Stony Brook University

Phylogenomic Analysis of Sex-Dependent Gene Expression in the *Drosophila* Genus to Assess Alternative Models of Evolution in Relation to Sex

A Dissertation Presented

by

Nicole M. Lashbrook

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Ecology and Evolution

Stony Brook University

December 2011

Stony Brook University

The Graduate School

Nicole M. Lashbrook

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

Walter F. Eanes - Advisor

Professor, Department of Ecology and Evolution

John R. True - Chairperson of Defense

Associate Professor, Department of Ecology and Evolution

Joshua Rest

Assistant Professor, Department of Ecology and Evolution

Dannie Durand

Associate Professor, Departments of Biological Sciences and Computer Science,

Carnegie-Mellon University

This dissertation is accepted by the Graduate School

Lawrence Martin

Dean of the Graduate School

Abstract of the Dissertation

Phylogenomic Analysis of Sex-Dependent Gene Expression in the *Drosophila* Genus to Assess Alternative Models of Evolution in Relation to Sex

by

Nicole M. Lashbrook

Doctor of Philosophy

in

Ecology and Evolution

Stony Brook University

2011

When Charles Darwin formulated his theory of sexual selection in *The Descent of Man, and Selection in Relation to Sex,* sex differences were understood largely in terms of morphology, physiology and behavior. Recent genomic studies have demonstrated that sex differences are apparent at the molecular level. In the *Drosophila* study system, 'faster-male' evolution is widespread, as is indicated by enhanced rates of amino acid substitution for male-biased as compared to nonsex- and female-biased genes. While the role of intrasexual male competition has been highlighted in the interpretation of these observations, the alternative hypothesis that historical sex differences may account for these patterns has not been addressed. To address the alternative hypothesis that relative constraints account for 'faster-male' evolution in *Drosophila*, multivariate matching techniques were used. Male and female-biased genes were matched with nonsex-biased genes on several sets of proposed confounders and male and female effects were estimated. Adjustment for confounding due to historical sex differences resulted in a significant decrease in estimates of male effects on rates of molecular evolution and marginal increase or no difference in female effect estimates. After adjustment, male and female effects on rates of molecular evolution were similar suggesting that 'faster-male' evolution may be largely due to sex differences in the efficacy of selection given historical constraints, rather than to contemporary selection.

Multi-species data derived from the application of high-throughput technology on multiple closely related species of *Drosophila* permits the modeling of traits on a phylogeny on a genome-wide scale. This framework is used to study the role of intersexual coevolution in affecting genome evolution. Ancestral inference on sex-dependent gene expression is used to empirically test theoretical expectations about how sex-biased gene expression should arise, to identify candidate genes involved in ongoing intersexual developmental conflict, and to study the role of expression evolution on the 'demasculinization' of X chromosomes. Functional analysis of candidate genes suggests that enhanced selection on male secondary sexual traits may be a common source of sexually antagonistic developmental conflict in *Drosophila*. Modeling of transition dynamics among states of sex-biased gene expression for X-linked and autosomal-linked genes indicates that expression evolution can account for the underrepresentation of male-biased genes on the X chromosome. Analysis of transition dynamics for candidate genes provides support for the hypothesis that the 'demasculinization' of X chromosomes in *Drosophila* is partially due to the effects of sexually antagonistic coevolution on the evolution of gene expression.

To Kenneth, my love

To my mother and father

An die Pferde der Hohen Schule:

Mögen Eure Sprünge aus der Freude entstehe!

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Appendix B

Appendix D

List of Tables

Chapter 2

Chapter 3

Appendix B

Acknowledgments

I wish to acknowledge the *Drosophila* community in general for prioritizing full and open access to high quality datasets and analytical tools. Special thanks to A. Larracuente for providing a lion's share of pre-computed data for use in this dissertation. I am grateful also to my committee members, especially to my doctoral advisor, Dr. Walter F. Eanes, for demonstrating steady support even when I wavered.

Chapter 1

An Introduction to the Thesis: Phylogenomic Analysis of Sex-Dependent Gene Expression in the *Drosophila* genus to Assess Alternative Models of Evolution in Relation to Sex

GENERAL INTRODUCTION

This dissertation has two aims. The first is to investigate the evolutionary processes that give rise to sex differences. The second is to advance analytic methods for the study of sex differences in the area of evolutionary genomics. No originally sourced data were analyzed in completion of this dissertation; instead, the focus has been to develop generalizable methods for ready use in the highly developed *Drosophila* system, which may be later used in other study systems as relevant datasets become available. The *Drosophila* study system provides a rich source of functional data derived from decades of research on the *melanogaster* species. More recently the study system has been expanded to include multiple species within the genus, making the system also ideal for the study of comparative biology.

Multi-species data derived from the application of high-throughput technology on multiple closely related species of *Drosophila* permits the modeling of traits on a phylogeny on a genome-wide scale. In this dissertation, a multi-species microarray analysis of sex-biased gene expression for 7 species of *Drosophila* is studied in conjunction with data on spatio-temporal expression and sequence evolution to address several outstanding questions about the evolution of sexual dimorphism in *Drosophila*. Data on sex-dependent gene expression is first analyzed in a multivariate context using the data sources available for *D. melanogaster*. The first aim is to address questions regarding the role of historical sex differences in contributing to 'faster-male' evolution. In this context, historical sex differences refer to differences in the kinds of genes that have become sex-biased over evolutionary time. As it turns out, the genes that are differentially expressed between the sexes are dramatically different on several variables that may affect rates of evolution.

In the third chapter, this same dataset is used as the basis for a phylogenomic analysis of sex-dependent gene expression, where the evolution of sex-biased gene expression is modeled on the *Drosophila* phylogeny for thousands of orthologous gene families. This framework is used to address questions about how sex-biased gene expression typically arises, to identify candidate genes involved in long-term intra-locus sexual conflicts, and to study the evolution of sex-biased gene expression as a dynamic process.

The development of methods for dealing with the problem of *multiple causes* is a central concern of this dissertation. This problem arises first in the study of 'faster-male' evolution, where the effects of sex-specific environmental exposure on gene evolution are estimated after adjusting for confounding due to historical factors. The problem of confounding factors is treated by employing methods from the field of observational studies. The issue of multiple causes

arises again in the second chapter where evidence is advanced in support of the view that expression evolution contributes to the 'demasculinization' of X chromosomes. 'Demasculinization' refers to a widely observed pattern in *Drosophila* whereby male-biased genes are underrepresented on the X chromosome. The modeling of expression evolution as a dynamic process permits estimation of the relative contributions of various subsets of genes to the 'demasculinization' phenomenon. In particular, the contribution of candidate sexually antagonistic genes to 'demasculinization' is analyzed. This same technique can be adapted to estimate contributions for other gene sets of interest for further study of this phenomenon in *Drosophila* and other study systems.

INTRODUCTION TO CHAPTER 2

Towards causal inference in evolutionary genomics: accounting for historical constraint in the study of 'faster-male' evolution

Sex-biased gene expression has been studied at a genome-wide level using microarrays where the relative amount of mRNA or cDNA transcribed by a given gene can be measured from samples of male and female tissue. Relative gene expression measures derived from microarray analysis can be used to group genes into male-biased, female-biased, or nonsex-biased classes for further investigation.

Early genome level studies in *D. melanogaster* have revealed that male-biased genes are evolving at enhanced rates and are more likely to experience positive selection than are female- and nonsex-biased genes (Zhang et al. 2007, Lemos et al. 2005a, Meiklejohn et al. 2003, Ranz et al. 2003, Zhang, Hambuch and Parsch 2004). Leading explanations relate this observation to the theory of sexual selection. Sexual selection theory predicts that males may often experience more intense intrasexual competition due to sex differences in reproductive strategy (Darwin 1871). Differences in reproductive strategy are epitomized by and may be ultimately derived from the phenomenon of anisogamy (Trivers 1972), whereby males and females produce gametes in of different size and in different amounts. Since males characteristically produce smaller and more gametes than females, they are expected to have a higher potential number of progeny. This may lead to sexual conflicts between males and females, as well as to enhanced competition among males for a relatively limited number of reproductive opportunities (Darwin 1871). Even though the effects of intersexual coevolution on molecular evolution is not well understood, asymmetries in genome-wide rates of molecular evolution are often attributed to asymmetries in the expected intensity of intrasexual competition as experienced by males and females.

An understanding of the cause (s) of the pattern of enhanced rate of molecular evolution seen for male-biased genes is complicated by the myriad of factors that are known to affect gene evolution. While it is widely acknowledged that historical sex differences may facilitate or constrain gene evolution, past analyses have not attempted to distinguish between the effects of historical sex differences and contemporary selective processes. This is partially because of the inherent difficulties in processing multiple large data-sets including all of the potentially relevant historical factors, but is also due to the lack of an appropriate framework for grappling with issues of historical confounding in evolutionary genomic analyses.

In the following chapter, an observational design framework is adopted for the study of 'faster-male' evolution in *Drosophila*. Specifically, a multivariate matching technique involving subclassification on a propensity score is used to compare groups of male or female-biased genes with nonsex-biased genes in order to estimate male and female effects on rates of molecular evolution. While the implementation of the propensity score methods will be described in detail

in the methods section of the subsequent chapter, the motivations for using this approach are presented here.

A natural experiment on a phylogeny

The methods adopted for use in the following chapter extend from a point of view that takes the randomized experiment as the ideal basis for scientific inference. The quasi–experimental approach to causality, which has the insight of *randomization* at its core, extends from the work of R.A. Fisher (Rosenbaum 2002). The quasi-experimental approach has been extensively developed for applications in medical and political sciences, where researchers are often limited in their ability to conduct controlled experiments. In evolutionary biology, phylogenetic comparative methodologies are often touted as a framework for making inferences about adaptive processes from natural experiments, yet the issue of multiple causes is an ongoing challenge. Increasingly, data are available for dealing with multiple causes, but the issue must also be addressed through advances in methodology. Advances in quasi-experimental methodological approaches that have occurred in other disciplines can be utilized to advance research in this area, where application of the propensity score model to the study of 'faster-male' evolution is a case in point.

According to the potential outcomes model, an individual causal effect is the difference between potential outcomes for any given unit under study (Rubin 2005). In studying the effects of exposure to a male or female environment on molecular evolution, the study unit is a single gene. Potential outcomes correspond to response measures under exposure to various treatment regimes. In the case of controlled randomized experiment, these regimes include a 'treatment' (exposed) and a 'control' (not exposed). The 'treatment' may correspond to any of a variety of experimental manipulation whose causal effect is of interest. In the second chapter, the treatments correspond to exposure to a male or female environment for which sex-biased gene expression is proxy. In any given experiment, although there may be multiple treatment regimes, only one outcome is ever observable. Causal inference methodologies make explicit the assumptions inherent in moving from the interpretation of cause as a difference between potential outcomes for an individual data unit to that in which a causal effect is estimated for a population on the basis of the mean difference between a 'treatment' and a 'control' group. A key assumption is that the 'treatment assignment' (that is, the assignment into a control or treatment group) of the individuals under study should be random with respect to potential outcomes. While individual study units may in fact differ in their association with other causes, if a treatment is randomly assigned, the effects of other potential causes should be randomly distributed in the treatment and control groups and should therefore not lead to systematic bias when a mean difference between observables is taken. In this way, the randomized experiment deals effectively with problem of multiple causes, permitting the estimation of isolated effects even where multiple causes are present.

Observational study methods are designed to deal specifically with conditions under which treatment assignments cannot be randomly assigned. Manipulation of treatment assignments is not always possible, and in these cases independence of the treatment assignment with the potential outcomes cannot be assured. In particular, biological processes of evolution may lead to associations between other causes affecting gene evolution and states of sex-biased gene expression. In the case of 'faster-male' evolution, it is proposed that historical associations between states of sex-biased gene expression may lead to misinterpretation of the relationship between gene evolution and the exposure to a male environment per se. For example, a gene

may evolve at an enhanced rate because it is free from pleiotropic constraints as a result of having limited expression breadth, rather than because of the effects of exposure to a sex-specific environment. If male-biased genes are generally more narrowly expressed, then restricted spatial expression may be the cause of 'faster-male' evolution rather than exposure to a male-specific environment per se.

Matching is a useful approach for eliminating or reducing bias due to other causes. Matching can be difficult when the number of confounders is large. In these cases a useful technique for achieving matching utilizes the 'propensity score', the conditional probability of treatment given a set of confounders (Rosenbaum and Rubin 1983). This technique reduces a multidimensional matching problem to a much easier one involving matching on a single variable. This is the approach taken in the second chapter.

Although not fully pursued in the following chapter, with accumulating multi-species data, there is the potential to interpret key methodological assumptions through the lens of common descent. Specifically, potential outcomes may be interpreted in a phylogenomic context. A schematic is depicted (**Fig 1**). This schematic assumes that the outcome variable is defined as a unit of change (for example, evolutionary rate, $\omega = dN/dS$), where the issue of phylogenetic non-independence is not an issue. In future applications, the investigation of methodological assumptions in a phylogenomic context may provide an effective means of testing model assumptions, for identifying an appropriate set of confounders, or for exploring the appropriate level of generality of inferences on a phylogeny.

Scientific findings

First, various classes of sex-biased genes are compared to nonsex-biased genes for several proposed determinants of evolutionary rate (Larracuente et al. 2008). Several covariates are analyzed including: the number of protein-protein interactors, protein length, gene expression, intron number, rate of recombination, tissue-biased and stage-biased gene expression (measures of spatial and temporal expression breadth). Male-biased genes are found to differ from nonsex-biased genes for all variables except the number of protein-protein interactors and the rate of recombination, while female-biased genes differ from nonsex-biased genes only for spatial expression breadth and the number of protein-protein interactors. To reduce bias due to proposed confounders, propensity score models are utilized to find an optimal subclassification such that the covariate distribution of proposed confounders is as similar as possible for comparisons of male-biased and female-biased genes with nonsex-biased genes. Several model criteria are considered, each of which is defined in terms of a distinct set of proposed confounders. Effect estimates with adjustment for historical confounders result in considerably reduced estimates of male effects on gene evolution and marginally increased female effects on gene evolution. This suggests that historical factors facilitate evolution in males and constrain evolution in females. Male and female effects are comparable with adjustment, suggesting that 'faster-male' evolution is largely a result of sex differences in the efficacy rather than the intensity of selection. With adjustment, differences between male effects and female effects on gene evolution are shown to be marginal or nonexistent.

INTRODUCTION TO CHAPTER 3

Phylogenomic analysis of sex-dependent gene expression to assess intra-locus sexual conflict

The objective of the third chapter of the thesis is to better understand the role of intra-

locus sexual conflict in affecting the evolution of sexual dimorphism. While the role of intrasexual male competition has been highlighted in the interpretation of genomic patterns of evolution, sexual selection theory predicts a concomitant role for intersexual coevolution, yet little progress has been made in discerning how these dynamics affect genome evolution.

Theoretical considerations suggest that sex-biased expression is likely to evolve when expression is advantageous in one sex and neutral or costly in the other (Rhen 2000). In response to the appearance of an allele with sexually antagonistic fitness effects, one plausible scenario is that selection will cause expression to be down-regulated in the disfavored sex to the point of sex-limitation, thereby eliminating further antagonistic selection pressures. However, when conflicts arise over a shared function, resolution through sex-limited expression may not be possible. As evolutionary opportunity permits, a tug-of-war may ensue over optimal expression and function as mean population fitness is maximized alternately at the expense of one sex or the other (Rice and Chippindale 2001).

To coordinately examine expression divergence in the unbiased and biased sex, a phylogenomic analysis is conducted using data from the seven-species genome-wide microarray analysis of sex-biased expression in the genus *Drosophila* (Zhang, Sturgill et al. 2007). More than 3,500 orthologous gene families were selected for phylogenomic anlaysis. Given the published species tree (Powell 1997) relating the seven *Drosophila* species and observed sex-specific expression values derived from microarray data, ancestral male and female expression values are estimated for each gene family using a Brownian model of expression evolution (Schluter et al. 1997). Ancestral male (M_i) and female (F_i) expression values for each orthologous gene family are inferred based on observed male and female expression values derived from Zhang *et al.'s* seven species microarray dataset. Sex expression divergence (ΔM_{ij}, ΔF_{ij}) along each branch is evaluated as a difference between expression values between nodes or between nodes and tips.

On the basis of ancestral inference on sex-dependent gene expression, several indices are devised to study intra-locus sexual conflict. First, a sex comparative index is introduced. The sex-comparative index is a signed difference in the magnitude of sex-specific expression divergence for male and female expression: $I = \frac{|\Delta F_{ij}| - |\Delta M_{ij}|}{\sqrt{\Delta F_{ij}{}^2 + \Delta M_{ij}{}^2}}$, where the sign of this index is indicative of the sex for which absolute expression divergence is larger; that is, if I >> 0, divergence is occurring disproportionately in the female sex, while if I << 0, divergence is occurring disproportionately in the male sex. Furthermore, states of the sex comparative index (**f, n, m**) are derived from statistical hypothesis testing on the sex comparative index. Ancestral states of sex-biased (**F, N, M**) gene expression are also assigned to each node on the basis of ancestral inferences.

The sex comparative index is used to study the evolution of sex-biased gene expression on a phylogeny, where the distribution of the sign of the sex comparative index is expected under neutral expectation to be binomially distributed on each branch. The distribution on the comparative sex index is studied for the different types of transition in sex-biased gene expression (for example, from ancestor to descendent: **FF, FN, FM, NF, NN, NM, MF, MM, MN**). Analysis based on the signs of the index indicates that the evolution of sex-limited expression most typically involves a loss of expression in the alternative sex, which is in accord with theoretical expectations. Where gene expression is not sex-biased, disproportionate divergence in the male sex is more common; also, divergence in the male sex for female-biased genes is more common than is disproportionate divergence in the female sex for male-biased

genes. These observations suggest that gene expression is more likely to be costly in the female sex.

It is proposed that enhanced lability on the sex comparative index may be associated with genes involved in sexual developmental conflict. This metric is evaluated using states of the sex comparative index (**f, n, m**) and is defined for each family of orthologs. Deviations from neutrality are used to select candidate genes for further analysis.

Both sex-biased and candidate genes are further studied using functional data derived from the *D. melanogaster* study system. Male-biased genes show restricted spatial and temporal expression patterns as compared to nonsex-biased genes. These observations apply even to genes that are not specifically or maximally expressed in reproductive tissues. Candidate genes are found to be overrepresented among genes expressed specifically at the adult stage and in the carcass tissues, and enhanced lability is associated with genes that are sometimes male-biased. Based on these observations, it is suggested that the most common source of developmental conflict may arise from the development of male secondary sexual traits.

