chr7S	1,619	0	0	0	0
chr8L-NSL	1,729	5	$2.89 \cdot 10^{-3}$	1	4
chr8L-SL	1,118	14	$1.25 \cdot 10^{-2}$	3	11
chr8S-NSL	1,035	0	0	0	0
chr8S-SL	1,080	4	$3.7 \cdot 10^{-3}$	0	4
chr9_10L	2,731	1	$3.7 \cdot 10^{-4}$	0	1
chr9_10S	2,283	1	$4.4 \cdot 10^{-4}$	1	0
Scaffolds	4,704	7	$1.49 \cdot 10^{-3}$	1	6

Table 5: Analyses of magnitude of expression divergence in *X. borealis* with different level of filtering stringency (DESeq2). Stage indicated the developmental stage, in which Tad st46 and Tad st48 represent tadpole at stage 46 and stage 48 respectively. FDR indicated the filtering FDR used in filtering of differential expression output. Int_coef represent interaction coefficient from the rrpp.lm model. Pr(>F) indicated the p-value for the interaction terms. Pr(>F) lower than 0.05 were in bold. NA indicated the model fitting failed due to lack of data or extremely small data set.

Stages	Tissue type	FDR	Int coef	Pr(>F)
Adult	liver	1	$-2.46 \cdot 10^{-2}$	0.38
Adult	liver	0.15	0.75	0.5
Adult	liver	0.1	0.88	0.53
Adult	liver	$5 \cdot 10^{-2}$	3.37	$9.4 \cdot 10^{-2}$
Adult	gonad	1	-0.16	$3.1 \cdot 10^{-2}$
Adult	gonad	0.15	-0.18	0.4
Adult	gonad	0.1	-0.2	0.42
Adult	gonad	$5 \cdot 10^{-2}$	-0.38	0.25
Tad st46	gonad	1	$6.6 \cdot 10^{-3}$	0.65
Tad st46	gonad	0.15	NaN	NaN
Tad st46	gonad	0.1	NaN	NaN
Tad st46	gonad	$5 \cdot 10^{-2}$	NaN	NaN
Tad st48	gonad	1	$-3.31 \cdot 10^{-2}$	0.12
Tad st48	gonad	0.15	1.27	0.56
Tad st48	gonad	0.1	NaN	NaN
Tad st48	gonad	$5 \cdot 10^{-2}$	NaN	NaN

Table 6: Permutation test of analyses of magnitude of expression divergence in *X. borealis* with different level of filtering stringency (DESeq2). Stage indicated the developmental stage, in which Tad st46 and Tad st48 represent tadpole at stage 46 and stage 48 respectively. FDR indicated the FDR used in each case. Total Perm indicated the total permutation done. Rank indicated the rank of the observed interaction coefficient among permutation interaction coefficient.

Stages	Tissue type	FDR	Total Perm	Rank	Rank5	Rank95
Adult	liver	1	70	1	3.5	66.5
Adult	liver	0.15	70	14	3.5	66.5
Adult	liver	0.1	70	26	3.5	66.5
Adult	liver	0.05	70	53	3.5	66.5
Adult	gonad	1	35	1	1.75	33.25
Adult	gonad	0.15	35	5	1.75	33.25
Adult	gonad	0.1	35	6	1.75	33.25
Adult	gonad	0.05	35	9	1.75	33.25
Tad st46	gonad	1	70	1	3.5	66.5
Tad st46	gonad	0.15	70	NA	3.5	66.5
Tad st46	gonad	0.1	70	NA	3.5	66.5
Tad st46	gonad	0.05	70	NA	3.5	66.5
Tad st48	gonad	1	1287	1	64.35	1222.65
Tad st48	gonad	0.15	1287	1	64.35	1222.65
Tad st48	gonad	0.1	1287	NA	64.35	1222.65
Tad st48	gonad	0.05	1287	NA	64.35	1222.65

Table 7: Analyses of magnitude of expression divergence in *X. laevis* with different level of filtering stringency (DESeq2). FDR indicated the filtering FDR used in filtering of differential expression output. Int_coef represent interaction coefficient from the rrpp.lm model. Pr(>F) indicated the p-value for the interaction terms.

Tissue type	FDR	Int coef	Pr(>F)
liver	1	$4.08 \cdot 10^{-2}$	0.25
liver	0.15	0.7	0.31
liver	0.1	0.72	0.32
liver	$5 \cdot 10^{-2}$	0.89	0.36

Table 8: Examine evolution rate of expression divergence between sexes using Brownian model – model selection using summed logLik and AICc of summed logLik. Rank indicated the ranking of observed AICc in 1000 permutations.

Model	Log-L sum	AICc for logL sum	AICc Rank in bootstrapping
1 rate model	−1,570.73	3,145.54	526
2 rate model	−1,458.38	2,922.94	513
3 rate model	−1,291.43	2,591.14	504
4 rate model	−877	1,764.44	522

Table 9: Examine evolution rate of expression divergence between sexes using Brownian model – pairwise models comparisons using likelihood ratio tests. Gene expression ratio between sexes of genes in *X. borealis* chr8L sex-linked regions and its paralogs and orthologies in *X. borealis* and *X. laevis* were fitted into BM models. The table presented the p-values of likelihood ratio tests against a chi square distribution.

Models	1 rate model	2 rate model	3 rate model
2 rate model	$1.52 \cdot 10^{-5}$		
3 rate model	$8.27 \cdot 10^{-20}$	$2.61 \cdot 10^{-17}$	
4 rate model	$2 \cdot 10^{-101}$	$5.65 \cdot 10^{-106}$	$4.21 \cdot 10^{-97}$

Table 10: Examine evolution rate of expression divergence between sexes using Brownian models – averaged evolution rates of branches from BM models. Gene expression ratio between sexes of genes in *X. borealis* chr8L sex-linked regions and its paralogs and orthologies in *X. borealis* and *X. laevis* were fitted into BM models. In two-parameters model, rate 1 was the non-sex-linked branch and rate 2 was the sex-linked branch. In three-parameters models, rate 1-3 were mean rates for background, S subgenome branches, and L subgenome branches. In four-parameters model, rate 1-4 were mean rates for background, S subgenome, *X. borealis* L branch, *X. laevis* L branch.

Models	rate 1	rate 2	rate 3	rate 4
1 rate model	13.96			
2 rate model	10.14	33.01		
3 rate model	0.26	13.3	21.65	
4 rate model	0.14	14.88	24.5	25.61

Table 11: Explore directionality of gene expression divergence between sexes using Brownian motion models. We used male expressions of genes in *X. borealis* chr8L sex-linked regions. In the 2-parameters model, rate 1 is the non-sex-linked branches and rate 2 is the sex-linked branch.

Model	Likelihood sum	Mean rate 1	Mean rate 2
1 rate model	−2,687.43	$2.31 \cdot 10^6$	
2 rate model	−2,656.19	$2.31 \cdot 10^6$	$2.31 \cdot 10^6$

Table 12: Explore directionality of gene expression divergence between sexes using Brownian motion models. We used male expressions of genes in *X. borealis* chr8L sex-linked regions. In the 2-parameters model, rate 1 is the non-sex-linked branches and rate 2 is the sex-linked branch.

Model	Likelihood sum	Mean rate 1	Mean rate 2
1 rate model	−2,251.45	$2.46 \cdot 10^6$	
2 rate model	−2,201.12	$2.46 \cdot 10^6$	$2.46 \cdot 10^6$

Table 13: Median(TPM) and mean (TPM) of normalized expression of genes in *X. borealis* chr8L sex-linked regions and their paralogs and orthologs in *X. borealis* and *X. laevis*.

Branches	Female Median	Male Median	Female Mean	Male Mean
X. borealis 8L	189.5	203	2493.7	2269.12
X. borealis 8S	82.5	94.77	1405.8	872.24
X. laevis 8L	105.06	149.93	622.62	1313.66
X. laevis 8S	58.67	85.09	581.66	917.85

Supplementary Figures

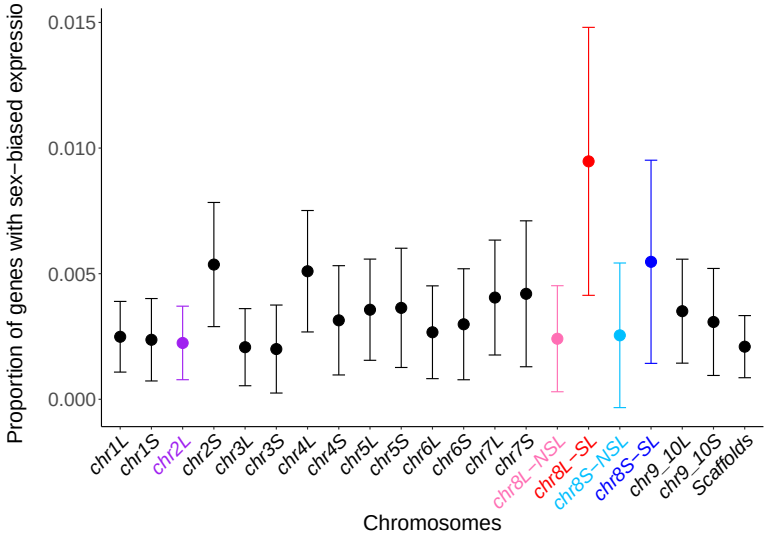

Figure S1: Proportions by chromosome and sex-linkage of sex-biased genes in adult liver (edgeR).

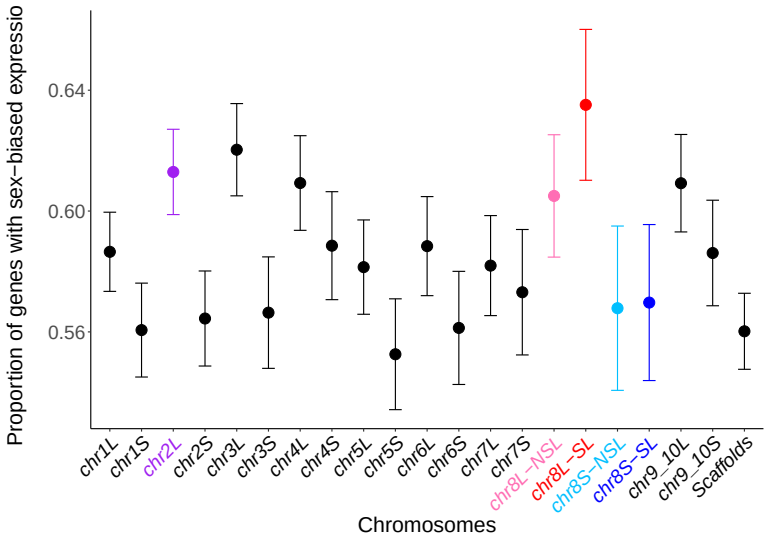

Figure S2: Proportions by chromosome and sex-linkage of sex-biased genes in adult gonad (edgeR).

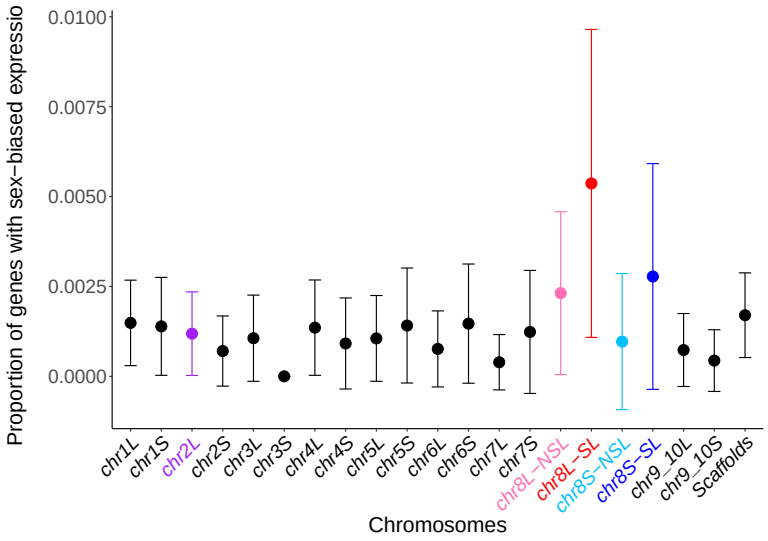

Figure S3: Proportions by chromosome and sex-linkage of sex-biased genes in tadpole stage46 gonad/mesenephros (edgeR).