An understanding of the processes that lead to the limited spatial and temporal expression profiles characteristic of male-biased genes is the key to understanding 'faster-male' evolution. It is observed that newly male-biased and female-biased genes come into being with different spatial profiles, but neither the proximate nor ultimate cause of this is clear. It is hypothesized that sexual developmental conflicts may play an ongoing role in driving the enhanced modularity of male-biased genes, but further research is required.

Sexually antagonistic coevolution and the X chromosome

In the second chapter, the phylogenomic framework is used to study the causes of the observed underrepresentation of male-biased genes on the X chromosome. Because the fixation rate of alleles with sexually antagonistic fitness effects is expected to differ as a function of dominance on the X chromosome as compared to the autosomes (Charlesworth, Coyne and Barton 1987, Rice 1984), it has been proposed that antagonistic coevolution may play a role in determining the distribution of sex-biased genes throughout the genome (Connallon and Knowles 2005, Oliver and Parisi 2004, Vicoso and Charlesworth 2006). While gene birth and death processes appear to play a substantial role in determining the distribution of sex-biased genes throughout the genome (Zhang et al. 2007, Bachtrog, Toda and Lockton 2010, Betran, Thornton and Long 2002), between species transitions among different categories of bias are also common (Ranz et al. 2003).

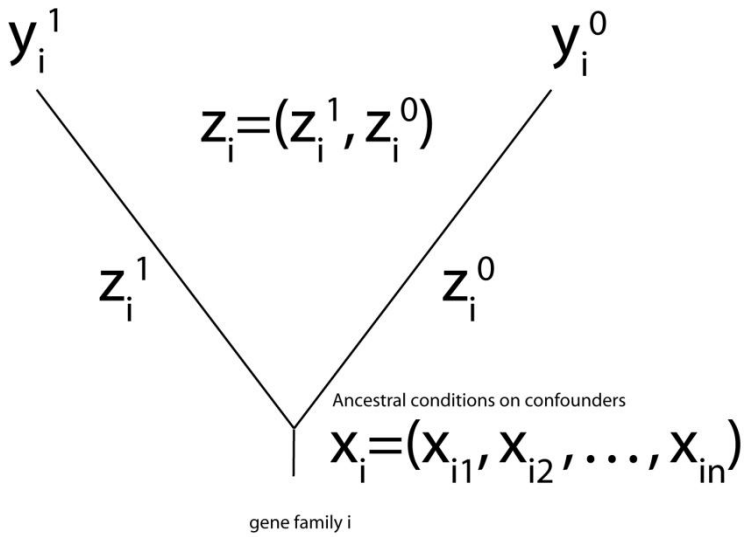
In the third chapter, the evolution of sex-biased gene expression is studied as a dynamic process where inferred states of sex-biased gene expression are used to construct a Markov model of genome evolution. Expression dynamics of sex-biased genes are investigated independently of gene birth and death processes and translocation events. Study of the stationary distributions associated with Markov models developed for each species lineage (from tip to root) suggests that sex differences in expression evolution contribute to the 'demasculinization' phenomenon.

The contribution of antagonistic genes to the 'demasculinization' phenomenon is further analyzed by separately modeling transition dynamics for X-linked and autosomal genes for genes showing enhanced lability on the sex comparative index. These results suggest that antagonistic coevolution may play a significant role in the 'demasculinization' of sex chromosomes via expression evolution, at least in certain lineages; however, it is not clear whether X-linkage facilitates antagonistic coevolution. These things can be said: candidate genes are not

overrepresented on the X chromosome and candidate genes are not more labile when X-linked. It is further predicted that an increase in the proportion of candidate genes on the X chromosome would lead to reduced ‘demasculinization’.

TABLES AND FIGURES

Figure 1 Schematic for a natural experiment on a phylogeny where multiple causes are considered. Above, $\mathbf{y}^0$ and $\mathbf{y}^1$ represent potential outcomes (for the study of 'faster-male' evolution, these may refer to rates of molecular evolution as observed in sister species), $\mathbf{z}_i$ denotes assignments to alternative treatment regimes (e.g., the possible treatments might include exposure to a male-specific environment or not), and $\mathbf{x}_i$ are ancestral conditions on confounders (e.g., ancestral states on tissue-bias and other supposed confounders of molecular evolution). In a phylogenomic context, the assumption of conditional independence can be evaluated across multiple orthologous gene families, where capital letters correspond to random variables describing the possible states on each variable for the individual study units (genes). It is assumed here that potential outcomes are defined as units of change where phylogenetic non-independence is not a concern.



Assumption of conditional independence:

$$(Y^0, Y^1) \perp Z \mid X$$

Chapter 2

Accounting for historical constraint in estimating the effects of sex-dependent environmental exposure on rates of molecular evolution

ABSTRACT

'Faster-male' evolution is widespread in *Drosophila*, as is evidenced by genome-wide analyses indicating that male-biased genes are evolving at enhanced rates of amino acid substitution as compared to nonsex- and female-biased genes. These observations lend support to the idea that genes that are disproportionately exposed to a male environment will experience more intense selection. This is consistent with sexual selection theory which predicts that intrasexual competition will often be stronger in males as a result of sex differences in the potential number of progeny. However, many other factors are known to affect gene evolution and may be associated with sex-biased gene classifications as a result of historical processes. To address the alternative hypothesis that relative constraints account for observed differences in the rates of gene evolution among classes of sex-biased genes in *Drosophila*, multivariate matching techniques were used. Male and female-biased genes were matched with nonsex-biased genes on several sets of proposed confounders and male and female effects were estimated. Adjustment for confounding due to historical sex differences resulted in a significant decrease in estimates of male effects on rates of molecular evolution and marginal increase or no difference in female effect estimates. This suggests that historical factors facilitate molecular evolution in males, and less so, constrain molecular evolution in females. After adjustment, male and female effects on rates of molecular evolution were similar suggesting that 'faster-male' evolution may be due to sex differences in the efficacy rather than in the intensity of selection.

INTRODUCTION

When Charles Darwin formulated his theory of sexual selection in The *Descent of Man, and Selection in Relation to Sex*, sex differences were understood largely in terms of morphology, physiology and behavior. Recent advances in modern genomics have permitted the study of sex differences at the molecular level, where they are apparent in the differential segregation of sex chromosomes and in the sex-biased expression of genes. On the basis of relative mRNA or cDNA expression levels in male and female tissues as measured in microarray experiments, genes have been classified into groups of male-, female-, or nonsex-biased genes. For a range of taxa, genome-wide analyses of sex-biased genes have provided evidence of sex-differences for several evolutionary parameters, including rate of gene evolution, typically estimated from comparative sequence data as the ratio of the rates of nonsynonymous to synonymous mutation (dN/dS) (Zhang et al. 2007, Ranz et al. 2003, Zhang et al. 2004, Zhang

and Parsch 2005, Ellegren and Parsch 2007). In the *Drosophila* study system, widespread observations that male-biased genes are evolving at enhanced rates as compared to nonsex-biased and female-biased genes has generally been interpreted as evidence that the genes that are expressed more highly in males are subject to more intense selection. This has led some to propose that intrasexual male competition may be the principal cause underlying patterns of sex difference in this genus (Zhang et al. 2004, Zhang et al. 2007).

While analyses in several taxa have purported to demonstrate a positive male effect on gene evolution, many other factors are supposed to affect rates of synonymous and nonsynonymous substitutions (Larracuente et al. 2008). Proposed determinants of evolutionary rate include: spatial (Zhang and Li 2004, Liao, Scott and Zhang 2006) and temporal expression (Cutter and Ward 2005), recombination rates (Comeron, Kreitman and Aguadé 1999), gene length (Comeron et al. 1999), intron number (Marais et al. 2005, Drummond et al. 2005), expression level (Drummond et al. 2005), gene essentiality (Fraser et al. 2002, Liao et al. 2006), and the extent of protein-protein interaction (Fraser et al. 2002). If any these factors covary with the evolution of sex-biased gene expression and also affect rates of gene evolution, confounding may lead to misunderstanding of the role of sex-dependent environmental exposure per se. Confounders are understood here as causal factors affecting evolutionary rate (dN/dS) that may be historically associated with the various classes of sex-biased gene expression: female-biased, male-biased, or nonsex-biased.

Mank *et al.* (2008) have hypothesized that tissue-bias may hamper the evolution of sex-biased gene expression. Tissue-bias is a generic measure of spatial expression breadth based on microarray assays of gene expression across different tissue types. High levels of tissue-bias indicate that gene expression is limited to a small number of tissues and low levels indicate ubiquitous expression throughout a given organism. Mank *et al.* reason that sex-differential fitness, which is supposed to drive the evolution of sex-biased gene expression (Rhen 2000), may be dampened as a gene is exposed to a broader range of tissue environments. Cross taxa observations suggest that tissue-bias may be a relatively conservative property of genes, whereas the evolution of sex-biased gene expression appears to be quite labile (Ellegren and Parsch 2007, Mank et al. 2008). If spatial expression of genes across tissue types often condition the evolution of sex-biased gene expression, this variable is likely to be associated with sex-biased gene expression (Mank et al. 2008). Since broad expression is supposed to lead to reduced evolutionary rates because of the potential for antagonistic pleiotropy to both inhibit positive selection and enhance purifying selection, the effects of spatial expression on gene evolution may account for differences in evolutionary rate seen among various classes of sex-biased genes.

Recently, attempts have been made to adjust for spatial expression profiles in analyzing the relationship between sex-biased gene expression and evolutionary rates (Meisel 2011). Tissue-bias was found to largely account for enhanced evolutionary rates seen for a pooled group of sex-biased genes (including both male- and female-biased genes). According to this analysis, sex-biased genes that are expressed in reproductive tissues still show evidence of enhanced rates of molecular evolution in comparison to nonsex-biased genes with comparable levels of tissue-bias.

However, spatial expression may not be the only historical factor confounding estimates of the effect of sex-dependent gene expression on evolutionary rates. Historical sex differences (that is ancestral conditions on other causes of molecular evolution) should be accounted for in attempts to estimate the effects of exposure to a sex-dependent selection regime (for which the class of sex-biased gene expression is proxy) on rates of molecular evolution. The question of

whether historical sex differences in expression profiles or other covariates of evolutionary rate may account for widely observed pattern of 'faster-male' evolution has not yet been addressed.

The aim of this analysis is to assess the role of historical constraint in influencing rates of gene evolution among various classes of sex-biased genes. The processes determining which kinds of genes become sex-differentially expressed may be highly biased making it difficult to distinguish between effects of ongoing exposure to male or female-specific environments and the effects of historical sex differences. In order to address the hypothesis that relative constraints account for observed differences in the rate of male and female gene evolution, an alternative statistical approach derived from the field of observational study methods is used. Proposed historical factors are treated as statistical confounders in an explicitly observational study design. This approach uses subclassification on a propensity score, a simple matching-based method for reducing bias in observational studies (Rosenbaum and Rubin 1983, Cochran 1968).

METHODS

The technique of subclassifying on a propensity score is used here to reduce bias associated with historical factors that may affect gene evolution and be associated with sex-biased gene groupings. These methods are based on a quasi-experimental design approach to the analysis of observational data, where the randomized experiment is taken as the ideal basis for scientific inference.

The basic idea of an observational study design is to eliminate associations between treatment assignments and hypothetical outcomes by comparing study units that are matched on confounders; that is, factors that may have independent causal effects on the outcome of interest and that affect the treatment selection process (Rosenbaum 2002). The aim here is to isolate the effects of exposure to a sex-dependent environment (for which sex-biased gene expression is proxy) on gene evolution (ω) by evaluating effect estimates for genes that are closely matched on multiple proposed confounders.

Propensity-score matching is a useful tactic when there are multiple confounders (Rosenbaum and Rubin 1983). Instead of matching groups of sex-biased genes with exactly the same values on each of a multitude of confounding factors, matching is based on an estimated propensity-score describing the probability that a given gene is in any particular class (male-biased, female-biased, or nonsex-biased) given its values on the set of confounders. It has been shown that an appropriately specified one-dimensional propensity-score—typically estimated using logistic regression on the covariates for which matching is desired—can be effectively used to find groups that are well-matched on multiple covariates at once. Estimating effects from groups that are well-matched with respect to confounders will eliminate or reduce bias due to these variables (Rosenbaum and Rubin 1983).

Part 1: Data Analysis

Explanatory variables

The explanatory variable in this study is exposure to a male or female environment, where data on male and female-biased gene expression derived from microarray analysis is used as a proxy measure for sex-specific environmental exposure. Genes were grouped into classes of sex-biased gene based on data derived from Zhang *et al.*'s multi-species study of sex-biased gene expression for multiple species of *Drosophila*. Genes were classified into male-, female-, or nonsex-biased classes based on genome-wide hypothesis testing comparing expression intensity

readings derived from samples of male and female tissues. A full description of the classification method is described in **Appendix A**. Only sex-dependent gene expression data from *D. melanogaster* was used in this analysis.

Confounders

Confounders are understood here as independent causal factors affecting contemporary measures of evolutionary rate (dN/dS) that may be associated with the classes of sex-biased gene expression due to historical processes. Many factors have been proposed to affect rates of gene evolution. In the most comprehensive study to date Larracuente *et al.* integrated data on several proposed determinants of evolutionary rate. These include: spatial breadth (usually measured as tissue-bias, which is defined below), gene expression, intron length, protein length, intron number, and rate of recombination, and the number of protein-protein interactors (PPI) (**Table 1**). Most of these variables were included in the present analysis, including all of those found to have a significant partial correlation with evolutionary rate based on a multivariate analysis (protein length, expression, intron number, and spatial expression breadth) (Larracuente et al. 2008). Although, the number of protein-protein interactors did not show a significant partial correlation with evolutionary rate, it was also analyzed. Intron length was not included since it is highly correlated with both protein length and intron number, both of which were included.

Additionally, several independently derived binary variables— including indicators of expression at the adult stage, in reproductive tissues, and of X-linkage—were considered as potential confounders in estimating the effects of ongoing sex-dependent exposure on evolutionary rates. Expression at adult stages and in reproductive tissues is likely to enhance exposure to sex-differential selection pressures which may lead to effects on gene evolution that are independent of those due to sex-differential environmental exposure per se (a gene does not have to be sex-biased to be exposed to sex-differential selection pressure). Theoretically X-linkage may lead to enhanced rates of evolution (Charlesworth et al. 1987). According to theoretical predictions that the X-chromosome should harbor sexually antagonistic fitness variation (Rice 1984, Jonathan, Adam and William 2002), X-linkage may be associated with the evolution of sex-biased gene expression, since sexual antagonism is a potential driver of sex-biased gene expression (Rhen 2000).

In an observational context, such as this one, where the 'treatment' variable (interpreted in this case, the class of sex-biased gene expression viewed as a proxy for sex-dependent environmental exposure) cannot be expected to have been randomly assigned to study units (genes), the appropriate choice of adjustment variables is the key to valid inference. Several assumptions are made here: It is assumed that the values measured on several variables in *D. melanogaster* reflect historical conditions; that is, they are similar or identical to values on these variables corresponding to the start of the measurement period over which the outcome variables correspond. Further discussion of the issues related to the appropriate choice of confounders and proposals for advancing causal inference in evolutionary genomics is taken up in the discussion section.

Outcome variables

Effects were estimated for two measures of evolutionary rate: ω_{Mel}, a lineage-specific estimate for the branch connecting *D. melanogaster* to the common ancestor of *D. simulans* and *D.melanogaster*; and ω_{M0}, a genus-wide estimate across 12 *Drosophila*. Assumptions about the historical status of covariates measured in *D. melanogaster* are more likely to hold for the former than the latter.

Genome-wide data on sex-biased gene classes (male-, female-, and nonsex-biased) analyzed for *D. melanogaster* (Zhang et al. 2007) were integrated with evolutionary and functional genomic datasets from several different sources. Most of the genomic variables relevant to the study of evolutionary rate were previously consolidated by Larracuente *et al.*'s in their multivariate analysis of the determinants of protein evolution, including: PPI, spatial breadth, intron number, recombination, protein length, gene expression, as well as estimates of evolutionary rate based on comparative sequence analysis in *Drosophila* (**Table 1**). Data on these variables were used without additional modification, with the exception that the outcome variables, ω_{mel} and ω_{M0}, were cleaned of values of dN/dS values above 100.

Several additional variables were independently derived from existing microarray datasets on spatial (Chintapalli, Wang and Dow 2007) and temporal (Arbeitman et al. 2002) gene expression. These included two binary variables (yes=1/no=0) indicating enhanced expression at adult stages and in the reproductive tissues. Both indicator variables were derived with reference to a generic measure of expression bias based on the average deviation from the maximum expression level, as measured over multiple tissues or life stages:

$$\rho_{\tau/\varphi} = \frac{\sum_{i=1}^{N}\left(1 - x_i / x_{\max}\right)}{N-1}$$

Values of ρ nearer to 0 indicate that gene expression is comparable across multiple sampled tissues or life-stages (incl., embryonic, larval, metamorphic or adult), while values nearer to 1 indicate that expression is more limited in its spatial or temporal extents. For the calculation of tissue-bias (ρ_τ), each x_i corresponds to a log ratio of expression measures for each sampled tissue to the measures based on assay of whole fly tissue samples. For the temporal data set, the x_i correspond to log ratios of expression at each stage sampled divided to a reference sample. In calculating both tissue and stage-bias, x_{max} refers to the maximum of the respective x_i. A gene was designated as reproductive if expression was maximal in the reproductive tissues (testis, ovary, spermatheca, accessory glands, or tubule) and if tissue-bias was above the 75th percentile across all genes. A gene was considered to be adult if expressed maximally in the adult stage and if stage-bias was above the 75th percentile across all genes.

Part 2: Statistical Analysis

In this study, the effects of sex-specific environmental exposure on rates of gene evolution were estimated on the basis of two comparisons: male-biased versus nonsex-biased genes and female-biased versus nonsex-biased genes. Comparisons of male and female effects were made indirectly. Effect estimates were based on several different model criteria specified in terms of distinct sets of proposed confounders.

Propensity score modeling

To construct matched groups for comparison, propensity scores were estimated on the basis of thousands of multinomial logistic regression models expressed in terms of proposed confounders of evolutionary rate. Each logistic model yielded a probability of being male-, female- or nonsex-biased for each gene as conditioned by its particular covariate background. For each comparison, genes were cssified into one of 6 groups based on quantiles of the

estimated propensity score. Quantiles were constructed from the propensity-score distribution corresponding to the group of sex-biased genes in each comparison. This was done to ensure adequate overlap between the sex-biased and nonsex-biased genes, where nonsex-biased genes outnumber sex-biased gene classes. Here, overlap refers to the number of data units within each subclass.