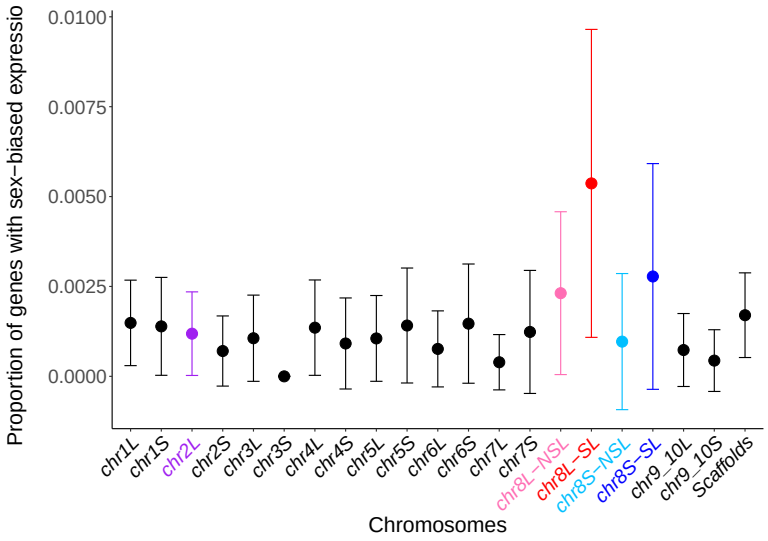

Figure S4: Proportions by chromosome and sex-linkage of sex-biased genes in tadpole stage48 gonad/mesenephros (edgeR).

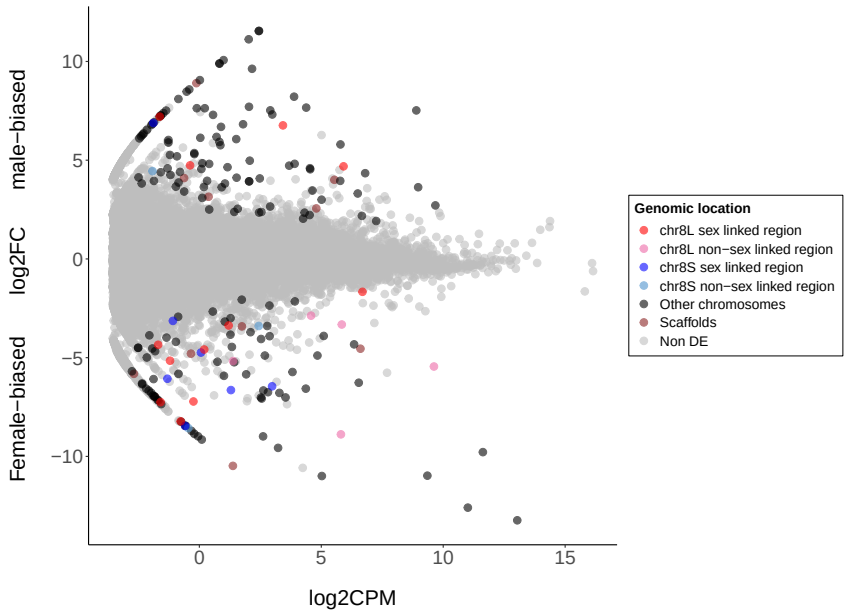

Figure S5: M-A plot of differentially expressed genes in adult liver (edgeR, FDR $\leq$0.05).

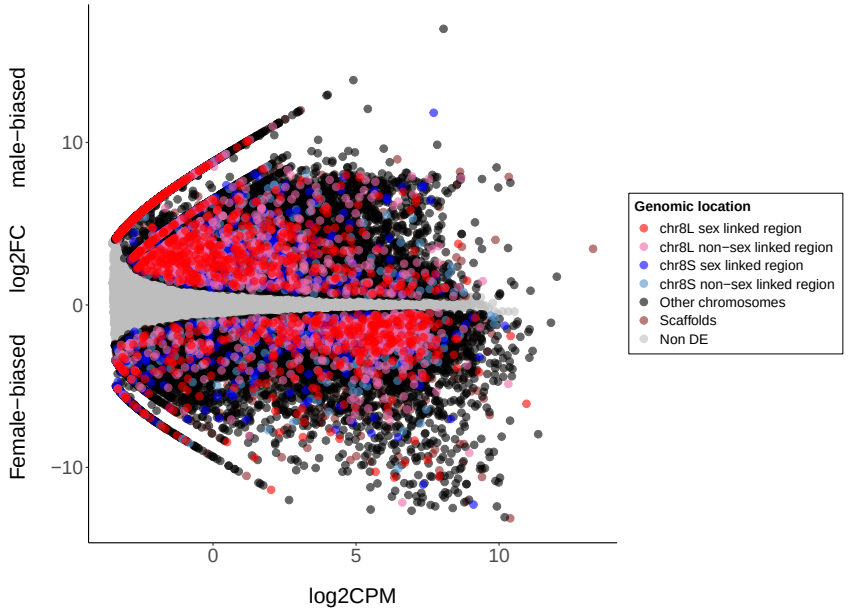

Figure S6: M-A plot of differentially expressed genes in adult gonad (edgeR, FDR $\leq$0.05).

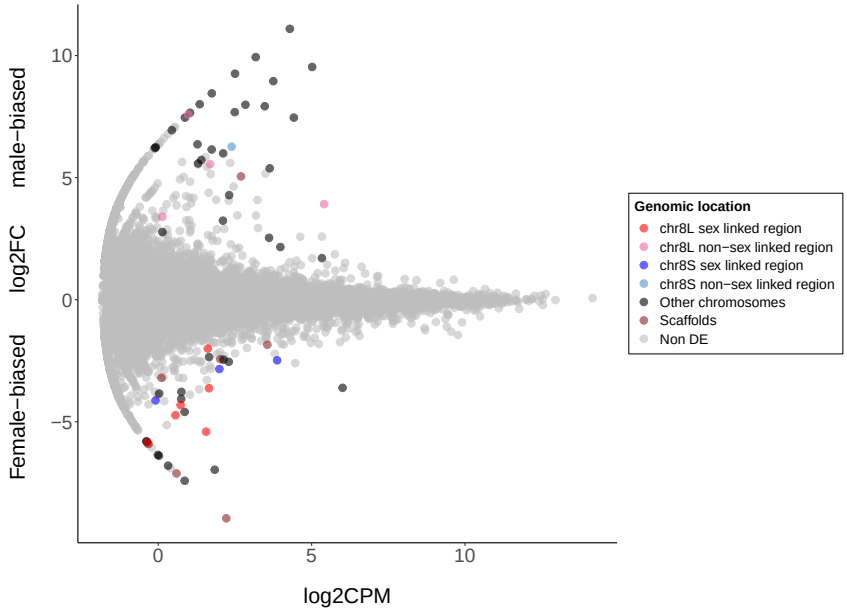

Figure S7: M-A plot of differentially expressed genes in tadpole stage46 gonad/mesenephros (edgeR, FDR $\leq$0.05).