To find the subclassification scheme that produced the best balance on a set of proposed confounders, multiple propensity score models were considered. Rather than evaluating separate binomial logistic regression models for each of the comparisons, propensity scores were estimated in a single swoop using multinomial logistic regression. While a single multinomial logistic regression model yields probabilities corresponding to each of the three classes of sex-biased genes, matching was evaluated separately for each comparison so that propensity scores could differ between comparisons; that is, the probability of being male or female-biased was not generally assumed to depend on the same set of variables.

Propensity-score models were specified in terms of several model criteria (M1-M3) stipulated in terms of distinct sets of proposed confounders of the effects of sex exposure on evolutionary rate (**Table 4**). For each of three sets of n proposed confounders (for M1, n = 2; for M2, n = 8; and for M3, n=9), the search space over which the optimal subclassification was selected consisted of all possible propensity score model specifications that are at most quadratic, allowing for no self-interaction terms (that is, for each of $\binom{n}{k}$ choices of k distinct variables for the linear term there are $2^{\binom{k}{2}}$ possible corresponding interaction terms). For each comparison, the optimally balanced model was selected according to the following criteria.

Model selection

Three criteria for model selection (M1-M3) were specified in terms of distinct sets of proposed confounders (**Table 4**). The first criterion (M1) seeks balance for the spatial breadth and reproductive variables alone. These are the variables that have been adjusted for in the most recent analyses of sex-biased genes in *D. melanogaster*, where it has been shown that adjustment for tissue-bias largely removes the effect of enhanced evolutionary rate except among reproductive genes (Meisel 2011). The purpose of analyzing this model here is to assess whether imbalance remains on other potential confounders, to compare the effect estimates based on this model to those for which more comprehensive adjustments are made, and to separately analyze male- and female-biased genes. The third and most comprehensive model criterion, M3, maximizes balance for all 9 of the proposed historical factors affecting evolutionary rates: PPI, spatial breadth, intron number, recombination, protein length, gene expression, reproductive, adult and X-linkage. The dataset from which adult indicator variable is derived (adult: yes = 1, no = 0) comes from a study of gene expression over the lifecycle of *Drosophila* which is less than genome-wide in scope. Imputation on this variable was not conducted. The effects of excluding or including this variable can be assessed by comparing the full model M3 to M2, which excludes this variable.

Model criteria were applied uniformly to each of the two comparisons: male-biased versus nonsex-biased, and female-biased versus nonsex-biased genes. Specifically, multiple propensity score models were estimated on the basis of the n proposed confounders associated with each model criterion. For any given propensity score model, its corresponding subclassification can be evaluated for 'balance' on each of multiple covariates. Balance is assessed for each covariate using a 'balance score' (β_i, defined for each covariate i). For a single

covariate, balance is maximized when the balance score is as near to zero as possible. The balance score is defined differently for binary and continuous variables. In the case of continuous variables, the balance score is calculated as the standardized weighted mean difference across subclasses, where the weights are the proportions of genes within each subclass and the differences are taken between comparison groups. For indicator variables the balance score is the mean difference in the proportion of each indicated variable as evaluated for each subclass. The best matched (or balanced) subclassification was selected by first maximizing the number of variables for which $\beta_i < 0.10$, and then minimizing the mean of the balance scores across covariates, hereafter referred to as the balancer ($B = \bar{\beta}$). On the basis of the optimally matched subclassifications selected for each comparison, male and female effects were estimated.

For models criteria M2 and M3, exhaustive analysis of all of the possible propensity score models was not conducted; instead, repeated bouts of random sampling (N = 10,000) of the propensity-score models were conducted. After each bout of sampling, balance for each of the proposed confounders was evaluated and the best model was determined according to criterion described above. If imbalance remained ($\beta_i > 0.10$ for any of the proposed confounders), subsequent iterations used the best fitting model of the prior bout of sampling, while selectively adding terms related to the still imbalanced covariates until matching was achieved ($\beta_i < 0.10$ for all confounders). For each comparison and criterion, a single propensity score model was selected for effect estimation (**Table 4**).

Estimating male and female effects on evolutionary rates

Male and female effects were based on direct comparisons between matched groups of each class of sex-biased genes with nonsex-biased genes. In each case, effects were calculated as the subclass-weighted mean difference between comparison groups, where the weights corresponded to the overall proportion of genes in each subclass. Male and female effects were indirectly compared; although direct matched comparison of male and female-biased genes was also attempted (further discussion can be found in the Results section).

For each direct comparison, simultaneous confidence levels were determined using the 0.83 and 99.2 percentiles of the bootstrap corresponding to Bonferroni-adjusted p-values of 0.05 (n = 3). The bootstrap distribution for each effect estimate was compiled by fixing each subclassification and resampling genes based on Monte Carlo simulation of 10,000 random permutations of the treatment assignments (the sex-biased gene classifications). Confidence intervals for the indirect adjusted comparisons were calculated from standard errors derived from the bootstrap simulation of the male and female effect estimates, $SE_{adj} = \sqrt{SE_{\text{female effect}}^2 + SE_{\text{male effect}}^2}$, and based on percentiles of the normal distribution.

Sensitivity Analysis

In order to evaluate how effect estimates varied for different threshold levels of the balancer B, sensitivity analyses on both outcome variables were conducted across all comparisons and models. Holding the sex-biased gene groupings fixed, pairs of subclass assignments were randomly shuffled and balancer and effect estimates were calculated for each of 1,000 such permutations. The range of effect estimates for fixed steps of size 0.01 away from the minimum balancer were compiled and compared.

RESULTS

Heterogeneity among groups before adjustment

Prior to propensity score analysis, heterogeneity among groups of sex-biased genes was analyzed for all 9 of the determinants of evolutionary rate. For continuous variables, the hypothesis of heterogeneity was evaluated using the Kruskal-Wallis test (**Figure 2; Table 3**). For variables yielding a significant result among groups, post-hoc comparisons were conducted for all possible dichotomous comparisons based on the Mann-Whitney U-statistic. The Fisher Exact Test was used to test for associations among sex-biased gene groupings and the indicator variables (adult, reproductive, and X-linked) (**Figure 2; Table 2**). Significance cutoffs were based on Bonferroni adjusted p-values (N = 3) for both tests.

With the exception of recombination rate, male-biased and female-biased genes differ markedly from one another for all of the covariates of evolutionary rate. Male-biased genes differ from non-sex biased genes in that: they are more tissue-biased, code for proteins of shorter length, and show evidence of having lower expression, where expression is evaluated using the first principle component of codon bias and the maximum of expression levels assayed across tissues (Larracuente et al. 2008). As is consistent with previous analyses, male-biased genes are also more likely to be expressed in reproductive tissues (Meiklejohn et al. 2003, Arbeitman et al. 2004), and are underrepresented on X chromosomes (Sturgill et al. 2007, Parisi et al. 2003a). Male-biased genes are more likely to be specifically expressed at the adult stage and have fewer introns than nonsex-biased genes. Female-biased are more widely expressed and have a larger number of protein-protein interactors as compared to nonsex-biased genes.

Propensity score matching

Male and female-biased gene classes are markedly different for a large number of covariates. In practice this makes it difficult to find matched groups for direct comparison of male-biased and female-biased genes, even with the aid of the propensity score. Rather than limit analysis to a reduced sample size of matched male and female-biased genes, the approach taken here was to estimate male and female effects by comparing each set of sex-biased genes to a matched control group of nonsex-biased genes. Male and female effects were then compared using indirect adjusted comparison. Indirect comparison was viewed as preferable for genome-level analysis. Direct comparisons of male and female-biased genes were nonetheless conducted and accompanying covariate balance tables and effect estimates for these comparisons are shown in **Appendix B** (**Figures 20, 23, and 26; Tables 32 and 33**).

The best propensity score model for each possible comparison and model criterion is shown in **Table 4**. Although the model criteria (M1-M3) were applied uniformly to each comparison, the best fitting model was assessed separately for each comparison. For each model criteria, propensity score models were specified only in terms of the proposed confounders listed (**Table 4**). While the application of model criteria M2 and M3 resulted in the selection of different propensity-score models for each of the comparisons, criteria M1 selected the same propensity score model for both. For the comparison of female and nonsex-biased genes, the application of criteria M2 was adequate for matching on all variables; inclusion of the adult variable in the propensity score did not improve overall balance for this comparison. As a result, the propensity score models for this comparison are identical for M2 and M3. Sample sizes were comparable for most models since genme-wide data is available on most of the variables used. However, data on temporal expression is still limited in *Drosophila*, so subclassification analysis

based on propensity scores specifications including the adult variable (M3) were based on a reduced sample size.

Covariate balance tables for each model (**Tables 5,7, and 9**) along with effect estimates (**Figures 3-5; Tables 6, 8, and 10**) are shown for all comparisons and model criteria. Additionally, overlap and balance can be examined graphically with reference to the Supplement (**Figures 20-28**). Within each covariate table, several statistics are shown: the difference in mean between comparison groups prior to adjustment, the adjusted mean difference, the percentage change in mean difference, and the balance score. The percentage change is calculated as: $100(|a| - |b|)/|a|$, where a is the unadjusted mean difference and b is the adjusted mean difference. The balance score is defined differently for binary and continuous variables. In the case of continuous variables, the balance score is calculated as the absolute value of the standardized weighted mean difference across subclasses, where the weights are the proportions of genes within each subclass and the differences are taken between comparison groups. For indicator variables the balance score is the absolute value of the mean difference in the proportion of each indicated variable evaluated for each subclass. Generally, balance scores exceeding 0.20 are considered problematic according to recommended guidelines for quasi-experimental designs (Stuart and Rubin 2007). These are highlighted in bold throughout the covariate tables. For each of the model criterion, adequate balance was achieved across all of the proposed confounders (see Table 5-6, where $\beta < 0.10$ for all associated proposed confounders). However, for both M1 and M2, imbalance remains on several of the covariates of evolutionary rate that were not included in the model criteria ($\beta > 0.10$).

Application of the first model criterion (M1), which treats only reproductive and spatial breadth variables as confounders, resulted in the selection of the same propensity score model for both comparisons. The best balanced model for both comparisons is specified in terms of the spatial breadth variable (**Table 4**). Although prior to adjustment, heterogeneity on both the reproductive and spatial breadth variables was indicated for the male versus nonsex-biased gene comparison, adjusting for tissue-bias effectively reduced bias on both variables (**Table 5, Figure 21**). Balance is achieved on the confounders included in criterion M1, but imbalance remains for several variables that were not included in the matching criterion. In the comparison of male-biased with nonsex-biased genes, nonsex-biased genes have a larger mean proportion of adult expressed genes and a larger mean protein length after adjustment ($\beta > 0.20$). While protein length was longer for nonsex-biased genes prior to adjustment, the procedure of matching has reversed the direction of imbalance on adult expressed genes. For the comparison of female-biased with nonsex-biased genes, the adjusted mean number of protein-protein interactors is larger for female-biased genes, as was the case prior to adjustment.

Despite imbalances in both the male and female comparisons, effects are estimated on the basis of M1 (**Figure 3**). With adjustment, the male effects on evolutionary rate are significantly reduced for both outcome variables (ω_{M0} and ω_{mel}), and are no longer statistically significant for the *Drosophila melanogaster* branch-specific estimates of evolutionary rate, though they remain so for genus-wide estimates of evolutionary rate. Estimates of female effects are marginally larger than pre-adjusted estimates, but the negative female effect on evolutionary rate is no longer statistically significant, though this is marginal. Where prior to adjustment the confidence intervals of the male effect and the female effect did not overlap on either outcome variable, after adjustment male and female effect estimates are within error for ω_{mel}, although they are nonoverlapping for ω_{M0}.

Effect estimates derived from M1 may be biased due to outstanding imbalance on variables other than those included in the criterion for matching (**Table 4**). While imbalance suggests that bias may exist for both comparisons, its expected direction is unclear. There is little consensus regarding how the number of protein-protein interactors will influence evolutionary rates (Fraser et al. 2002, Jordan, Wolf and Koonin 2003). While protein length is expected to have negative effects on evolutionary rate (Comeron et al. 1999, Lemos et al. 2005b), it seems likely that adult expression will have positive effects: adult expressed genes may experience stronger selection due to enhanced exposure to sexual selection, while restricted temporal expression, especially the later in ontogeny, may reduce exposure to pleiotropic constraint (Kirkwood and Rose 1991).

M2 adjusts for all of the potential confounders of evolutionary rate, except for the temporal variable indicating adult expression. By matching on the proposed confounders associated with M2, the female comparison is fully balanced with respect to all of the potential confounders (**Table 7b; Figure 25**); however, for the male comparison, imbalance remains for the variable indicating adult expression. Adult expressed genes are seen in higher proportions across the subclasses of nonsex-biased genes (**Table 7a; Figure 24**). Improved balance on PPI appears to have further boosted female effect estimates so that there is little support for a negative female effect on evolutionary rate for either measure (**Figure 4**). Male effects are not appreciably altered with improved balance on protein length but there may be outstanding bias due to imbalance on the adult expression variable.

Adding the adult variable to the propensity-score estimate achieves balance on this variable for the M vs. N comparison (**Table 9a**), but selection bias is apparent. This can be seen by comparing effect estimates prior to adjustment for analyses based M1 and M2, which involve larger sample sizes, with pre-adjusted effect estimates for M3. Estimates are about halved in each case (e.g., for ω_{mel}, the unadjusted male effect estimate is 0.90 based on the larger sample size, but is 0.49 for M3) (**cf. Table 8 and Table 10**). Furthermore, the control group of nonsex-biased genes is different for the male and female comparison, which further weakens inference based on indirect comparisons. However, if selection bias is on the same variables as those included in the criteria, indirect matching adjusted comparison should also reduce its effects. Estimates based on M3 (**Table 10, Figure 5**) are identical to M2 for the female effect since the same propensity score model was used. Adjustment for adult expression decreases estimates of male effect for the genus, but increases the lineage-specific estimate. Confidence is reduced due to the smaller sample size. Based on this model, male effects are not significant for either outcome variable.

Finally, sensitivity analyses were conducted for each comparison and model, yielding effect ranges as a function of the balancer B (**Tables 11-12**). Estimates of female effects on evolutionary rate generally show a wider range of values for both outcome variables as threshold levels of the balancer (Γ_B) are increased. Thus, female effect estimates are more sensitive to misspecifications of the propensity-score model.

CONCLUSIONS

Male and female-biased genes differ with respect to a wide range of variables, many of which may independently constrain or facilitate the rate of evolution, as measured here in terms of ratios of nonsynonymous to synonymous mutation. For all model specifications, estimates of male effects on evolutionary rate are reduced as compared to unadjusted estimates that do not match genes on potential confounders of evolutionary rate. Meanwhile estimates of female effects are largely unchanged with adjustment, or are marginally increased. For *D. melanogaster*,

there is little evidence for either positive male effects or negative female effects with adjustment for the full set of confounders.

These findings have implications for understanding the cause(s) of pervasive patterns of 'faster-male' evolution in *Drosophila*, and may have implications for other taxa in which 'faster-male' evolution is observed. Positive associations between evolutionary rates and male-biased gene expression have been widely interpreted as evidence that males experience more intense selection than females. This view is consistent with sexual selection theory which predicts the existence of sex-dependent selection regimes arising from differences in reproductive strategy. These reproductive strategies are characterized by anisogamy, where females invest more in a few gametes and males invest less in more gametes. As a result, the male has the potential to produce more offspring than the female, which has implications for both intersexual and intrasexual dynamics. Functional differences are expected to evolve for traits that mediate intersexual conflicts; for example, females may evolve preference traits enabling them to choose a high quality mate, while males may evolve signals that make them more attractive to the female sex. It is also expected that the intensity of intrasexual competition will be higher among males, since males are expected to compete among themselves for a limited number of mating opportunities. If coevolutionary dynamics can be expected to have reciprocal effects on genome-wide rates of evolution, then asymmetrical effects of intrasexual competition may lead to 'faster-male' evolution.

The results presented here provide support for the alternative view that 'faster-male' evolution primarily is due to differences the efficacy rather than in the intensity of selection. Male and female effects are seen as comparable after full adjustment for proposed historical factors that may facilitate or constrain selection. It appears that these factors act overall to facilitate evolution in males, and less so, to constrain evolution in females. Previous analyses of sex-biased genes that have not accounted for, or have only accounted for a few confounders, are likely to be biased. This is supported by the analysis of heterogeneity among groups prior to propensity score analysis (**Tables 2-3**) and by the results from M1 (**Figure 2, Table 6**), where outstanding imbalance is seen for both comparisons despite adjustments for spatial breadth and reproductive expression. For the models showing the best balance across the full set of covariates (M3), male and female effect estimates are comparable for both outcome variables.

These results do not preclude a role for sexual selection in explaining 'faster-male' evolution in the *Drosophila* genus; rather, its role is qualified as historical. It is clear that sexual selection has played a significant role in determining which kinds of genes have become sex-biased, as is evidenced by the different representation of reproductive and adult expressed genes among the sex-biased genes (**Table 3**). Further research should focus on understanding the processes that determine which kinds of genes are likely to become male versus female-biased and on elucidating mechanisms that account for the apparently greater evolvability of males in this genus.

Evolutionary genomics, even more so than most fields, is limited in its capacity to conduct experiments. Although considerable effort has been made to address potential confounding, it is important to acknowledge some of the limitations of any observational design framework. The omission of unidentified confounders or selection bias on measured variables may bias the effects estimated here and lead to erroneous conclusions. It is likely that temporal expression patterns influence gene evolution but it appears that discrete sampling over the *Drosophila* life cycle may miss the more quickly evolving genes. It seems likely that this is explained by a tendency for quickly evolving genes to be highly stage-biased. It is not clear how

better coverage on this variable would affect the conclusion presented here. The adjustments made based on available data suggest that male effects will be negligibly decreased for the genus, but increased for the *D. melanogaster* lineage (**cf. Table 8 and Table 9**).