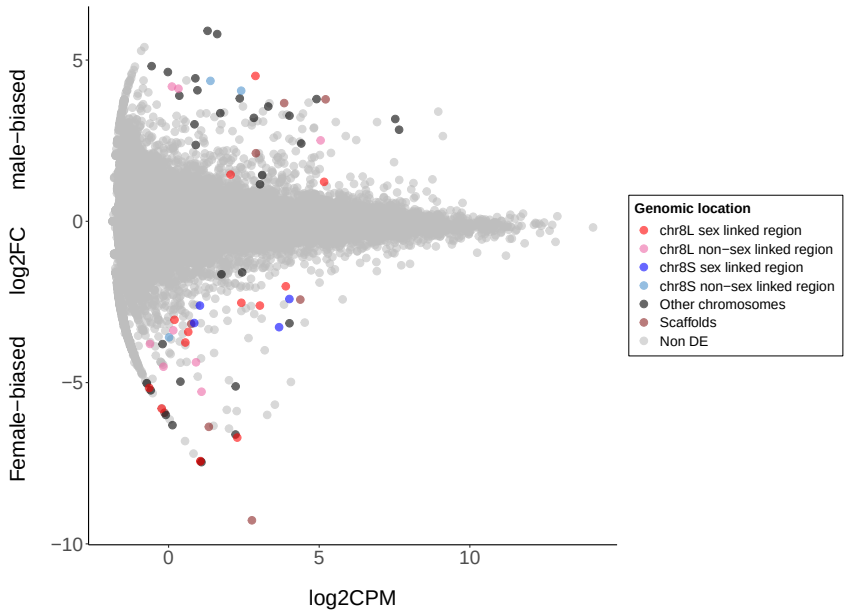

Figure S8: M-A plot of differentially expressed genes in tadpole stage48 gonad/mesenephros (edgeR, FDR $\leq$0.05).

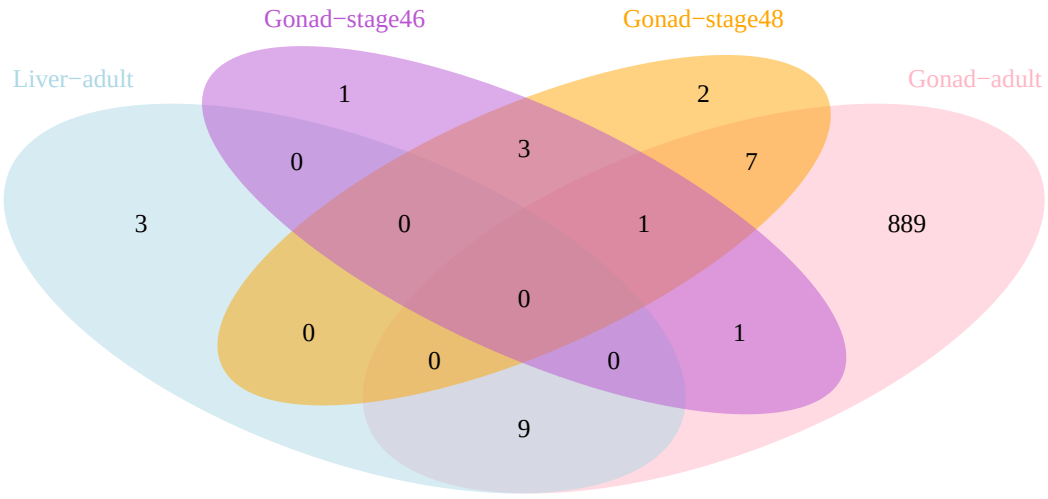

Figure S9: Venn Diagram of the number of sex-biased genes in each tissue type and developmental stage (edgeR, FDR $\leq$0.05).

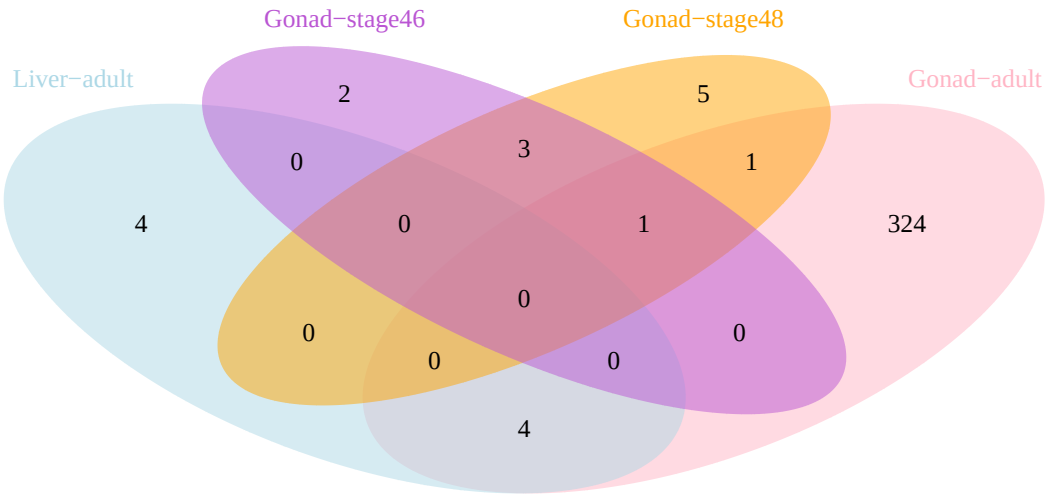

Figure S10: Venn Diagram of the number of female-biased genes in each tissue type and developmental stage (edgeR, FDR $\leq$0.05).

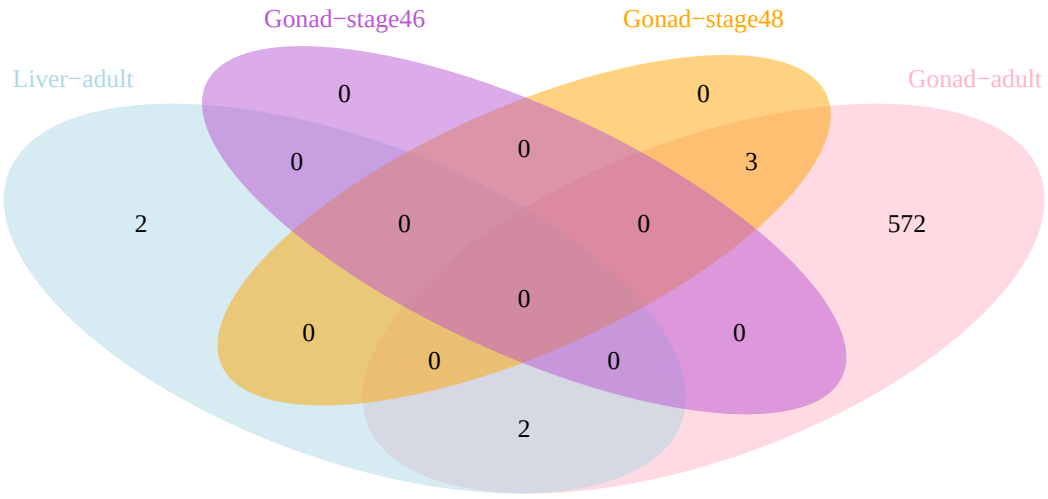

Figure S11: Venn Diagram of the number of male-biased genes in each tissue type and developmental stage (edgeR, FDR $\leq$0.05).