The conclusions presented here are based on assumptions about the ancestral conditions of sex-biased genes. It is assumed that the measures on the potentially confounding variables are similar to ancestral values at the start of the outcome measurement; that is, that they do not reflect response values that have been affected by differential exposure to a male or female environment over the duration over which the outcome is measured. If this were not the case, then it would be inappropriate to adjust for them since their effects should instead be subsumed in the effect estimate. Ideally, one could condition on accurately inferred ancestral values for each of the covariates of evolutionary rate; however, at this time, multi-species data is lacking for several key genomic variables and appropriate evolutionary models for several genomic variables are yet to be determined. Although two outcome variables were analyzed here, all of the covariate values for which adjustments were made were derived from measures taken in D. *melanogaster*. Thus, the assumption that these covariate values are similar to values prior to the period over which the outcome variable is measured is more likely to hold for estimates based on ω_{mel} than for those based on ω_{M0}. Similarly, sex-biased gene classifications can and do change over the time periods represented by the genus-wide estimates (Ranz et al. 2003, Meiklejohn et al. 2003), and so male and female effects based on this outcome variable may be biased by the different proportions of time a gene has spent in each state.

Finally, subtle issues arise in choosing adjustment variables in the context of observational studies; it is known for example, that adjustment on one covariate can lead to bias on another and that variables that proxy for or directly measure responses to exposure are problematic (Pearl 2009). An example of the former was seen when adjustment for tissue-bias and reproductive expression resulted in imbalance on adult expression for M1. This suggests that associational criteria are generally inadequate for choosing confounders; it is preferable that a mechanism of independent causal effect be advanced. Mechanisms have been advanced for the proposed confounders used in this analysis (Larracuente et al. 2008), but association analyses in too large a part form the basis of knowledge in the field of evolutionary genomics. Despite the considerable limitations of observational study, conclusions are presented here with the hope that further critical analysis will lead to the development of more robust methods for causal inference in evolutionary genomics.

FIGURES AND TABLES

Table 1 Data sources Details about of how independently derived variables (†) were calculated are described in the data analysis subsection of the methods. See original source for details regarding the derivations of the other genomic variables.

Data Type	Variables	Description	Source
Spatial expression	Reproductive†	Binary variable indicating that gene expression is maximal in a reproductive tissue (incl., testis, ovary, tubule , and accessory glands) and tissue-bias (see below) is above the 75th percentile	(Chintapalli et al. 2007)
Evolutionary rate	ω_{mel}, ω_{M0}	Ratio of the rate of nonynonymous to synonymous mutations estimated from comparative sequence data in *Drosophila*. ω_{mel} is the evolutionary rate estimate for the branch connecting *D. melanogaster* to the common ancestor of *D. melanogaster* and *D. simulans* and ω_{M0} is genus-wide estimate.	(Consortium 2007)
Various	PPI	Number of protein-protein interactors for each gene	(Larracuente et al. 2008)
	Tissue-bias	Measure of spatial expression breadth ranging from 0 to 1, based on multi-tissue expression data	
	Gene expression	First principle component of codon bias (as measured using the frequency of optimal codons) and the maximum of expression across the spatial expression profile based on an 11 tissue assay of expression from FlyAtlas	
	Intron Length	Average length of introns	
	Intron Number	The number of introns for each gene	
	X-linkage	Presence/absence on the X chromosome	
	Recombination	Rate of recombination based on physical and genetic maps from *D. melanogaster* release 4.3	
	Stage-bias†	Measure of spatial expression breadth ranging from 0 to 1, based on microarray expression assays throughout the life cycle of *D. melanogaster* (embryonic, larval, metamorphic, and adult)	
Temporal expression	Adult†	Binary variable indicating that gene expression is maximal in either adult tissues (male or female) and stage-bias is above the 75th percentile	(Arbeitman et al. 2002)

Figure 2 Heterogeneity among sex-biased genes First and second rows show boxplots of various genomic variables for female (red), male (blue), and nonsex-biased genes. Outliers have been omitted. Third row includes barplots showing the proportion of male-biased (M), female-biased (F) and nonsex-biased genes (N) that are reproductive, adult, and X-linked.

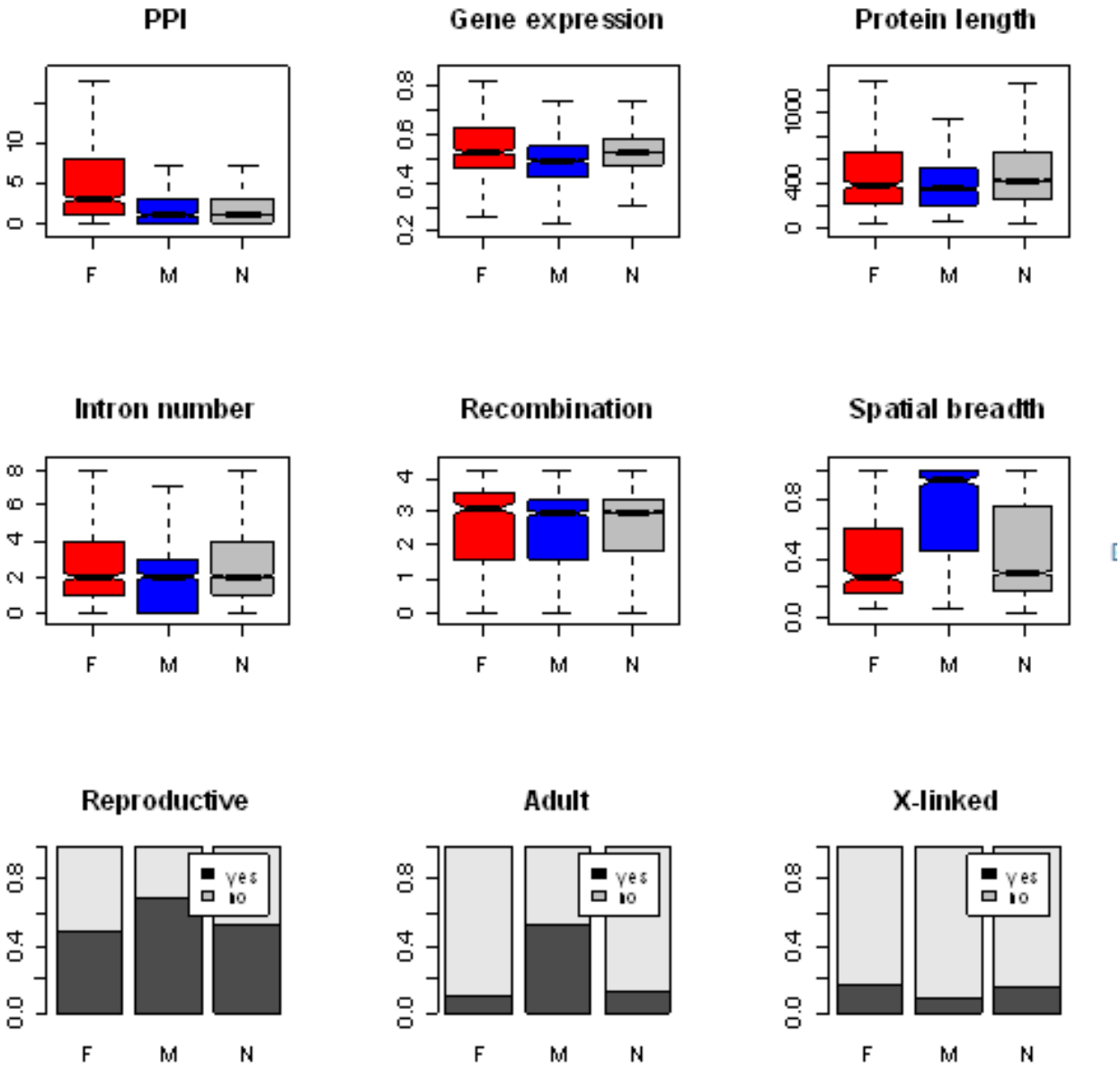


Table 2 Tests of independence for binary covariates of evolutionary rate. Table includes (from left to right) the covariate name, the number of sex-biased genes for which data on that covariate is available, results from two-tailed Fisher Exact Tests for all dichotomous comparisons.

Binary Variable	**(N_F,N_M,N_N)**	**Mult. Comp C1 vs C2**	**Odds ratio (C_1/C_2)**	**Sig.**
Reproductive	(286,822,4888)	F vs. M	0.42	***
		F vs. N	0.86	*ns*
		M vs. N	2.02	***
Adult	(155,353,2698)	F vs. M	0.10	***
		F vs. N	0.78	*ns*
		M vs. N	7.74	***
X-linked	(339,942,6700)	F vs. M	2.13	***
		F vs. N	1.10	*ns*
		M vs. N	0.52	***

Table 3 Heterogeneity among groups for continuous covariates of evolutionary rate Table includes (from left to right) the variable name, the number of sex-biased genes for which data on that variable is available, Kruskal-Wallis test results for differences among groups (X^2) with standard significance cutoffs, post-hoc test results for multiple comparisons based on the U-statistic. Significance is based on Bonferroni adjusted cutoffs for two-tailed hypothesis: * ($0.01 \leq p_{adj} < 0.05$); ** ($0.001 \leq p_{adj} < 0.01$); *** ($p_{adj} < 0.001$).

Continuous Covariate	**(N_F,N_M,N_N)**	**Comp.**	**X^2**	**Sig. (X^2)**	**U**	**Sig. (U)**
Spatial breadth	(324,915,6046)	F vs. M	606.78	***	2190.19	***
		F vs. N			397.38	**
		M vs. N			1792.81	***
PPI	(339,948,6716)	F vs. M	116.96	***	1366.77	***
		F vs. N			1307.73	***
		M vs. N			59.04	*ns*
Gene expression	(339,948,6716)	F vs. M	115.16	***	1064.73	***
		F vs. N			230.14	*ns*
		M vs. N			834.59	***
Protein length	(339,948,6716)	F vs. M	89.30	***	436.64	**
		F vs. N			308.68	*
		M vs. N			745.32	***
Intron number	(339,948,6716)	F vs. M	92.74	***	448.57	**
		F vs. N			303.56	*ns*
		M vs. N			752.13	***
Recombination	(334,945,6672)	F vs. M	2.34	*ns*	*(165.12)*	*ns*
		F vs. N			*(49.81)*	*ns*
		M vs. N			*(115.31)*	*ns*

Table 4 Best propensity score models for each comparison and model criteria. Note that the propensity score models are identical for the comparisons indicated ([a]†).

Crit.	Proposed confounders	(N_F,N_M,N_N)	Comp.	Best propensity-score model
M1	Spatial breadth, Reproductive	(0,915,6046) (324,0,6046)	M vs. N[a] F vs. N[a]	P-score = Spatial breadth P-score = Spatial breadth
M2	PPI Gene expression Recombination Protein length Intron number Spatial breadth X-linkage Reproductive	(0,915,6046) (324,0,6046)	M vs. N F vs. N†	P-score = Spatial breadth + Gene expression + Protein length + Intron number + Spatial breadth x Gene expression + Spatial breadth x Protein length + Gene expression x Protein length + Gene expression x Intron number P-score = Spatial breadth + PPI + Protein length + Intron number + Spatial breadth x Protein length + Spatial breadth x Intron number +PPI x Intron number
M3	PPI Gene expression Recombination Protein length Intron number Spatial breadth X-linkage Adult Reproductive	(0,344,2580) (324,0,6046)	M vs. N F vs. N†	P-score = Spatial breadth + PPI +Gene expression + Protein length + Intron number + Adult + Spatial breadth x Adult + PPI x Gene expression + PPI x Protein length + PPI x Adult + Gene expression x Protein length + Gene expression x Intron number + Protein length x Intron number + Intron number x Adult P-score = Spatial breadth + PPI + Protein length + Intron number + Spatial breadth x Protein length + Spatial breadth x Intron number + PPI x Intron number

Table 5 Covariate balance forM1. From left to right, list of covariates, mean difference prior to adjustment, mean difference after adjustment, the percentage change in the mean difference before and after adjustment, and the balance score (β). Proposed confounders included in the matching criteria for model selection are indicated (†).

Table 5a Covariate balance for M versus N (M1)

Covariates	Diff. (no adj.)	Diff. (adj.)	Perc. Change	β
PPI	0.08	-0.11	-39.88	0.01
Gene expression	-0.03	-0.01	81.25	0.07
Recombination	-0.07	-0.09	-33.95	0.07
Protein length	-119.46	-68.88	42.34	0.19
Intron number	-0.80	0.28	65.30	0.09
Spatial breadth†	0.29	0.01	93.80	0.05
X-linkage	-0.83	0.05	94.41	0.05
Adult	0.08	-0.47	-517.74	**0.47**
Reproductive†	0.38	0.02	93.79	0.02

Table 5b Covariate balance for F versus N (M1)

Covariates	Diff. (no adj.)	Diff. (adj.)	Perc. Change	β
PPI	2.54	2.06	19.00	**0.29**
Gene expression	0.02	0.00	74.83	0.09
Recombination	-0.06	-0.06	3.86	0.05
Protein length	-31.89	-8.83	72.31	0.04
Intron number	-0.55	-0.54	2.65	**0.21**
Spatial breadth†	-0.06	-0.01	84.54	0.00
X-linkage	-0.68	-0.01	98.44	0.01
Adult	-0.79	0.03	9.61	0.03
Reproductive†	-0.03	0.04	-45.95	0.03

Figure 3 Effect estimates for M1. Male effect estimates (M-N), female effect estimates (F-N), and the difference between effects estimates (M-F) are shown (from top to bottom within each plot). Mean differences in comparison groups before adjustment (solid line) and after adjustment are shown. Effect estimates based on branch-specific estimates (left) and genus-wide estimates of evolutionary rate (right) are shown.

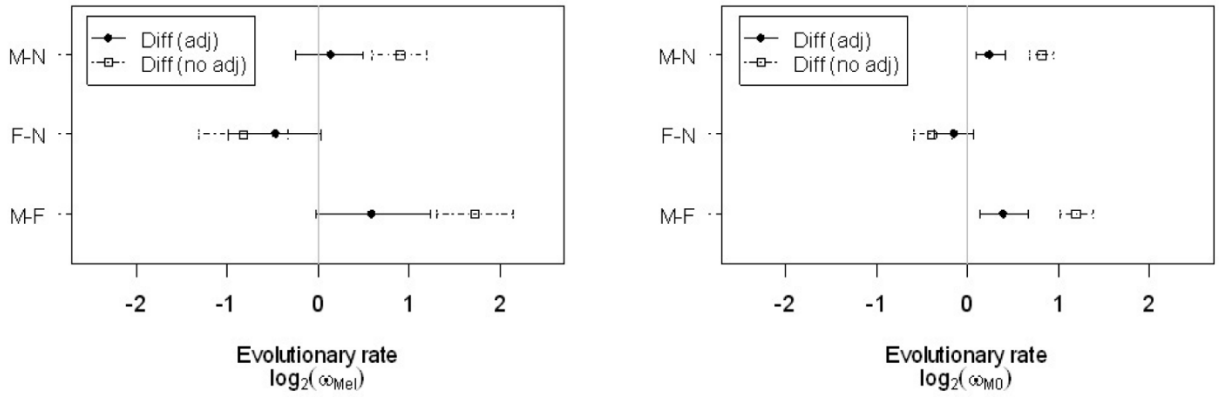


Table 6 Effect estimates for M1. From left to right, table shows outcome variables, comparisons, sample sizes, mean effect estimates prior to adjustment along with associated simultaneous confidence intervals based on bootstrapping approximation described in methods section.

Outcome	**Comp.**	**(N_F,N_M,N_N)**	**Diff. (no adj.)**	**C.I. (no adj.)**	**Diff. (adj.)**	**C.I. (adj.)**
ω_{mel}	M-N	(0,915,6046)	0.90	(0.60,1.20)	0.14	(-0.25, 0.50)
	F-N	(324,0,6046)	-0.82	(-1.32, -0.33)	-0.46	(-0.99,0.03)
	M-F[a]	(324,915,6046)	1.72	(1.30,2.14)	0.60	(-0.03,1.23)
ω_{M0}	M-N	(0,915,6046)	0.81	(0.68,0.93)	0.25	(0.08,0.41)
	F-N	(324,0,6046)	-0.38	(-0.59,-0.19)	-0.14	(-0.37,0.07)
	M-F[a]	(324,915,6046)	1.19	(1.01,1.37)	0.40	(0.13,0.66)

Table 7 Covariate balance for (M2). From left to right, list of covariates, mean difference prior to adjustment, mean difference after adjustment, the percentage change in the mean difference before and after adjustment, and the balance score (β). Proposed confounders included in matching criteria for model selection are indicated (†).

Table 7a Covariate balance for M vs. N (M2)

Covariates	Diff. (no adj.)	Diff. (adj.)	Perc. Change	β
PPI†	0.08	-0.03	60.95	0.00
Gene expression†	-0.03	0.00	98.22	0.01
Recombination†	-0.07	-0.07	-7.17	0.06
Protein length†	-119.46	-10.60	91.12	0.03
Intron number†	-0.80	0.30	62.69	0.10
Spatial breadth†	0.29	0.01	96.53	0.03
X-linkage†	-0.83	0.06	93.03	0.06
Adult	0.08	-0.43	-465.34	0.43
Reproductive†	0.38	0.01	97.28	0.01

Table 7b Covariate balance for F vs. N (M2)

Covariates	Diff. (no adj.)	Diff. (adj.)	Perc. Change	β
PPI†	2.54	0.67	73.58	0.09
Gene expression†	0.02	0.00	84.85	0.03
Recombination†	-0.06	-0.03	40.94	0.03
Protein length†	-31.89	15.69	50.78	0.04
Intron number†	-0.55	-0.16	70.97	0.06
Spatial breadth†	-0.06	0.02	69.72	0.06
X-linkage†	-0.68	0.00	99.39	0.00
Adult	-0.79	0.02	97.15	0.03
Reproductive†	-0.03	0.03	1.89	0.03

Figure 4 Effect estimates for M2. Male effect estimates (M-N), female effect estimates (F-N), and the difference between effects estimates (M-F) are shown (from top to bottom within each plot). Mean differences in comparison groups before adjustment (solid line) and after adjustment are shown. Effect estimates based on branch-specific estimates (left) and genus-wide estimates of evolutionary rate (right) are shown.