Supplementary Tables

Table S1: Differentially expressed genes in *X. borealis* adult liver (edgeR). Chr8L was divided into the sex-linked (SL) region and non-sex-linked (NSL) region; the same division was done for its homeologous region on chr8S.

Regions	Total gene	DE count	DE proportion	Male biased count	Female biased count
chr1L	4,821	12	$2.49 \cdot 10^{-3}$	7	5
chr1S	3,376	8	$2.37 \cdot 10^{-3}$	6	2
chr2L	4,013	9	$2.24 \cdot 10^{-3}$	4	5
chr2S	3,356	18	$5.36 \cdot 10^{-3}$	12	6
chr3L	3,375	7	$2.07 \cdot 10^{-3}$	5	2
chr3S	2,500	5	$2 \cdot 10^{-3}$	3	2
chr4L	3,335	17	$5.1 \cdot 10^{-3}$	5	12
chr4S	2,545	8	$3.14 \cdot 10^{-3}$	3	5
chr5L	3,365	12	$3.57 \cdot 10^{-3}$	4	8
chr5S	2,473	9	$3.64 \cdot 10^{-3}$	5	4
chr6L	2,998	8	$2.67 \cdot 10^{-3}$	6	2
chr6S	2,344	7	$2.99 \cdot 10^{-3}$	3	4
chr7L	2,963	12	$4.05 \cdot 10^{-3}$	11	1
chr7S	1,905	8	$4.2 \cdot 10^{-3}$	3	5
chr8L-NSL	2,074	5	$2.41 \cdot 10^{-3}$	0	5
chr8L-SL	1,267	12	$9.47 \cdot 10^{-3}$	4	8
chr8S-NSL	1,178	3	$2.55 \cdot 10^{-3}$	1	2
chr8S-SL	1,279	7	$5.47 \cdot 10^{-3}$	1	6
chr9_10L	3,136	11	$3.51 \cdot 10^{-3}$	11	0
chr9_10S	2,599	8	$3.08 \cdot 10^{-3}$	5	3
Scaffolds	5,251	11	$2.09 \cdot 10^{-3}$	5	6

Table S2: Differentially expressed genes in *X. borealis* adult gonad (edgeR). Labels correspond to Table S1

Regions	Total gene	DE count	DE proportion	Male biased count	Female biased count
chr1L	5,427	3,183	0.59	2,161	1,022
chr1S	3,919	2,197	0.56	1,565	632
chr2L	4,573	2,803	0.61	1,865	938
chr2S	3,818	2,155	0.56	1,445	710
chr3L	3,882	2,408	0.62	1,596	812
chr3S	2,772	1,570	0.57	987	583
chr4L	3,729	2,272	0.61	1,459	813
chr4S	2,914	1,715	0.59	1,151	564
chr5L	3,837	2,231	0.58	1,561	670
chr5S	2,823	1,560	0.55	1,055	505
chr6L	3,462	2,037	0.59	1,382	655
chr6S	2,699	1,515	0.56	1,018	497
chr7L	3,411	1,985	0.58	1,342	643
chr7S	2,181	1,250	0.57	817	433
chr8L-NSL	2,238	1,354	0.61	926	428
chr8L-SL	1,428	907	0.64	577	330
chr8S-NSL	1,275	724	0.57	466	258
chr8S-SL	1,413	805	0.57	543	262
chr9_10L	3,516	2,142	0.61	1,382	760
chr9_10S	3,054	1,790	0.59	1,145	645
Scaffolds	5,980	3,350	0.56	2,427	923

Table S3: Differentially expressed genes in *X. borealis* tadpole stage 46 gonad/mesenephros (edgeR). Labels correspond to Table S1

Regions	Total gene	DE count	DE proportion	Male biased count	Female biased count	6
chr1L	4,036	6	$1.49 \cdot 10^{-3}$	4	2	
chr1S	2,877	4	$1.39 \cdot 10^{-3}$	4	0	
chr2L	3,370	4	$1.19 \cdot 10^{-3}$	3	1	
chr2S	2,840	2	$7 \cdot 10^{-4}$	1	1	
chr3L	2,830	3	$1.06 \cdot 10^{-3}$	1	2	
chr3S	2,161	0	0	0	0	
chr4L	2,954	4	$1.35 \cdot 10^{-3}$	3	1	
chr4S	2,188	2	$9.1 \cdot 10^{-4}$	1	1	
chr5L	2,844	3	$1.05 \cdot 10^{-3}$	3	0	
chr5S	2,122	3	$1.41 \cdot 10^{-3}$	1	2	
chr6L	2,620	2	$7.6 \cdot 10^{-4}$	2	0	
chr6S	2,046	3	$1.47 \cdot 10^{-3}$	1	2	
chr7L	2,548	1	$3.9 \cdot 10^{-4}$	1	0	
chr7S	1,619	2	$1.24 \cdot 10^{-3}$	1	1	
chr8L-NSL	1,729	4	$2.31 \cdot 10^{-3}$	4	0	
chr8L-SL	1,118	6	$5.37 \cdot 10^{-3}$	0	6	
chr8S-NSL	1,035	1	$9.7 \cdot 10^{-4}$	1	0	
chr8S-SL	1,080	3	$2.78 \cdot 10^{-3}$	0	3	
chr9_10L	2,731	2	$7.3 \cdot 10^{-4}$	1	1	
chr9_10S	2,283	1	$4.4 \cdot 10^{-4}$	1	0	
Scaffolds	4,704	8	$1.7 \cdot 10^{-3}$	1	7	

Table S4: Differentially expressed genes in X. borealis tadpole stage 48 gonad/mesenephros (edgeR). Labels correspond to Table S1

Regions	Total gene	DE count	DE proportion	Male biased count	Female biased count
chr1L	4,036	3	$7.4 \cdot 10^{-4}$	3	0
chr1S	2,877	1	$3.5 \cdot 10^{-4}$	1	0
chr2L	3,370	1	$3 \cdot 10^{-4}$	0	1
chr2S	2,840	0	0	0	0
chr3L	2,830	3	$1.06 \cdot 10^{-3}$	2	1
chr3S	2,161	2	$9.3 \cdot 10^{-4}$	1	1
chr4L	2,954	2	$6.8 \cdot 10^{-4}$	2	0
chr4S	2,188	3	$1.37 \cdot 10^{-3}$	1	2
chr5L	2,844	1	$3.5 \cdot 10^{-4}$	1	0
chr5S	2,122	3	$1.41 \cdot 10^{-3}$	0	3
chr6L	2,620	2	$7.6 \cdot 10^{-4}$	2	0
chr6S	2,046	6	$2.93 \cdot 10^{-3}$	4	2
chr7L	2,548	2	$7.8 \cdot 10^{-4}$	1	1
chr7S	1,619	0	0	0	0
chr8L-NSL	1,729	8	$4.63 \cdot 10^{-3}$	3	5
chr8L-SL	1,118	13	$1.16 \cdot 10^{-2}$	3	10
chr8S-NSL	1,035	3	$2.9 \cdot 10^{-3}$	2	1
chr8S-SL	1,080	4	$3.7 \cdot 10^{-3}$	0	4
chr9_10L	2,731	2	$7.3 \cdot 10^{-4}$	1	1
chr9_10S	2,283	1	$4.4 \cdot 10^{-4}$	1	0
Scaffolds	4,704	9	$1.91 \cdot 10^{-3}$	3	6

Bibliography

Andersson, M. B. (1994). *Sexual selection*. Princeton University Press.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Arunkumar K. P.and Mita, K. and Nagaraju, J. (2009). The silkworm Z chromosome is enriched in testis-specific genes. *Genetics* 182(2), 493–501. ISSN: 00166731.

Assis, R., Zhou, Q., and Bachtrog, D. (2012). Sex-Biased Transcriptome Evolution in Drosophila. *Genome Biology and Evolution* 4(11), 1189–1200. ISSN: 1759-6653.

Baines, J. F., Sawyer, S. A., Hartl, D. L., and Parsch, J. (2008). Effects of X-Linkage and Sex-Biased Gene Expression on the Rate of Adaptive Protein Evolution in Drosophila. *Molecular Biology and Evolution* 25(8), 1639–1650. ISSN: 0737-4038.

Bergero, R. and Charlesworth, D. (2009). The evolution of restricted recombination in sex chromosomes. *Trends in Ecology Evolution* 24(2), 94–102. ISSN: 0169-5347.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15), 2114–2120. ISSN: 1460-2059.

Bray, N. L., Pimentel Haroldand Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* 34(5), 525–527. ISSN: 1087-0156.

Bull, J. J. et al. (1983). *Evolution of sex determining mechanisms.* The Benjamin/Cummings Publishing Company, Inc.

Butler, M. A. and King, A. A. (2004). Phylogenetic Comparative Analysis: A Modeling Approach for Adaptive Evolution. *Source: The American Naturalist* 1647064(6), 683–695.

Charlesworth, B., Coyne, J. A., and Barton, N. H. (1987). The Relative Rates of Evolution of Sex Chromosomes and Autosomes. *The American Naturalist* 130(1), 113–146.

Charlesworth, B., Jordan, C. Y., and Charlesworth, D. (2014). The evolutionary dynamics of sexually antagonistic mutations in pseudoautosomal regions of sex chromosomes. *Evolution* 68(5), 1339–1350. ISSN: 15585646.

Charlesworth, D., Charlesworth, B., and Marais, G. (2005). Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 95(2), 118–128. ISSN: 0018-067X.

Cheng, C. and Kirkpatrick, M. (2016). Sex-Specific Selection and Sex-Biased Gene Expression in Humans and Flies. *PLOS Genetics* 12(9). Ed. by M. W. Nachman, e1006170. ISSN: 1553-7404.

Chippindale, A. K., Gibson, J. R., and Rice, W. R. (2001). Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in Drosophila. *Proceedings of the National Academy of Sciences* 98(4), 1671–1675. ISSN: 00278424.

Clayton, S., Ginger, C., Qianxing, M., Marina, V., and Adam, J. (2009). A microarray analysis of sex- and gonad-biased gene expression in the zebrafish: Evidence for masculinization of the transcriptome.

Collyer, M. L. and Adams, D. C. (2018). RRPP: An r package for fitting linear models to high-dimensional data using residual randomization. *Methods in Ecology and Evolution* 9(7). Ed. by R. Freckleton, 1772–1779. ISSN: 2041210X.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X., and al, et (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology* 17(1), 13. ISSN: 1474-760X.

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(12), 3417–3442. ISSN: 00143820.

Connallon, T. and Knowles, L. L. (2005). Intergenomic conflict revealed by patterns of sex-biased gene expression. *Trends in Genetics* 21(9), 495–499. ISSN: 01689525.

Darolti, I., Wright, A. E., Pucholt, P., Berlin, S., and Mank, J. E. (2018). Slow evolution of sex-biased genes in the reproductive tissue of the dioecious plant Salix viminalis. *Molecular ecology* 27(3), 694–708. ISSN: 1365-294X.

Darwin, C. (1872). *The descent of man and selection in relation to sex.* Vol. 1.

Drummond, D. A., Bloom, J. D., Adami, C., Wilke, C. O., and Arnold, F. H. (2005). Why highly expressed proteins evolve slowly. *PNAS* 102(40), 14338–14343.

Ellegren, H., Hultin-Rosenberg, L., Brunström Björnand Dencker, L., Kultima, K., and Scholz, B. (2007). Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. *BMC Biology* 5(1), 40. ISSN: 1741-7007.

Ellegren, H. and Parsch, J. (2007). The evolution of sex-biased genes and sex-biased gene expression. *Nature Reviews Genetics* 8(9), 689–698. ISSN: 1471-0056.

Emlen, D. J. (2008). The Evolution of Animal Weapons. *Annual Review of Ecology, Evolution, and Systematics* 39(1), 387–413. ISSN: 1543-592X.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection.* Clarendon Press.

Furman, B. L. S. and Evans, B. J. (2016). Sequential Turnovers of Sex Chromosomes in African Clawed Frogs (Xenopus) Suggest Some Genomic Regions Are Good at Sex Determination.

Furman, B. L. S. and Evans, B. J. (2018). Divergent Evolutionary Trajectories of Two Young, Homomorphic, and Closely Related Sex Chromosome Systems. *Genome Biology and Evolution* 10(3), 742–755.

Gelbart, M. E. and Kuroda, M. I. (2009). Drosophila dosage compensation: A complex voyage to the X chromosome. *Development* 136(9), 1399–1410. ISSN: 09501991.