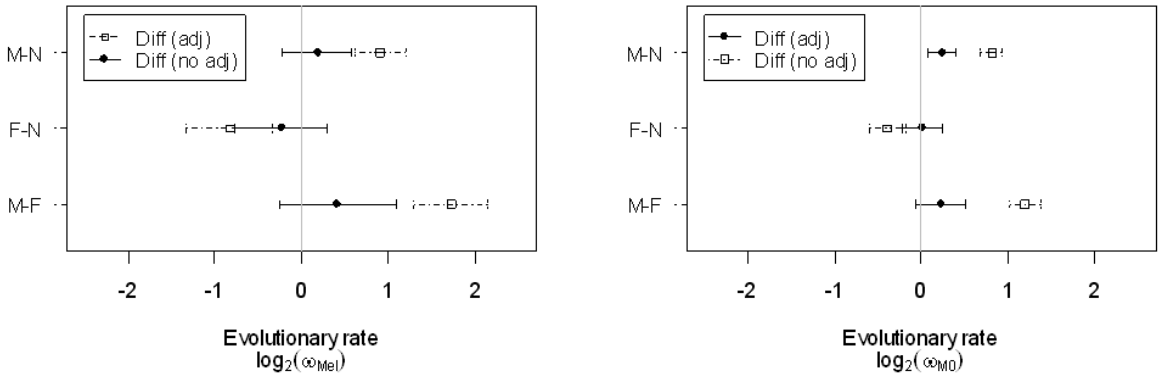


Table 8 Effect estimates for M2 From left to right, table shows outcome variables, comparisons, sample sizes, mean effect estimates prior to adjustment along with associated simultaneous confidence intervals based on bootstrapping approximation described in methods section. Indirect adjusted comparisons are indicated ([a]).

Outcome	Comp.	(N_F,N_M,N_N)	Diff. (no adj.)	C.I. (no adj.)	Diff. (adj.)	C.I. (adj.)
ω_{mel}	M-N	(0,915,6046)	0.90	(0.60,1.20)	0.19	(-0.23, 0.57)
	F-N	(324,0,6046)	-0.82	(-1.32, -0.33)	-0.23	(-0.80,0.30)
	M-F[a]	(324,915,6046)	1.72	(1.30,2.14)	0.41	(-0.26,1.08)
ω_{M0}	M-N	(0,915,6046)	0.81	(0.68,0.93)	0.24	(0.06,0.40)
	F-N	(324,0,6046)	-0.38	(-0.59,-0.19)	0.02	(-0.21,0.23)
	M-F[a]	(324,915,6046)	1.19	(1.01,1.37)	0.22	(-0.06,0.51)

Table 9 Covariate balance for M3. From left to right, list of covariates, mean difference prior to adjustment, mean difference after adjustment, the percentage change in the mean difference before and after adjustment, and the balance score (β). Proposed confounders included in matching criteria for model selection are indicated (†).

Table 9a Covariate balance for M versus N (M3)

Covariates	Diff. (no adj.)	Diff. (adj.)	Perc. Change	β
PPI†	-0.40	0.01	98.53	0.00
Gene expression†	0.00	0.00	55.10	0.02
Recombination†	-0.07	0.00	97.56	0.00
Protein length†	-119.92	-12.92	89.23	0.03
Intron number†	-0.47	0.32	31.61	0.09
Spatial breadth†	0.25	0.00	99.26	0.01
X-linkage†	-0.84	0.07	91.58	0.07
Adult†	0.08	-0.02	67.33	0.02
Reproductive†	0.21	-0.01	94.37	0.01

Table 9b Covariate balance for F versus N (M3)

Covariates	Diff. (no adj.)	Diff. (adj.)	Perc. Change	β
PPI†	2.54	0.67	73.58	0.09
Gene expression†	0.02	0.00	84.45	0.03
Recombination†	-0.06	-0.03	40.94	0.03
Protein length†	-31.89	15.69	50.78	0.04
Intron number†	-0.55	-0.16	70.97	0.06
Spatial breadth†	1.03	0.02	69.72	0.06
X-linkage†	-0.68	0.00	99.39	0.00
Adult†	-0.79	0.02	97.15	0.02
Reproductive†	-0.03	0.03	1.89	0.03

Figure 5 Effect estimates for M3. Male effect estimates (M-N), female effect estimates (F-N), and the difference between effects estimates (M-F) are shown (from top to bottom within each plot). Mean differences in comparison groups before adjustment (solid line) and after adjustment are shown. Effect estimates based on branch-specific estimates (left) and genus-wide estimates of evolutionary rate (right) are shown.

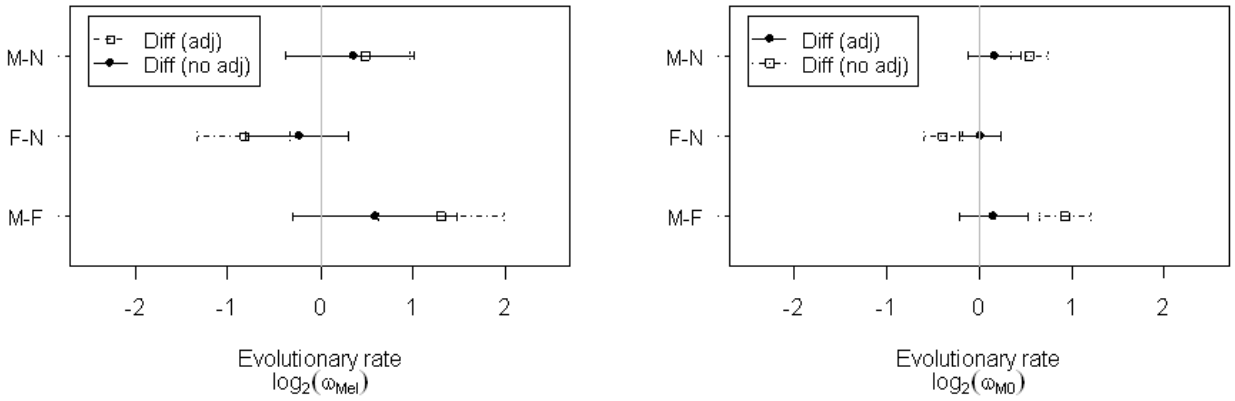


Table 10 Effect estimates for M3 From left to right, table shows outcome variables, comparisons, sample sizes, mean effect estimates prior to adjustment and after adjustment along with simultaneous confidence intervals based on bootstrapping approximation described in methods section. Indirect adjusted comparisons are indicated ([a]).

Outcome	Comp.	(N_F,N_M,N_N)	Diff. (no adj.)	C.I. (no adj.)	Diff. (adj.)	C.I. (adj.)
ω_{mel}	M-N	(0,344,2580)	0.49	(0.00,0.97)	0.36	(-0.37, 1.00)
	F-N	(324,0,6046)	-0.82	(-1.32, -0.33)	-0.23	(-0.80,0.30)
	M-F[a]	(324,344,6046/2580)	1.30	(0.62,1.99)	0.59	(-0.30,1.48)
ω_{M0}	M-N	(0,355,2580)	0.54	(0.34,0.74)	0.15	(-0.13,0.45)
	F-N	(324,0,6046)	-0.38	(-0.59,-0.19)	0.02	(-0.21,0.23)
	M-F[a]	(324,344,6046/2580)	0.92	(0.64,1.20)	0.15	(-0.21,0.52)

Table 11 Sensitivity analyses for ω_{mel} for M1-M3 (top-bottom) For each sub-table (a-c), from left to right: a threshold balancer increasing in fixed steps of size 0.001 (Γ_B); the corresponding range of effect estimates, including lower and upper values, and their difference (Δ); and the number of subclass simulations for which the balancer is below the threshold (N).

Sensitivity analysis for M1 (ω_{mel})

	M vs. N					F vs. N			
Γ_B	**Lower**	**Upper**	**Δ**	**N**	**Γ_B**	**Lower**	**Upper**	**Δ**	**N**
0.035	0.14	0.14	0.00	2	0.036	-0.46	-0.46	0.00	1
0.036	0.14	0.14	0.00	2	0.037	-0.46	-0.46	0.00	31
0.037	0.14	0.14	0.00	2	0.038	-0.46	-0.46	0.00	31
0.038	0.14	0.14	0.00	7	0.039	-0.46	-0.46	0.00	31
0.039	0.14	0.14	0.00	7	0.04	-0.46	-0.46	0.00	31
0.040	0.14	0.15	0.01	14	0.041	-0.46	-0.46	0.00	31
0.041	0.14	0.15	0.01	24	0.042	-0.50	-0.46	0.03	63
0.042	0.14	0.15	0.01	24	0.043	-0.50	-0.46	0.03	63
0.043	0.13	0.15	0.01	26	0.044	-0.50	-0.46	0.04	84
0.044	0.13	0.15	0.01	27	0.045	-0.50	-0.46	0.04	94
0.045	0.13	0.15	0.01	36	0.046	-0.50	-0.46	0.04	100

Sensitivity analysis for M2 (ω_{mel})

	M vs. N					F vs. N*			
Γ_B	**Lower**	**Upper**	**Δ**	**N**	**Γ_B**	**Lower**	**Upper**	**Δ**	**N**
0.037	0.19	0.19	0.00	1	0.040	-0.26	-0.26	0.00	1
0.038	0.19	0.21	0.03	9	0.041	-0.26	-0.26	0.00	3
0.039	0.19	0.22	0.04	51	0.042	-0.28	-0.24	0.04	34
0.040	0.19	0.22	0.04	60	0.043	-0.28	-0.24	0.04	126
0.041	0.19	0.22	0.04	61	0.044	-0.32	-0.23	0.09	203
0.042	0.19	0.22	0.04	63	0.045	-0.32	-0.23	0.09	244
0.043	0.19	0.22	0.04	74	0.046	-0.34	-0.23	0.11	250
0.044	0.19	0.22	0.04	78	0.047	-0.35	-0.23	0.12	261
0.045	0.19	0.22	0.04	78	0.048	-0.35	-0.23	0.12	342
0.046	0.19	0.22	0.04	81	0.049	-0.38	-0.23	0.15	362
0.047	0.19	0.22	0.04	88	0.050	-0.41	-0.23	0.18	365

Sensitivity analysis for M3 (ω_{mel})

	M vs. N					F vs. N			
Γ_B	**Lower**	**Upper**	**Δ**	**N**	**Γ_B**	**Lower**	**Upper**	**Δ**	**N**
0.032	0.36	0.36	0.00	2	0.040	-0.26	-0.26	0.00	1
0.033	0.36	0.36	0.00	5	0.041	-0.26	-0.26	0.00	3
0.034	0.36	0.36	0.00	5	0.042	-0.28	-0.24	0.04	34
0.035	0.36	0.36	0.00	5	0.043	-0.28	-0.24	0.04	126
0.036	0.36	0.36	0.00	8	0.044	-0.32	-0.23	0.09	203
0.037	0.36	0.36	0.00	8	0.045	-0.32	-0.23	0.09	244
0.038	0.36	0.36	0.00	8	0.046	-0.34	-0.23	0.11	250
0.039	0.36	0.36	0.00	8	0.047	-0.35	-0.23	0.12	261
0.040	0.36	0.36	0.00	8	0.048	-0.35	-0.23	0.12	342
0.041	0.36	0.37	0.00	12	0.049	-0.38	-0.23	0.15	362
0.042	0.36	0.37	0.00	12	0.050	-0.41	-0.23	0.18	365

Table 12 Sensitivity analyses for ω_{M0} for M1-M3 (top-bottom). For each sub-table (a-c), from left to right: a threshold balancer increasing in fixed steps of size 0.001 (Γ_B); the corresponding range of effect estimates, including lower and upper values, and their difference (Δ); and the number of subclass simulations for which the balancer is below the threshold (N).

Sensitivity analysis for M1 (ω_{M0})

M vs. N					F vs. N				
Γ_B	Lower	Upper	Δ	N	Γ_B	Lower	Upper	Δ	N
0.035	0.25	0.25	0.00	2	0.036	-0.14	-0.14	0.00	1
0.036	0.25	0.25	0.00	2	0.037	-0.14	-0.14	0.00	31
0.037	0.25	0.25	0.00	2	0.038	-0.14	-0.14	0.00	31
0.038	0.25	0.25	0.00	7	0.039	-0.14	-0.14	0.00	31
0.039	0.25	0.25	0.00	7	0.040	-0.14	-0.14	0.00	31
0.040	0.25	0.25	0.00	14	0.041	-0.14	-0.14	0.00	31
0.041	0.25	0.25	0.00	24	0.042	-0.16	-0.14	0.01	63
0.042	0.25	0.25	0.00	24	0.043	-0.16	-0.14	0.01	63
0.043	0.25	0.26	0.01	26	0.044	-0.16	-0.14	0.02	84
0.044	0.25	0.26	0.01	27	0.045	-0.16	-0.14	0.02	94
0.045	0.25	0.26	0.01	36	0.046	-0.16	-0.14	0.02	100

Sensitivity analysis for M2 (ω_{M0})

M vs. N					F vs. N				
Γ_B	Lower	Upper	Δ	N	Γ_B	Lower	Upper	Δ	N
0.037	0.24	0.24	0.00	1	0.04	0.00	0.00	0.00	1
0.038	0.24	0.24	0.00	9	0.041	0.00	0.00	0.00	3
0.039	0.24	0.25	0.01	51	0.042	-0.02	0.01	0.03	34
0.040	0.24	0.25	0.01	60	0.043	-0.02	0.01	0.03	126
0.041	0.24	0.25	0.01	61	0.044	-0.05	0.02	0.06	203
0.042	0.24	0.25	0.01	63	0.045	-0.05	0.02	0.06	244
0.043	0.24	0.25	0.01	74	0.046	-0.06	0.02	0.08	250
0.044	0.24	0.25	0.01	78	0.047	-0.07	0.02	0.08	261
0.045	0.24	0.25	0.01	78	0.048	-0.07	0.02	0.08	342
0.046	0.24	0.26	0.02	81	0.049	-0.08	0.02	0.10	362
0.047	0.24	0.26	0.02	88	0.050	-0.08	0.02	0.10	365

Sensitivity analysis for M3 (ω_{M0})

M vs. N					F vs. N				
Γ_B	Lower	Upper	Δ	N	Γ_B	Lower	Upper	Δ	N
0.032	0.17	0.17	0.00	2	0.040	0.00	0.00	0.00	1
0.033	0.17	0.17	0.00	5	0.041	0.00	0.00	0.00	3
0.034	0.17	0.17	0.00	5	0.042	-0.02	0.01	0.03	34
0.035	0.17	0.17	0.00	5	0.043	-0.02	0.01	0.03	126
0.036	0.17	0.18	0.01	8	0.044	-0.05	0.02	0.06	203
0.037	0.17	0.18	0.01	8	0.045	-0.05	0.02	0.06	244
0.038	0.17	0.18	0.01	8	0.046	-0.06	0.02	0.08	250
0.039	0.17	0.18	0.01	8	0.047	-0.07	0.02	0.08	261
0.040	0.17	0.18	0.01	8	0.048	-0.07	0.02	0.08	342
0.041	0.17	0.18	0.01	12	0.049	-0.08	0.02	0.10	362
0.042	0.17	0.18	0.01	12	0.050	-0.08	0.02	0.10	365

Chapter 3

Phylogenomic analysis of sex-dependent gene expression to assess intra-locus sexual conflict in *Drosophila*

ABSTRACT

Phylogenomics provides a powerful framework for studying the evolution of sex-biased gene expression. Ancestral inference on sex-dependent gene expression is used here to empirically test theoretical expectations about how sex-biased gene expression should arise, to identify candidate genes involved in ongoing intersexual developmental conflict, and to study the role of expression evolution on the 'demasculinization' of X chromosomes. Functional analysis of candidate genes suggests that enhanced selection on male secondary sexual traits may be a common source of sexually antagonistic developmental conflict in *Drosophila*. Modeling of transition dynamics among states of sex-biased gene expression for X-linked and autosomal genes indicates that expression evolution can account for the underrepresentation of male-biased genes on the X chromosome. Analysis of transition dynamics for candidate genes provides support for the hypothesis that the underrepresentation of male-biased genes on the X chromosome is in part due to the effects of sexually antagonistic coevolution, although not to the exclusion of other proposed causes.

INTRODUCTION

Multi-species genome-wide assays of sex-biased gene expression have demonstrated that genes can exhibit different forms of sex-biased gene expression across species; that is, sex-biased gene expression is a remarkably labile trait (Ranz et al. 2003). Phylogenomic analysis of sex-biased expression is an as yet underutilized approach for extracting biological information about the evolution of sex differences from microarray datasets. Available data on genome-wide sex-biased gene expression for seven species of *Drosophila* permits the use of phylogenetic methods on a genome-wide scale. Ancestral inference using multi-species sex expression data can be used to empirically test theoretical predictions about the evolution of sexual dimorphism, and may be useful for identifying candidate genes involved in intersexual developmental conflict.

Theoretical considerations suggest that sex-biased expression is likely to evolve when expression is advantageous in one sex and neutral or costly in the other (Rhen 2000, Rice 1984). In response to the appearance of an allele with sexually antagonistic fitness effects, alternative scenarios may occur. If expression is costly in the disfavored sex, and function is dispensable, selection may cause expression to be down-regulated in the disfavored sex to the point of sex-limited expression, thereby eliminating detrimental fitness effects. However, when conflicts arise over a shared function, resolution through the evolution of sex-limited expression may not be possible (Connallon and Knowles 2005). As evolutionary opportunity permits, a tug-of-war may

ensue over optimal expression and or/ function as mean population fitness is maximized alternately at the expense of one sex or the other (Rice and Chippindale 2001).

Gains and losses of sex-biased gene expression between species may involve disproportionate changes in male and female expression (Connallon and Knowles 2005). Knowledge of the ancestral and derived states of sex-biased gene expression (male-, female-, or nonsex-biased) is insufficient for determining whether a gain or loss of sex-biased gene expression has primarily involved sex expression divergence in the biased or the unbiased sex. In order to determine this, it is necessary to infer ancestral expression for both sexes separately. A loss of male-biased gene expression (or a transition from an ancestral state in which male-biased gene expression is inferred to a derived state for which nonsex-biased gene expression is inferred) may be described by one of three scenarios: it may primarily involve a loss of expression in the male sex, a gain of expression in the female sex, or may involve a roughly equal combination of changes in both sexes (**Fig 6**). To assess relative expression change and to discriminate between scenarios of gain or loss of expression, a comparative sex expression index is introduced. The comparative sex index is the normalized signed difference in female and male expression divergence:

$$CSI = \frac{|\Delta F| - |\Delta M|}{\sqrt{\Delta F^2 + \Delta M^2}},$$

If CSI >> 0, evolutionary change in the state of sex-biased gene expression has primarily involved expression divergence in the female sex, whereas if CSI << 0, change in the state of sex-biased gene expression has primarily involved divergence in the male sex.

Multi-species microarray data on sex expression permits ancestral inference across the *Drosophila* phylogeny. Inferred male and female expression values can be used to assign discrete states of sex-biased expression (**M**, **F**, or **N**, indicating male-, female-, or nonsex-biased expression respectively) to each ancestral node and a discrete state of the sex comparative index (**m**, **f**, or **n)**, indicating disproportionate female, male divergence, or comparable levels of divergence in both sexes) to each branch. Assuming a Brownian model of both male and female expression, the sign of the sex comparative index is not expected to depend on the state of sex-biased gene expression at the endpoints of each branch. In order to test how sex-biased gene expression typically evolves, the frequency of the sign of comparative index across multiple gene families can be evaluated for each transition type and compared to neutral expectations. This framework can be used to study how sex-biased gene expression typically evolves; that is to discriminate between alternative scenarios such as those depicted in **Fig 6**.

In addition to permitting study of the evolution of sex-biased gene expression, it is proposed that enhanced lability on the comparative sex index, as measured in terms of the frequency of transitions over a phylogeny, may be a useful criterion for identifying genes and gene functions that are likely to experience intersexual developmental conflicts over evolutionary time. Intra-locus sexual conflict occurs when there is a negative correlation between the selection coefficients of the same allele when expressed in either sex. When there are sexual conflicts over phenotypic optima related to a shared gene function, and where sex-specific fitness optima cannot be mutually attained, an allele that is detrimental to one sex may be selected for if advantages outweigh costs when averaged over the population. Conflicts of this kind may perpetuate evolutionary change as opportunity permits: a shift towards the male optimum results in increased selection in females for mutations that shift back toward the female optimum (Rice and Chippindale 2001). It is expected that this process may also lead to enhanced lability of the comparative sex index, since expression is also expected to diverge when there are antagonistic

fitness effects (Rhen 2000). Several tree-based lability statistics are devised to measure the extent of tug-of-war dynamics occurring over evolutionary time for a single orthologous gene family. Since both male and female expression are assumed to evolve according to a simple Brownian model of evolution, the sign of the sex comparative index is expected to be binomially distributed across the branches of a phylogeny. Candidate genes that may be experiencing ongoing intra-locus sexual conflict can be selected by comparing observed values on these lability statistics to expected values based on a null model. Candidate genes can be further studied to assess the role of antagonistic coevolution on various aspects of genome evolution, including demasculinization.

Using additional functional and evolutionary genomic data, candidate gene sets can be evaluated for consistency with several other predictions that have been made for sexually antagonistic genes. Because the fixation rate of alleles with sexually antagonistic fitness effects is supposed to be enhanced for both recessive male-favoring and dominant female-favoring antagonistic genes (Charlesworth et al. 1987), the X-chromosome is expected to harbor sexually antagonistic fitness variation. If the X chromosome harbors sexually antagonistic fitness variation, labile genes may be more likely to be X-linked and may be more labile when X-linked. Candidate genes should be overrepresented among genes involved in reproduction and expressed at the adult stage of life where sex differential fitness pressure is expected to be the most intense. Although there is evidence that antagonistic coevolution leads to enhanced rates of sequence evolution between species (Paterson et al. 2010) and that sexual conflicts may have led to rapid evolution of reproductive genes (Swanson and Vacquier 2002, Gavrilets 2000, Panhuis and Swanson 2006, Swanson et al. 2001), whether sequence divergence is expected to be enhanced for genes specifically involved in intra-locus *developmental* conflicts is less clear, since these genes are expected to experience opposing selection pressures depending on the sex in which they are expressed and the direction of sex-differential fitness effects.

It has been proposed that developmental conflicts may be quite common for genes involved in reproduction because mutations that are beneficial to one sex may reduce fitness in the other sex through negative pleiotropic effects (Rice and Chippindale 2001). When gene function is not shared, sex-limitation may readily evolve since there is selective pressure to reduce antagonistic fitness consequences in the alternative sex. This is likely to result in a loss of expression in the disfavored sex. However, even when function is not shared, opportunities may exist for ongoing antagonistic coevolution, particularly if sex-limitation is incomplete, for example, for X-linked genes that escape dosage compensation. Furthermore, the evolution of sex-limited expression may be less likely to occur for broadly expressed genes, since it is thought to be increasingly unlikely that an allele will confer net positive sex differential fitness as the potential for negative pleiotropy is increased (Mank et al. 2008). If this is the case then genes that show sex-biased gene expression should be more narrowly expressed than genes that are not sex-biased and ongoing conflicts may persist especially among broadly expressed genes that are nonsex-biased. Functional and genomic analysis of candidate genes may help to identify gene for which developmental conflicts are most pervasive.

Phylogenomic analysis of sex-biased gene expression can also be used to evaluate the contribution of gene expression evolution to the widely observed phenomena whereby male-biased genes are underrepresented on the X chromosome for several species of *Drosophila* (Parisi et al. 2003b, Zhang et al. 2007). Although selective retrotransposition of male-biased genes from X-linked to autosomal regions (Khil, Oliver and Camerini-Otero 2005, Betran et al. 2002), as well as differential birth and death rates of male-biased genes (Zhang et al. 2007,

Sturgill et al. 2007) are known to contribute to the underrepresentation of X-linked male-biased genes, it has been suggested that expression evolution may not, mainly because changes in sex-biased gene expression are not observed among genes that have experienced chromosomal translocations (Betran et al. 2002, Sturgill et al. 2007). However, the evolution of sex-biased gene expression across the genome is a dynamic process, and has yet to be modeled as such. Phylogenetic inference on sex-biased gene expression can be used to evaluate whether expression evolution also contributes to the 'demasculinization' of X chromosomes.

Proposed causes for the underrepresentation of male-biased are as multifarious as the mechanisms by which it is affected. Evidence has been advanced to support several hypotheses about why male-biased genes may be disfavored on the X chromosome, including interference of dosage compensation mechanisms with male-biased gene expression (Bachtrog et al. 2010), and inhibiting effects of X-inactivation on male-biased genes expressed during early spermatogenesis (Parisi et al. 2003a). It has also been speculated that sexually antagonistic coevolution may play a role in affecting genomic patterns of sex-biased gene expression (Connallon and Knowles 2005, Oliver and Parisi 2004, Vicoso and Charlesworth 2006). Theoretical considerations suggest that X-linkage may facilitate antagonistic coevolution (Jonathan et al. 2002), since both dominant female-benefiting and recessive male-benefiting antagonistic alleles are expected to fix at higher rates on the X chromosome. However, since recessive mutations are thought to be more common in general, the hypothesis that antagonistic coevolution may also contribute to the underrepresentation of male-biased genes on the X chromosome has been in doubt (Oliver and Parisi 2004).

Explanations for the underrepresentation of male-biased genes on X chromosomes have generally aimed to explain differential birth and death dynamics of X-linked male-biased genes (for exception, see Bachtrog et al. 2010). If dynamic trends in expression evolution also play a role in the 'demasculinization' phenomenon, several additional hypotheses can be considered. Directional trends in expression evolution that lead to biased transference of genes from the male-biased state to a nonsex-biased or female-biased state may lead to 'demasculinization'. If the male environment facilitates the fixation of new mutations either because of enhanced selection or because of a relatively reduced potential for negative pleiotropy (see Chapter 1), novel alleles may commonly rise to fixation in an antagonistic male-benefiting form. Even if initially sex-limited in expression, adaptive processes occurring in males may potentially lead to reduced sexual antagonism over evolutionary time. Transference (specifically of the form of **MN** transitions) may be more common on the X by the general 'faster-X' mechanisms. If transference is indeed more common on the X, then loss of male-biased gene expression events may be more likely to involve gain of expression in the female sex for X-linked genes. Ongoing antagonistic coevolution may also lead to trends in expression evolution over evolutionary time. Further study of expression trends for candidate genes may provide insights about the impact of antagonistic coevolution on genome evolution.

METHODS

Ancestral Inference

Using pre-computed analysis of orthology available from FlyBase, a total of 4096 orthologous gene families were selected for phylogenomic analysis. Study was confined to orthologous gene families for which a single member could be identified in each of the seven species for which sex-biased expression data is available, including: *D. melanogaster, D. simulans, D. yakuba, D. ananassae, D. pseudoobscura, D. virilis, and D. mojavensis.* Since gene

duplication and loss events may be associated with rapid evolution (Powell 1997), care was taken to eliminate gene families showing evidence of duplication or loss in any of the 12 sequenced species belonging to the *Drosophilidae.* Also, gene families involved in X-autosomal translocation events were specifically removed from the analysis out of concern inadequacies of dosage compensation could obscure the patterns under study here. In all, 3,505 gene families were considered.

Ancestral male and female expression values were inferred in a maximum likelihood framework assuming a phylogenetic model of gene expression evolution. For each orthologous gene family, ancestral male and female expression values were estimated *separately.* Given the species tree relating the seven *Drosophila* species and male or female expression data for the species at each tip, ancestral values were estimated under a Brownian motion model by maximizing the following likelihood function, parameterized in terms of the ancestral values (x_a) and the variance of the Brownian motion process (σ^2, interpreted as the mutational variance) (Schluter et al. 1997):

$$L(\sigma^2, \boldsymbol{x_a} | \dagger, \boldsymbol{x}) = \frac{1}{\sigma^n} \exp\left(\frac{1}{2\sigma} \sum \frac{(x_i - x_j)^2}{t_{ij}}\right) \qquad \textbf{[1]}$$

Here, † is the seven-species phylogeny and $\boldsymbol{x}$ corresponds to observed male or female expression values as measured in each of the seven species. The sum is taken over all contrasts (of which there are $n - 1$ for a tree including n species), where x_i and x_j correspond to the expression values at the node or tip defined by each contrast, and t_{ij} refers to the span of evolutionary time separating the nodes or tips defined by the contrast. The species phylogeny relating the seven species used in this analysis was obtained from FlyBase (Powell 1997). Ancestral inference was implemented in R using the **ape** package (Paradis, Claude and Strimmer 2004).

States of sex-biased gene expression (**N, F**, or **M**) and the sex comparative index (**n, f, or m**) were assigned based on randomization tests conducted on data derived from Zhang *et al.*'s multi-species microarray study of sex-biased gene expression. A full description of the methods used can be found in **Appendix A**.

Evolution of sex-biased gene expression

The evolution of sex-biased gene expression was examined by considering the probability of a positive or negative index, as conditioned on each of the different types of transition between states of sex-biased gene expression. Transition types are defined for each branch and include: gains of sex-biased gene expression, **NM** and **NF** (where the state of sex-biased gene expression is nonsex-biased in the ancestral state); losses of sex-biased gene expression, **MN** and **FN** (where the state of sex-biased gene expression is nonsex-biased in the descendent species); maintenance of sex-biased gene expression, **MM, FF** or **NN**; and switches between sex-biased states, **MF** and **FM**. Under the null model, the sign of the index is expected to be binomially distributed across the branches of the tree, and, for a given branch, does not a priori depend upon the states of the sex-biased gene expression that are inferred at the ancestral and derived nodes. Deviations from binomial expectations may reflect biologically-based biases in the way that sex expression typically evolves. These can be analyzed at a variety of taxonomic levels. Where changes in the state of expression occur (all cases but **MM, FF** and **NN**), the sign of the index provides information about the net direction of change. For example,

if there is a loss of male expression event **MN**, and the sign of the comparative sex index is positive, it can be deduced that expression has been gained in the female sex. By contrast, if the index is negative, indicating disproportionate expression change for the male sex, expression has been gained in the male sex.

Candidate gene analysis

A lability index (λ) was used to select candidate genes for further analysis. The lability index is evaluated as the sum of the number of changes in the sign of the sex comparative index evaluated for each lineage and summed over all 7 lineages. Specifically, for each of the 7 lineages under consideration, transition opportunities were counted at each internal node if the index differed in sign for the branches adjacent to the node in question. For the *D. melanogaster* lineage, there are four opportunities for transition corresponding to the total number of internal nodes from tip to root (the root node is excluded), while for *D. virilis* there is a single opportunity for transition (**Fig 7**). There 16 opportunities for transitions, where the total number of opportunities is the sum of the number of transition opportunities across all 7 species lineages. Note that this procedure involves repeated counting of transitions occurring at the most basal nodes. For example, if a transition occurs at node 9, then it will be counted 5 times since it appears in the paths linking 5 of the species (*D. pseudoobscura*, *D. ananassae*, *D. yakuba*, *D.simulans*, and *D. melanogaster*) to the root. This has the effect of more heavily weighting transitions that occur more basally in the tree, as well as more heavily weighting the evolutionary histories associated with species that are removed from the root by a larger number of internal nodes. Alternative measures of lability on the sex comparative index were also considered. For example, a conservation index can be calculated as the absolute value of the difference in the number of branches associated with a positive index and the number of branches associated with a negative index. The conservation index ranges from 12 to 0 (where 12 is the total number of branches in the phylogeny). High values of the conservation index are indicate that the sign of the sex comparative index is the same throughout most or all of the tree, while low values suggest more disparity in the sign of index. The conservation index can be further generalized for the case of multiple states, $x_1, \ldots, x_n$, where $\lambda = \sum_{i \neq j} \left| I_{x_i} - I_{x_j} \right|$, where I_{x_i} is denotes the count of the total number of branches assigned to state x_i across the phylogeny. Although this statistic does not bias the evolutionary history of any single species lineage over another, it cannot resolve between sustained conflicts involving multiple transitions occurring along a continuous line of descent and those that occur only once in several different lineages. Since instances of multiple transitions occurring over a continuous line of descent are thought to be indicative of ongoing intra-locus sexual conflicts, a path-based statistics was ultimately chosen as the basis for selection of candidate genes.

The lability index can also be calculated in terms of the sign of the comparative index. Under the null model, each branch has an equal likelihood $(1/2)$ of having a negative or a positive index so the expected value of the lability index is 8. Alternatively, each branch can be coded based on hypothesis testing on the sex comparative index. In this case, three transition types are possible ($\boldsymbol{n} \leftrightarrow \boldsymbol{f}, \boldsymbol{m} \leftrightarrow \boldsymbol{f}, \boldsymbol{n} \leftrightarrow \boldsymbol{m}$) with probabilities that are based on the significance cut-off used for hypothesis testing (α, for two-tailed test). Under null expectations, the probability that any branch is in the **f** or **m** state is $\alpha/2$, and the probability of being **n** is $(1 - \alpha/2)$. If the neutral state is ignored in calculating the lability index—that is, if one counts a lineage **m-n-f** as having a single transition but **m-n-m** as having none—then $\lambda_{exp} = \alpha^2/2[10 - 2 \cdot \alpha + 10 \cdot \alpha^2 - 2 \cdot \alpha^3] \approx 0.0005$, where a false-discovery error rate of α = 0.01 is used (See

Appendix C for derivation of λ_{exp}). To reduce the potential for genes with noisy fluctuations on the comparative sex index to be overrepresented among candidate genes, candidacy was determined on the basis of assigned states of the sex comparative index. Specifically, candidate genes were assigned to the labile class if $\lambda > \lambda_{exp} = 0.0005$.

Expression evolution as a dynamic process

In order to study the expression evolution as a dynamic process and to compare X-linked and autosomal-linked genes, transition probabilities were estimated for each lineage from the ancestral states of sex-bias inferred for the root and the derived states of sex-biased gene expression. A probability transition matrix was determined by counting the total number of genes associated with each of the possible transitions (for example, **FF** is the number of genes inferred to have female-biased gene expression at the root and female-biased gene expression at a tip), arranging them into a 3 x 3 matrix and dividing by the sum of each row:

$$\begin{bmatrix} \frac{FF}{FF+FM+FF} & \frac{FM}{FF+FM+FF} & \frac{FN}{FF+FM+FF} \\ \frac{MF}{MF+MM+NM} & \frac{MM}{MF+MM+NM} & \frac{MN}{MF+MM+NM} \\ \frac{NF}{NF+NM+NN} & \frac{NF}{NF+NM+NN} & \frac{NF}{NF+NM+NN} \end{bmatrix} = \begin{bmatrix} P_{FF} & P_{FM} & P_{FN} \\ P_{MF} & P_{MM} & P_{MN} \\ P_{NF} & P_{NM} & P_{NN} \end{bmatrix}$$

For each lineage, separate probability matrices were calculated for X-linked genes and autosomal genes and their associated stationary distributions were compared to assess whether expression evolution alone can account for the underrepresentation of male-biased genes on the X chromosome. Sensitivity analyses were also performed by modifying the p-value cut offs used to determine the states of sex-biased gene expression. Several possible cutoffs were used to evaluate the sensitivity of stationary distribution estimates to the state assignments: 0.01, 0.05, 0.10. No qualitative differences were found for analyses based on these different criteria. Only analysis based on $\alpha = 0.01$ is shown in the main text.

Data on ancestral sex expression values were further integrated with functional and evolutionary genomic data on sequence evolution (Consortium 2007), as well as data on several indicator variables derived from spatial (Chintapalli et al. 2007) and temporal expression (Arbeitman et al. 2002) assays on adult and reproductive expression. A description of how the indicator variables on adult and reproductive expression were derived can be found in the methods section of the first chapter of this dissertation.

RESULTS

The evolution of sex-biased gene expression

The evolution of sex-biased gene expression was studied by considering the distribution of the sign of the index as conditioned on each of the possible transition types. For each transition type and branch, the number of branches inferred to have a positive index was compared to the number of branches inferred to have a negative index and tested for deviation from binomial expectations on the distribution of signs (**Table 13**). Significance testing was conducted at the genus-wide level, where the signs of the index are expected to be binomially distributed under the null model. Genus-wide totals were obtained by summing the counts across all 12 branches of the phylogeny. At the genus wide level, the sign of the index is associated with the transition type (two-tailed Fisher Exact Test, simulated p-value < 0.001).

Significance testing was conducted at the genus-wide level for each transition type based on the signs test. P_{tot} corresponds to Bonferroni adjusted p-values for multiple comparisons across each of the transition types (N = 7). Significant associations were seen between branch and the sign of the index for each transition type, where significance testing was based on Fisher Exact Tests, where $p_{among} < 0.001$ for all transition types. P-values were simulated and adjusted for multiple testing across transition types (N = 7). Within transition types, significance testing was conducted for each branch based on the signs test, where p-values were Bonferroni adjusted for multiple testing across branches (N = 84). In total, 3,505 orthologous gene families were analyzed (corresponding to 42,060 branches). Additionally, counts of the number of genes with positive and negative index were tested for independence across branches.

For each of the transition types analyzed, the sign of the comparative index is highly branch dependent (Chi-squared goodness of fit test for: $p_{adj} < 0.001$ for all transition types, **Table 14**). Genus-wide, inferred gains of sex-biased gene expression typically involve disproportionate expression change in the sex for which sex-biased gene expression is gained, implying that gains of sex-biased gene expression typically involve a loss of expression in the unbiased sex. This is the case both for gains of female-biased gene expression (two-tailed Binomial Test: $p_{\mathbf{tot}} < 0.001$), and for gains of male-biased gene expression (two-tailed Binomial Test: $p_{\mathbf{tot}} = 0.012$). Although gain of male-biased gene expression is more likely to involve loss of expression in the female sex at the genus-wide level, the branch which has *D. melanogaster* at its tip shows the reverse.