Gibson, J. R., Chippindale, A. K., and Rice, W. R. (2002). The X chromosome is a hot spot for sexually antagonistic fitness variation. *Proc. R. Soc. Lond. B* 269, 499–505.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., and al., et (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29(7), 644–652. ISSN: 1087-0156.

Grath, S. and Parsch, J. (2016). Sex-Biased Gene Expression. *Annu. Rev. Genet* 50, 29–44.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., and al., et (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols* 8(8), 1494–512. ISSN: 1750-2799.

Haldane, J. B. S. (1927). *A mathematical theory of natural and artificial selection, part V: selection and mutation.* Vol. 23. 7. Cambridge University Press.

Hamilton, W. and Zuk, M. (1982). Heritable True Fitness and Bright Birds: A Role for Parasites? *Science* 218(4570), 384–387.

Harrison, P. W., Wright, A. E., Zimmer Fabianand Dean, R., Montgomery, S. H., Pointer, M. A., and Mank, J. E. (2015). Sexual selection drives evolution and rapid turnover of male gene expression. *Proceedings of the National Academy of Sciences* 112(14). ISSN: 0027-8424.

Hvilsom, C., Qian, Y., Bataillon, T., Li, Y., Mailund, T., Sallé, B., Carlsen, F., Li, R., Zheng, H., Jiang, T., and al., et (2012). Extensive X-linked adaptive evolution in central chimpanzees. 109(6), 2054–2059.

Ingleby, F. C., Flis, I., and Morrow, E. H. (2015). Sex-Biased Gene Expression and Sexual Conflict throughout Development. *Cold Spring Harb Perspect Biol* 7.

Itoh, Y., Melamed, E., Yang, X., Kampf, K., Wang, S. a., Yehya, N., VanNas, A., Replogle, K., Band, M. R., Clayton, D. F., and al., et (2007). Dosage compensation is less effective in birds than in mammals. *Journal of Biology* 6(1), 2. ISSN: 14754924.

Karimi, K., Fortriede, J. D., Lotay, V. S., Burns, K. A., Wang, D. Z., Fisher, M. E., Pells, T. J., James-Zorn, C., Wang, Y., Ponferrada, V. G., and al., et (2018). Xenbase: a genomic, epigenomic and transcriptomic model organism database. *Nucleic Acids Research* 46(D1), D861–D868. ISSN: 0305-1048.

Kayserili, M. A., Gerrard, D. T., Tomancak, P., and Kalinka, A. T. (2012). An Excess of Gene Expression Divergence on the X Chromosome in Drosophila Embryos: Implications for the Faster-X Hypothesis. *PLoS Genetics* 8(12). Ed. by A. Kopp, e1003200. ISSN: 1553-7404.

Khil, P. P., Smirnova, N. A., Romanienko, P. J., and Camerini-Otero, R. D. (2004). The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nature Genetics* 36(6), 642–646. ISSN: 1061-4036.

Leder, E. H., Cano, J. M., Leinonen, T., O'Hara, R. B., Nikinmaa, M., Primmer, C. R., and Merilä, J. (2010). Female-Biased Expression on the X Chromosome as a Key Step in Sex Chromosome Evolution in Threespine Sticklebacks. *Molecular Biology and Evolution* 27(7). 1,268 genes that were differentially expressed between sexes, a highly significant proportion (23%) was concentrated on chromosome 19, corresponding to the recently described nascent sex chromosomes. The, 1495–1503. ISSN: 1537-1719.

Lenth, R. (2019). *emmeans: Estimated Marginal Means, aka Least-Squares Means.* R package version 1.4.2.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15(12), 550. ISSN: 1474-760X.

Lucotte, E. A., Laurent, R., Heyer, E., Ségurel, L., and Toupance, B. (2016). Detection of allelic frequency differences between the sexes in humans: A signature of sexually antagonistic selection. *Genome Biology and Evolution* 8(5), 1489–1500. ISSN: 17596653.

Ma, W.-J., Veltsos, P., Sermier, R., Parker, D. J., and Perrin, N. (2018a). Evolutionary and developmental dynamics of sex-biased gene expression in common frogs with proto-Y chromosomes. *Genome Biology* 19(1), 156. ISSN: 1474-760X.

Ma, W.-J., Veltsos, P., Toups, M. A., Rodrigues, N., Sermier, R., Jeffries, D. L., and Perrin, N. (2018b). Tissue Specificity and Dynamics of Sex-Biased Gene Expression in a Common Frog Population with Differentiated, Yet Homomorphic, Sex Chromosomes. *Genes* 9(6). ISSN: 2073-4425.

Mank, J. E., Nam, K., Brunstrom, B., and Ellegren, H. (2010). Ontogenetic Complexity of Sexual Dimorphism and Sex-Specific Selection. *Molecular Biology and Evolution* 27(7), 1570–1578. ISSN: 0737-4038.

Mank, J. E. (2009). Sex chromosomes and the evolution of sexual dimorphism: Lessons from the genome. *American Naturalist* 173(2), 141–150. ISSN: 00030147.

Mank, J. E. and Ellegren, H. (2009). Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds. *Evolution* 63(6), 1464–1472. ISSN: 00143820.

Mank, J. E., Hultin-Rosenberg, L., Webster, M. T., and Ellegren, H. (2008). The unique genomic properties of sex-biased genes: Insights from avian microarray data. *BMC Genomics* 9(1), 148. ISSN: 1471-2164.

Meisel, R. P., Malone, J. H., and Clark, A. G. (2012). Faster-X Evolution of Gene Expression in Drosophila. *PLoS Genetics* 8(10). Ed. by D. Bachtrog, e1003013. ISSN: 1553-7404.

Naurin, S., Hasselquist, D., Bensch, S., and Hansson, B. (2012). Sex-Biased Gene Expression on the Avian Z Chromosome: Highly Expressed Genes Show Higher Male-Biased Expression. *PLoS ONE* 7(10). Ed. by D. I. Nurminsky, e46854. ISSN: 1932-6203.

O'meara, B. C., Ané, C. C., Sanderson, M. J., and Wainwright, P. C. (2006). *TESTING FOR DIFFERENT RATES OF CONTINUOUS TRAIT EVOLUTION USING LIKELIHOOD*. Vol. 60. 5.

Orr, H. A. (2003). The Distribution of Fitness Effects Among Beneficial Mutations. *Genetics* 163, 1519–1526.

Otto, S. P., Pannell, J. R., Peichel, C. L., Ashman, T.-L., Charlesworth, D., Chippindale, A. K., Delph, L. F., Guerrero, R. F., Scarpino, S., and Mcallister, B. F. (2011). About PAR: The distinct evolutionary dynamics of the pseudoautosomal region. *Trends in Genetics* 27(9).