Loss of female-biased expression most frequently involves a loss of expression in the female sex for several branches belonging to the *Sophophora* subgenus, including branches with *D. simulans* ($p_{adj} < 0.001$) and *D. ananassae* ($p_{adj} < 0.001$) at the tip. For the *Drosophila* subgenus, including *D.virilis* ($p_{adj} < 0.001$), loss of female-bias more often involves gain of expression in the male. A similar pattern is seen for loss of male-biased gene expression, which typically involves loss of expression in the biased sex for branches in the *Sophophora* subgenus, but more typically involves a gain of expression in the unbiased sex for members of the *Drosophila* subgenus.

For the branches that are not associated with transitions in the state of sex-biased gene expression, the sign of comparative sex index generally indicates that the expression changes that occur within a state of sex-biased gene expression predominately involve expression changes in the biased sex. This pattern is highly significant for male-biased genes across all branches except for that seen for the branch connecting the root to the common ancestor of the *Sophophora*, where the trend is also apparent for *D. pseudoobscura*. For female-biased genes, a positive index is most common ($p_{\mathbf{tot}} < 0.001$), but reversals of the genus-wide pattern also occur. Divergence within the state of nonsex-biased gene expression shows a phylogenetic pattern where divergence is larger in the female sex for members of the *Drosophila* subgenus ($p_{adj} < 0.001$) and is larger in the male sex for the *Sophophora* subgenus ($p_{adj} < 0.001$).

Whether sex-biased gene expression is gained, lost, or maintained, disproportionate divergence in the unbiased sex is more likely if the unbiased sex is the male (**Tables 15-17**). The sign of the index can also be tested for independence with chromosomal location for each transition type. Transitions involving male-biased gene expression were examined for X-effects on the sign of the index, but none were detected (**Table 18**).

The magnitude of the sex comparative index was further compared among genes showing conserved sex-biased gene expression throughout the 7-species phylogeny (**Fig 8b**), as well as among genes expressed maximally in different tissue-types (**Fig 9**). While the mean of the sex comparative index for genes with conserved female-biased gene expression does not differ

significantly from zero for any branch, male-biased genes show disproportionate divergence in the male sex for several branches (**Fig 8b**). Differences were also seen as a function of the tissue of maximal expression. In particular, genes showing maximal expression in sex-limited tissues—including the testis, accessory glands, tubule and ovary—all show deviation from zero, where genes expressed in male sex-limited tissues show the largest deviations from zero. Disproportionate male expression (CSI < 0) is also indicated for genes expressed maximally in the carcass and hindgut.

Spatio-temporal expression and the evolution of sex-biased gene expression

Genes that are inferred to have experienced at least one transition in the state of sex-biased gene expression in recent history are more broadly expressed than conserved genes (Kruskal-Wallis Test indicates heterogeneity among groups: 134.05, df = 3, $p_{adj} < 0.001$) (**Fig 10**). Post-hoc testing comparing conserved and nonconserved groups indicates that genes that transition between states of sex-biased gene expression are more broadly expressed than genes that are conserved (two-tailed Mann Whitney Test: W = 1325910, $p_{adj} < 0.001$). Genes that are consistently female-biased are more broadly expressed than other conserved genes (Mann-Whitney Test: W = 59186, $p_{adj} < 0.001$), while genes with conserved male-biased gene expression are more narrowly expressed than other conserved genes (Mann-Whitney Test: W = 20616, $p_{adj} < 0.001$). P-values were Bonferroni adjusted for multiple testing (N = 3). Distributions on tissue-biased gene expression can be compared for all possible transition types.

It has been proposed that core promoters may evolve more readily in spermatocytes because of the higher levels of RNA polymerase that is expressed there (Swanson and Vacquier 2002), suggesting that male-biased genes may more readily originate as narrowly expressed testis-specific genes. Similarly it has been suggested that genes expressed during oogenesis may originate with enhanced pleiotropic constraint (Zhang et al. 2007). However, differences in spatial expression are not confined to genes expressed either maximally or specifically in the reproductive tissues (**Fig 12**). Similarly, while male-biased genes show reduced stage-biased gene expression as compared to female-biased genes, this difference cannot be restricted to genes expressed in sex-limited tissues.

Candidate gene selection and analysis

It is proposed that enhanced lability on the sex comparative index may be characteristic of genes that undergo tug-of-war dynamics over long spans of evolutionary time. Since there is no clear a priori basis for distinguishing between noisy and biologically relevant sex comparative divergence, lability was analyzed in two ways corresponding to the extremes of interpretation. First, the raw sign of the sex comparative index was used to calculate the lability index. The second more stringent criterion used the state of the sex comparative index as derived from statistical hypothesis testing. For each case, corresponding neutral expectations for the lability index were used as cut-offs to classify genes into labile and non-labile gene classes. Data presented in the main text are derived from analysis based on the more stringent statistical criterion, but a summary of results from candidate genes analysis derived from raw sign of the sex comparative index is included in **Appendix C**.

Labile genes ($\lambda > \lambda_{exp} = \alpha^2/2 = 0.0008$) are not more tissue-biased than non-labile genes (Mann Whitney Test: W = 273618, p-value = 0.3213), but they are significantly more stage-biased than non-labile genes (W = 51863, p-value = 0.002) (**Fig 13**).When lability is assessed in terms of the state of the sex comparative index, X-linked genes are not more labile than

autosomal genes (Kruskal-Wallis test: one-tailed Mann Whitney Test: W = 868185, p = 0.80, H_1: $\lambda_X > \lambda_A$) (**Fig 14, Table 19**), and labile genes ($\lambda > \lambda_{exp} = \alpha^2/2 = 0.0008$) show no evidence of overrepresentation on the X chromosome (one-tailed Fisher Exact Test: $(\text{Odds Ratio})_{X/A} = 0.82$, p = 0.84, H_1; $(\text{Odds ratio})_{X/A} > 1$). Fisher Exact Tests of over-/underrepresentation (two-tailed) were conducted across 30 tissue- and stage-specific expression indicators, where p-values were false discovery rate adjusted (N = 30) (**Table 20**). Genes that are labile on the sex comparative index are overrepresented among genes that are specifically expressed in the carcass, which includes the residual tissues of abdomen and thorax with the gut and sexual tracts removed (p_{adj} = 0.04), and in adult tissues (p_{adj} = 0.001). Labile genes are underrepresented among genes expressed specifically in the brain (p_{adj} = 0.001), thoracic-abdominal ganglion (p_{adj} = 0.002), the larval central nervous system (p_{adj} = 0.006), and in embryonic tissues (p_{adj} = 0.004). Adjusting for tissue-bias, labile genes show no evidence of enhanced rates of gene evolution within any subclass of tissue-biased gene expression (two-tailed Mann Whitney applied within each subclasses of tissue-biased gene expression: H_{low}: $\omega_{lab} = \omega_{non\text{-}lab}$ (W = 32772.5, p-value = 0.14), H_{mid}: $\omega_{lab} = \omega_{non\text{-}lab}$ (W = 30827, p-value = 0.97) H_{high}: $\omega_{lab} = \omega_{non\text{-}lab}$ (W = 30109, p-value = 0.08) (**Fig 15**), nor is the lability index correlated with rates of sequence evolution (Kendall-Tau Test: $\tau = 0.0086$, z = 0.75, p-value = 0.45). Labile genes are more broadly expressed for genes that are consistently nonsex-biased throughout the phylogeny (one-tailed Mann Whitney: W = 7214 ; p = 0.04) (**Fig 16**). The lability index is highly correlated with the number of transitions in sex-biased gene expression (Kendall Tau Test: z = 12.86, p-value < 0.001).

The majority of orthologous gene families, 2531 of 4096 sampled (~62%), show conserved sex-biased gene expression throughout the genus: around 95% (N_N = 2413) show conserved nonsex-biased gene expression, 2% (N_F = 40) show conserved female-biased gene expression, and around 3% show conserved male-biased gene expression (N_M = 78). Of the remaining 48% of orthologous gene families, 54% ($N^T_{F\neg M}$ = 846) are sometimes female-biased (never male-biased), 42% are sometimes male-biased (never female-biased) ($N^T_{M\neg F}$ = 664). Four percent of genes are expressed in both the male and female-biased state ($N^T_{F\cap M}$ = 55).

There is no evidence for differences in lability for genes that show conserved male-biased gene expression as compared to those that exhibit conserved female-biased gene expression throughout the phylogeny (two-tailed Mann-Whitney test: W = 1711, p-value = 0.21; **Fig 17**), nor are conserved male-biased genes more likely to be labile than conserved female-biased genes (two-tailed Fisher Exact Test: Log odds = 0.54, p-value = 0.22) (**Table 21**). However, genes that spend some time in the male-biased state but none in the female-biased state are more labile than sometimes female-biased genes (two-tailed Mann-Whitney Test: W = 340107, p = 0.02; **Fig 12**). Sometimes male-biased genes are more likely to be labile than sometimes female-biased genes (Fisher Exact Test: Log odds = 0.70, p-value = 0.03; **Fig 18, Table 22)**. Sometimes male genes that are labile show enhanced rates of sequence evolution at the adult stage (H_0: $\omega_L = \omega_{NL}$, W = 59.5, p_{adj} = 0.02), while no evidence of enhanced rates are seen for sometimes female-biased genes expressed at the adult stage (H_0: $\omega_L = \omega_{NL}$, W = 7337, p_{adj} = 1.0) (**Fig 19**).

Modeling of the evolution of gene expression

Transition dynamics were compared for X-linked and autosomal linked genes by considering the evolution of states of sex-biased gene expression as a Markov process. Transition matrices were derived for each of the 7 species lineages on the basis of states of sex-biased gene expression inferred for the root of the 7-species phylogeny and those observed at each tip. Transition matrices were derived separately for both X-linked and autosomal genes, as

well as for labile and non-labile genes. Expected distribution of sex-biased genes were evaluated at the stationary distribution and compared for X-linked genes and autosomal genes (**Table 23**). For all species except *D. ananassae* the expected proportions of male-biased genes at equilibrium is higher for autosomal genes than X-linked genes. Using estimated proportions at equilibrium and the observed counts of the number of X-linked genes and autosomal genes based on the dataset (N_X = 591, N_A = 2914), hypothetical contingency tables were evaluated to assess whether expectations at equilibrium are consistent with statistical evaluations of the underrepresentation male-biased genes. Significant underrepresentation of male-biased genes is expected at equilibrium for 5 of the 7 species (**Table 24**), where *D. simulans* shows an insignificant trend, and *D. ananassae* shows overrepresentation instead.

Stationary distributions for X-linked and autosomal genes were further decomposed to estimate the relative contributions of labile and non-labile genes to the estimated underrepresentation of male-biased genes at equilibrium. The relationships between the various quantities used to calculate this contribution of labile and non-labile genes to 'demasculinization' are summarized (**Table 27**). At equilibrium (or at the stationary distribution), the expected difference in the proportion of male-biased genes is often lower for labile than it is for non-labile genes (cf. $\pi^{L}_{X-A}(M)$ and $\pi^{\neg L}_{X-A}(M)$, **Table 28**). Although the labile gene class comprises a mere 5% of the genome, their contribution to underrepresentation can be substantial ranging from 0-73% (**Table 28**).

Expected differences in the proportion of male-biased genes for X-linked and autosomal-linked genes were also studied under several additional scenarios (**Tables 29-31**). First, effects were considered for the scenario in which the full proportion of labile genes is on the X chromosome instead of roughly evenly distributed on the X and the A, as is observed. Overrepresentation of labile genes on the X chromosome would result in reduced 'demasculinization' (**Table 29**), while overrepresentation on the autosomes would enhance 'demasculinization' (**Table 30**). Finally, when the overall proportion of candidate genes is increased 5-fold, enhanced 'demasculinization' is predicted (**Table 31**).

CONCLUSIONS

Expression of sex-limited genes is typically costly in the (unbiased) alternative sex; expression in the female sex may more often be costly

Genus-wide patterns of evolution of sex-biased gene expression throughout the *Drosophila* clade generally provide support for the notion that expression of sex-biased genes in the alternative sex is most often costly. Specifically, gains of sex-biased expression are more likely to involve loss of expression in the unbiased sex and losses of sex-biased gene expression are most likely to involve loss of expression in the biased sex. However, considerable taxonomic variation is apparent, and deviations from these general patterns are characteristic of several species. Given the amount of taxonomic variation observed in patterns sex dependent evolution, it seems likely that global scale changes are responsible for general trends affecting large numbers of genes between species. Sex-dependent gene expression is assayed based on whole adult samples across multiple species and expression divergence indicates only relative change as compared to the whole organism. Nothing can be deduced about whether expression changes are due to modification in *cis- or trans*-acting regulators, or indeed, whether expression change is due to effects on transcription for the genes in question. For example allometric changes in the relative size of reproductive tissues due to growth factors affecting all genes but those expressed

in reproductive tissues could lead to sex comparative divergence that would result in high levels of sex-specific expression divergence for large numbers of sex-biased genes expressed in reproductive tissues. In particular, allometric changes in the relative size of reproductive organs, where sex-biased genes are disproportionately represented (Ranz et al. 2003), are likely to be responsible for trends in the evolution sex-dependent gene expression. Marked differences in the magnitude of the comparative sex index seen for genes that are consistently male biased may be due to relative changes in testis size between species, where the magnitude of sex comparative index is largest (**Figure 3**).

Although sex-biased gene expression appears to be generally costly for the alternative sex, several observations suggest that male expression divergence is less constrained. Within classes of sex-biased gene expression, enhanced male divergence (CSI < 0) is relatively more common for female-biased genes than is enhanced female divergence (CSI > 0) for male-biased genes (**Table 5**). Secondly, when sex-biased gene expression is gained or lost, expression in the unbiased sex is more likely to occur when the unbiased sex is the male (**Tables 3 & 4**).

Patterns of male expression divergence mimic those seen for molecular evolution (Chapter 1). Expression divergence may be generally enhanced in the male sex because of enhanced selection pressure in the male environment, perhaps due to asymmetries in the relative intensity of intrasexual competition between the sexes. Alternatively, enhanced expression divergence may be explained by relative constraints arising from historical sex differences, where relative to the female, there may be less opportunity for negative pleiotropic interaction when a gene is expressed in the male (Chapter 1). Although conserved male-biased genes are more numerous (in this study, they are twice is common), conserved female-biased genes are expressed more broadly, are longer, and are involved in more protein-protein interactions, all of which may relatively enhance the potential for negative pleiotropy (Chapter1).

Spatially broad expression does not inhibit the evolution of female-biased gene expression

Sex-biased gene expression is expected to evolve when expression is beneficial in one sex and neutral or costly in the alternate sex. Considering a new sexually antagonistic allele which is expressed in both sexes, population mean fitness will only be positive if benefits to the favored sex offset costs to the alternative sex when averaged across the population. It has been proposed that the evolution of sex-limited gene expression may be hampered by spatial breadth, since increased expression breadth may magnify the effects of negative pleiotropy and thereby reduce the magnitude of sex-differential fitness effects (Mank et al. 2008). Several observations suggest that spatial breadth may indeed inhibit the evolution of male-biased gene expression, but this view is not consistent with observed patterns of evolution for female-biased genes. Female-biased genes are among the most broadly expressed genes, and genes that are inferred to have gained female-biased gene expression over the evolutionary time span under examination are no less tissue-biased than are nonsex-biased genes. Differences in the spatio-temporal expression patterns of sex-biased genes cannot be attributed to differences in the expression patterns of genes expressed in male versus female reproductive tissues, since these differences also characterize genes that are not expressed specifically or maximally in reproductive tissues (**Fig 6**).

Hypotheses about the relationship between spatial expression and the evolution of sex-biased genes assume antagonism; however conflicts may not always exist. One possible explanation is that female-biased genes may be more broadly expressed because alleles conferring female benefits may often be commensal. Analysis of gain and loss of female-biased

gene expression events does not support this view, rather the evolution of female-biased gene expression more often involves a disproportionate loss of expression in the male sex than vice versa (**Table 3**). Alternatively, ongoing antagonistic selection pressures may lead to selection bias on genes with restricted spatial (and temporal) expression profiles since successive allelic replacements will be advantaged at every stage if exposure to negative pleiotropy in the alternative sex is minimal. Observed sex differences in spatial and temporal gene expression may result if genes involved in ongoing tug-of war dynamics are more likely to appear in the male-biased state. Analysis of candidate genes provides some support for this view, though further research is required.

Candidate gene analysis: the development of male secondary sexual traits may be the primary source of intra-locus sexual conflicts between the sexes

Tug-of-war dynamics are not expected to develop when sexual antagonism can be resolved. There are several ways in which antagonisms may be resolved; for example, essential gene functionality may be compensated for via gene duplication or other mechanisms (Rice and Holland 1997). In cases where gene function is inessential in the disfavored sex, antagonisms may be resolved through the evolution of sex-limited gene expression. Enhanced lability is expected in the particular case of shared function, where evolutionary opportunity does not permit resolution, or where conflicts may occur with especially high frequency. If enhanced lability on the sex comparative index is indeed an effective means of detecting genes mediating ontogenetic sexual conflict, functional analysis suggests that secondary sexual traits may be the primary source of ongoing developmental antagonism in *Drosophila*: Firstly, overrepresentation of candidate genes is indicated for genes expressed in the carcass and at the adult stages of the life cycle, while genes expressed in reproductive tissues are not overrepresented. Functional analysis of candidate genes suggests a role for sexual selection: candidate genes are overrepresented in places (in the carcass) and at times (during the adult stage) when sexual selection pressures are expected to be disproportionate. Although the genetic basis of sexual ornaments in *Drosophila* is rarely well-documented enough to permit functional attribution of specific genes, it seems plausible that these genes may be overrepresented among those expressed specifically in the carcass, which includes several aspects of external morphology that are associated with courtship and other secondary sexual functions (Cobb, Connolly and Burnet 1985, Ewing 1964), including the sex combs, tarsus, and wings. Secondly, candidate genes are significantly overrepresented among genes that are sometimes expressed in the male-biased state but are never expressed in the female-biased state, suggesting that candidate genes spend more time as male-biased.