Paris, M., Villalta, J. E., Eisen, M. B., and Lott, S. E. (2015). Sex Bias and Maternal Contribution to Gene Expression Divergence in Drosophila Blastoderm Embryos. *PLOS Genetics* 11(10). Ed. by A. G. Clark, e1005592. ISSN: 1553-7404.

Parisi, M., Nuttall, R., Edwards, P., Minor, J., Naiman, D., Lü, J., Doctolero, M., Vainer, M., Chan, C., and Malley, J. (2004). A survey of ovary-, testis-, and soma-biased gene expression in Drosophila melanogaster adults. *Genome Biology* 5(6), R40. ISSN: 14656906.

Parker, G. (1979). Sexual selection and sexual conflict. In: *In Sexual Selection and Reproductive Competition in Insects*. Ed. by M. S. Blum and N. A. Blum. Academic Press, 123–166.

Parsch, J. and Ellegren, H. (2013). The evolutionary causes and consequences of sex-biased gene expression. *Nature Reviews Genetics* 14(2), 83–87. ISSN: 1471-0056.

Patwa, Z. and Wahl, L. M. (2008). The fixation probability of beneficial mutations. *Journal of the Royal Society, Interface* 5(28), 1279–89. ISSN: 1742-5689.

Petrie, M. (1994). Improved growth and survival of offspring of peacocks with more elaborate trains. *Nature* 371(6498), 598–599. ISSN: 0028-0836.

Ponnikas, S., Sigeman, H., Abbott, J. K., and Hansson, B. (2018). Why Do Sex Chromosomes Stop Recombining? *Trends in Genetics* 34(7), 492–503.

Prum, R. (1990). Phylogenetic Analysis of the Evolution of Display Behavior in the Neotropical Manakins (Aves: Pipridae). *Ethology* 84, 202–231.

Quinlan, A. R. and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6), 841–842. ISSN: 1460-2059.

Ranz, J. M., Castillo-Davis, C. I., Meiklejohn, C. D., and Hartl, D. L. (2003). Sex-dependent gene expression and evolution of the Drosophila transcriptome. *Science* 300(5626), 1742–1745. ISSN: 00368075.

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3(2), 217–223. ISSN: 2041210X.

Rice, W. R. and Chippindale, A. K. (2001). Intersexual ontogenetic conflict. *Journal of Evolutionary Biology* 14(5), 685–693. ISSN: 1010061X.

Rice, W. R. (1984). SEX CHROMOSOMES AND THE EVOLUTION OF SEXUAL DIMORPHISM. *Evolution* 38(4), 735–742. ISSN: 00143820.

Rice, W. R. (1987). *THE ACCUMULATION OF SEXUALLY ANTAGONISTIC GENES AS A SELECTIVE AGENT PROMOTING THE EVOLUTION OF REDUCED RECOMBINATION BETWEEN PRIMITIVE SEX CHROMOSOMES; THE ACCUMULATION OF SEXUALLY ANTAGONISTIC GENES AS A SELECTIVE AGENT PROMOTING THE EVOLUTION OF REDUCE*. Vol. 41. 4.

Rice, W. R. (1992). Sexually Antagonistic Genes: Experimental Evidence. *Science* 256(5062), 1436–1439.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1), 139–140. ISSN: 1367-4803.

Schurch, N. J., Schofield, P., Gierliński, M., Cole, C., Sherstnev, A., Singh, V., Wrobel, N., Gharbi, K., Simpson, G. G., Owen-Hughes, T., and al., et (2016). How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA* 22(6), 839–851. ISSN: 1355-8382.

Seger, J. and Trivers, R. (1986). Asymmetry in the evolution of female mating preferences. *Nature* 319(6056), 771–773.

Session, A. M., Uno, Y., Kwon, T., Chapman, J. A., Toyoda, A., Takahashi, S., Fukui Akimasaand Hikosaka, A., Suzuki, A., Kondo, M., and al., et (2016). Genome evolution in the allotetraploid frog Xenopus laevis. *Nature* 538(7625), 336–343. ISSN: 0028-0836.

Storchová, R. and Divina, P. (2006). Nonrandom representation of sex-biased genes on chicken Z chromosome. *Journal of Molecular Evolution* 63(5), 676–681. ISSN: 00222844.

Straub, T. and Becker, P. B. (2007). Dosage compensation: The beginning and end of generalization. *Nature Reviews Genetics* 8(1), 47–57. ISSN: 14710056.

Tymowska, J. (1991). Polyploidy and cytogenetic variation in frogs of the genus Xenopus. In: *Amphibian Cytogenetics and Evolution*. Ed. by D. M. Green and S. K. Sessions. Academic Press San Diego, 259–297.

Vicoso, B. and Bachtrog, D. (2015). Numerous Transitions of Sex Chromosomes in Diptera. *PLOS Biology* 13(4). Ed. by H. S. Malik, e1002078. ISSN: 1545-7885.

Vicoso, B. and Charlesworth, B. (2006). Evolution on the X chromosome: unusual patterns and processes. *NATURE REVIEWS | GENETICS* 7, 645–653.

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? *Journal of Molecular Evolution* 68(5), 576–583. ISSN: 0022-2844.

Vicoso, B., Kaiser, V. B., and Bachtrog, D. (2013). Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. *PNAS* 110(16), 6453–6458.

Wu, T. D. and Watanabe, C. K. (2005). GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* 21(9), 1859–1875. ISSN: 1367-4803.

Yoshimoto, S., Okada, E., Umemoto, H., Tamura, K., Uno, Y., Nishida-Umehara, C., Matsuda, Y., Takamatsu, N., Shiba, T., and Ito, M. (2008). *A W-linked DM-domain gene, DM-W, participates in primary ovary development in Xenopus laevis DEVELOPMENTAL BIOLOGY*. Vol. 105. 7.

Zha, X., Xia, Q., Duan, J., Wang, C., He, N., and Xiang, Z. (2009). Dosage analysis of Z chromosome genes using microarray in silkworm, Bombyx mori. *Insect Biochemistry and Molecular Biology* 39(5–6), 315–321. ISSN: 0965-1748.

Zhou, Y., Zhu, J., Tong, T., Wang, J., Lin, B., and Zhang, J. (2019). A statistical normalization method and differential expression analysis for RNA-seq data between different species. *BMC Bioinformatics* 20(1), 1–10. ISSN: 14712105.

Zhu, A., Ibrahim, J. G., and Love, M. I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* 35(12). Ed. by O. Stegle, 2084–2092. ISSN: 1367-4803.