When there is shared function, costs are expected to derive from negative pleiotropic effects in the alternative sex, which must be offset by benefits to the favored sex. If selection on genes with shared function is more intense in the male environment, male-biased gene expression may often evolve despite negative fitness consequences to the female. This may explain why labile genes are more often male-biased. Alternatively, differences in the potential for negative pleiotropy between the sexes may facilitate the evolution of male benefiting mutations making it more likely that new mutations will be beneficial when expressed in the male sex. Functional analysis suggest a role for sexual selection: it is not clear why genes that may be involved in the development of external morphology, and that are specific to the adult would be overrepresented because of biases derived from sex differences in historical conditions alone, although this hypothesis can be further tested (for example, using methods introduced in

the first chapter). It is notable that enhanced rates of sequence evolution are seen for candidate genes that are sometimes male-biased when expressed specifically at the adult stage, though no enhancement is seen for genes that are sometimes female-biased (**Fig 9**). It is not clear whether enhanced selective pressures arise predominately from intrasexual male competition or intersexual coevolution, though if females experience more often experience the negative consequences of developmental antagonism, mechanisms of female choice may favor the enhanced compartmentalization of genes that will not benefit the female but will benefit the female's male progeny.

Enhanced tug-of-war dynamics arising from selection on secondary secondary sexual traits may in part explain the underrepresentation of male-biased genes on the X chromosome

Phylogenomic analysis of sex-biased gene expression provides support for the idea that expression evolution plays a role in the de-masculinization of sex chromosomes. Furthermore, analysis of candidate genes suggests that transition dynamics among labile genes may contribute substantially to the 'demasculinization' phenomena (ranging between 0% in *D. virilis* to 73% in *D. mojavensis*), despite comprising a relatively small proportion of the overall genome (~5%). In *Drosophila*, candidate genes do not appear to be overrepresented on the X chromosome. Counterintuitively, it appears that increasing the relative proportion of labile genes on the X chromosome would lead to reduced 'demasculinization', while a relative increase in the proportion of labile genes on the autosomes would lead to enhanced 'demasculinization' of the X chromosome.

FIGURES AND TABLES

Figure 6 Evolutionary change in the state of sex-biased gene expression may involve disproportionate change in male or female expression (Connallon and Knowles 2005). Lines of equal bias are indicated with dashed lines, where a line of equal male-bias is shown in blue and a line of equal nonsex-biased gene expression is shown as a dashed black line. Several alternative scenarios for a loss of male-biased gene expression are shown. Scenario 1 involves a disproportionate increase in female expression, corresponding to positive values of the sex-comparative index defined in the methods section (CSI). The second scenario shows a change in sex-biased gene expression involving proportionate change in both sexes, while the third scenario involves a disproportionate gain of expression in the male. Determination of which of these evolutionary pathways has taken place is based on phylogenetic analysis of sex-biased gene expression using male and female expression intensities assayed from multi-species microarray analysis.

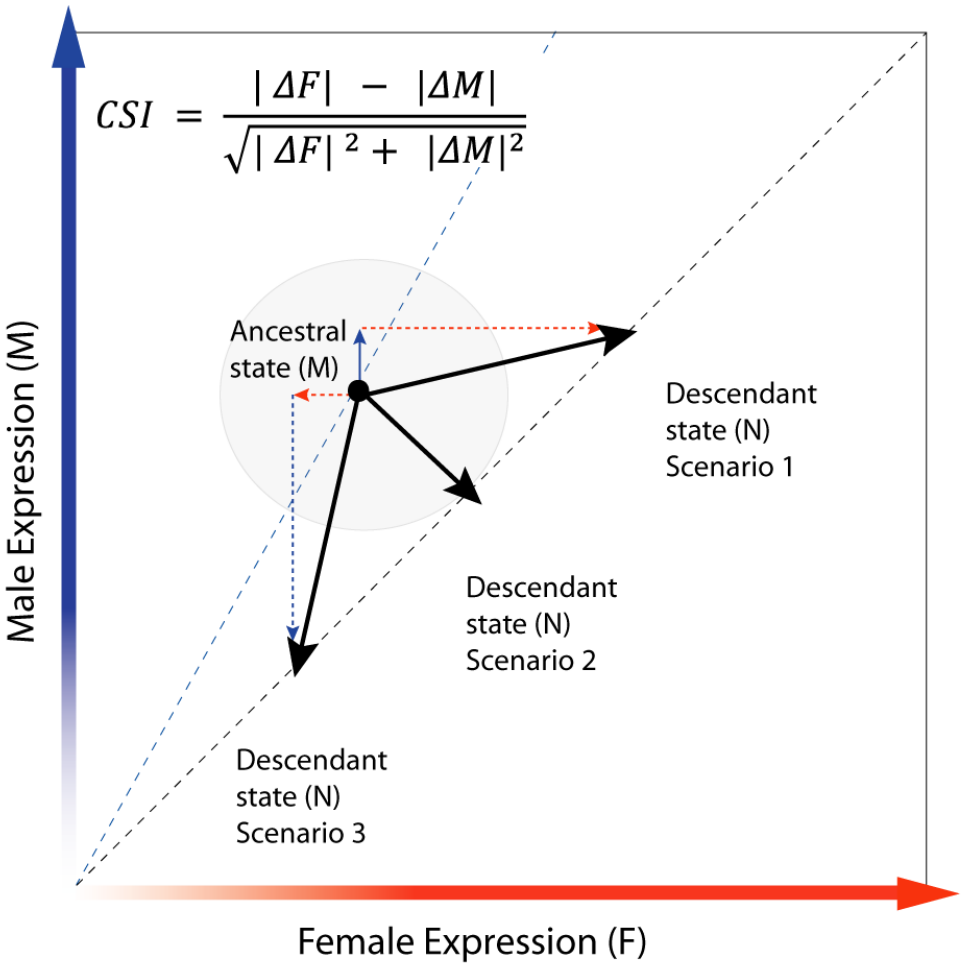
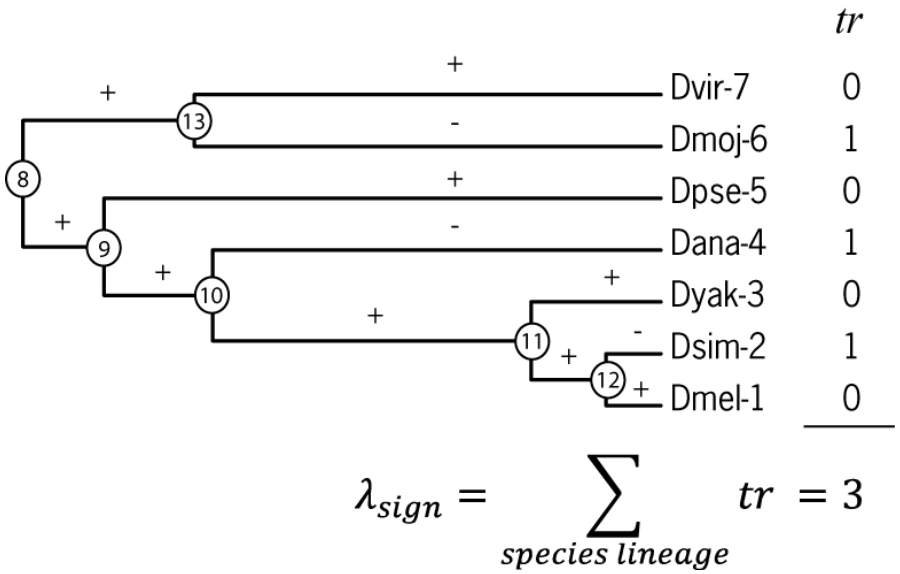


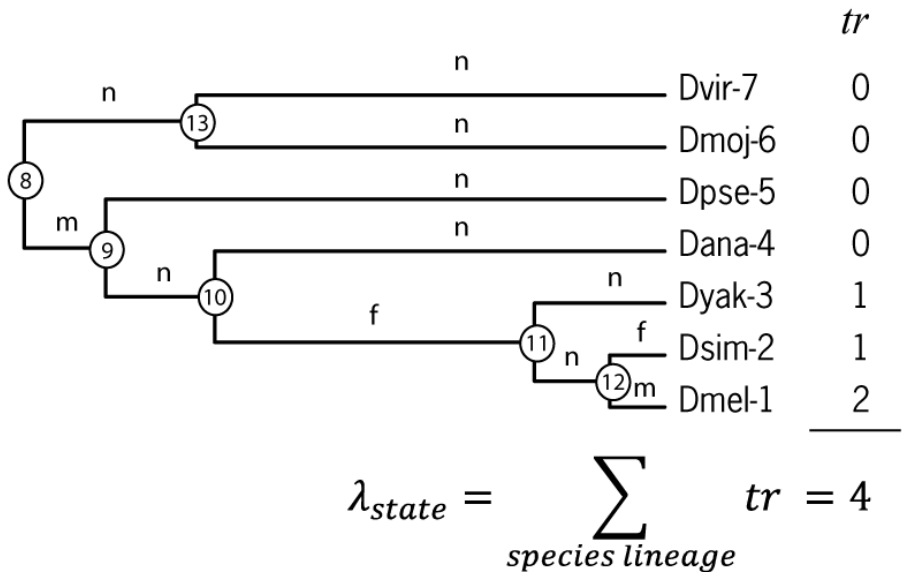
Figure 7 Graphic demonstrating the method for calculating the lability index. Lability can be calculated based on the sign of the index (top) and the state of the sex-comparative index (bottom). Nodes and tips are labeled numerically for reference. For calculation of the lability index based on the sign of the index, the lability index is the sum of the number of transition (*tr*) counted for each species lineage evaluated individually. At each internal node, a transition is counted if adjacent branches have different signs. In calculating the lability index based on the state of the sex-comparative index, branches for which the **n** state is indicated are ignored; that is, the transition number for each lineage can be obtained by collapsing the lineage wherever **n** occurs. Below, all nodes and tips are labeled for reference.

Figure 7a Graphic shows calculation based on the sign of the sex comparative index



$$\lambda_{sign} = \sum_{species\ lineage} tr = 3$$

Figure 2b Graphic shows calculation based on the state of the sex comparative index



$$\lambda_{state} = \sum_{species\ lineage} tr = 4$$

Figure 8 Plots showing the mean comparative sex index for genes that show conserved sex-biased gene expression for each branch of the *Drosophila* phylogeny

Figure 8a Phylogeny relating the seven species of Drosophila. Nodes and tips are labeled numerically for reference.

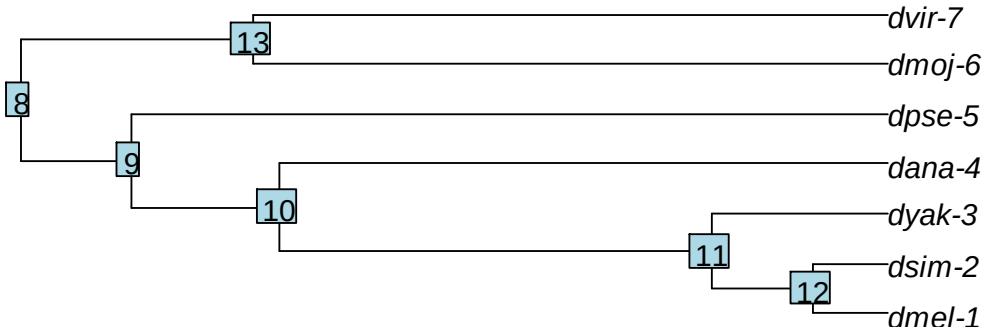


Figure 8b Plots showing mean comparative sex index. Plot of the mean comparative-sex index for groups of male-biased (blue), female-biased (red), and nonsex-biased (gray) genes is shown. Confidence intervals are based on Bonferroni adjusted p-values (n = 3). Analysis is based on a sample size of 72 conserved male-biased, 2,052 conserved nonsex-biased, and 34 conserved female-biased orthologous families.

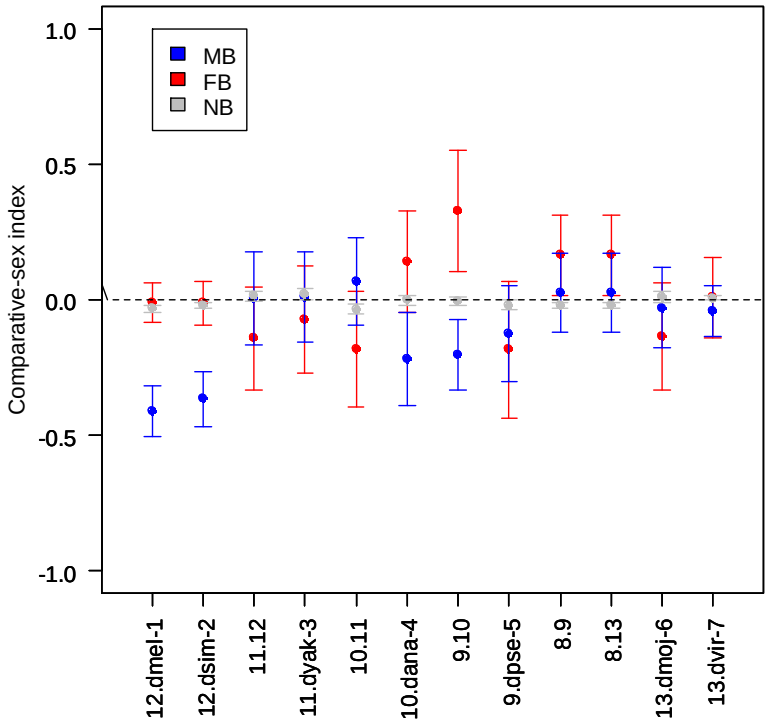


Table 13 Relationship between sex-biased gene expression and the sex comparative index for each type of transition in sex-biased gene expression. Tables below show the total number of genes with a positive or negative index for each of the possible transition types (including gain of sex-biased gene expression events (NM and NF) loss of sex-biased gene expression events (MN and FN), and maintenance of sex-biased gene expression (FF, MM, and NN).

Table 13a Gain of sex-biased gene expression

NF	***12.dsim-2***	***12.Dmel-1***	***11.12***	***11.Dyak-3***	***10.11***	***Dana-4***	***9.10***	***9.dpse-5***	***8.9***	***8.13***	***13.dmoj-6***	***13.dvir-7***	***Genus-wide***	
-1	7	1	21	7	60	17	13	41	29	19	21	1	∑	218
+1	1	13	6	4	22	24	4	7	10	19	6	8	∑	124
P_{adj}	1.00	0.15	0.50	1.0	0.00	1.0	0.59	<0.001	0.28	1.0	0.50	1.0	P_{TOT}	<0.001
Sig.	*ns*	*ns*	*ns*	*ns*	**	*ns*	*ns*	***	*ns*	*ns*	*ns*	*ns*	*Sig.*	***

NM	***12.dsim-2***	***12.Dmel-1***	***11.12***	***11.Dyak-3***	***10.11***	***Dana-4***	***9.10***	***9.dpse-5***	***8.9***	***8.13***	***13.dmoj-6***	***13.dvir-7***	***Genus-wide***	
-1	0	26	6	5	32	2	3	25	4	7	4	8	∑	122
+1	7	5	14	1	72	3	10	27	21	1	6	1	∑	168
P_{adj}	1.0	0.02	1.0	1.0	0.009	1.0	1.0	1.0	0.081	1.0	1.0	1.0	P_{TOT}	0.012
Sig.	*ns*	*	*ns*	*ns*	**	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*Sig.*	**

Table 13b Loss of sex-biased gene expression

FN	***12.dsim-2***	***12.Dmel-1***	***11.12***	***11.Dyak-3***	***10.11***	***Dana-4***	***9.10***	***9.dpse-5***	***8.9***	***8.13***	***13.dmoj-6***	***13.dvir-7***	***Genus-wide***	
-1	56	81	44	118	79	114	16	106	0	69	95	208	∑	1008
+1	129	92	44	114	66	252	54	123	22	28	91	70	∑	1085
P_{adj}	<0.001	1.0	1.0	1.0	1.0	<0.001	<0.001	0.77	1.00	0.003	1.00	<0.001	P_{TOT}	0.0021
Sig.	***	*ns*	*ns*	*ns*	*ns*	**	**	*ns*	*ns*	**	*ns*	***	*Sig.*	**

MN	***12.dsim-2***	***12.Dmel-1***	***11.12***	***11.Dyak-3***	***10.11***	***Dana-4***	***9.10***	***9.dpse-5***	***8.9***	***8.13***	***13.dmoj-6***	***13.dvir-7***	***Genus-wide***	
-1	192	55	30	93	37	153	35	63	8	11	85	32	∑	794
+1	32	24	31	116	24	92	10	43	0	57	48	133	∑	610
P_{adj}	<0.001	0.053	1.00	1.00	1.00	0.009	0.021	0.77	1.00	<0.001	0.14	<0.001	P_{TOT}	0.001
Sig.	***	*ns*	*ns*	*ns*	*ns*	**	*	*ns*	*ns*	***	*ns*	***	*Sig.*	**

Table 13 (cont.)

Table 13c Maintenance of sex-biased gene expression

FF	*12.dsim-2*	*12.Dmel-1*	*11.12*	*11.Dyak-3*	*10.11*	*Dana-4*	*9.10*	*9.dpse-5*	*8.9*	*8.13*	*13.dmoj-6*	*13.dvir-7*	*Genus-wide*	
-1	98	37	186	84	186	51	139	217	195	126	169	44	∑	1532
+1	61	134	131	89	137	52	313	76	288	282	72	105	∑	1740
P_{adj}	0.35	<0.001	0.20	1.00	0.63	1.00	<0.001	<0.001	<0.001	0.0023	<0.001	<0.001	P_{TOT}	<0.001
Sig.	*	***	*ns*	*ns*	*ns*	*ns*	***	***	***	***	***	***	*Sig.*	***

MM	***12.dsim-2***	***12.Dmel-1***	***11.12***	***11.Dyak-3***	***10.11***	***Dana-4***	***9.10***	***9.dpse-5***	***8.9***	***8.13***	***13.dmoj-6***	***13.dvir-7***	***Genus-wide***	
-1	147	310	181	135	158	111	279	181	174	171	107	140	∑	2094
+1	52	34	222	120	201	61	128	165	253	196	134	79	∑	1636
P_{adj}	<0.001	<0.001	1.00	1.00	1.00	0.014	<0.001	1.00	0.013	1.00	1.00	<0.001	P_{TOT}	<0.001
Sig.	***	***	*ns*	*ns*	*ns*	*	***	*ns*	*	*ns*	*ns*	***	*Sig.*	***

NN	***12.dsim-2***	***12.Dmel-1***	***11.12***	***11.Dyak-3***	***10.11***	***Dana-4***	***9.10***	***9.dpse-5***	***8.9***	***8.13***	***13.dmoj-6***	***13.dvir-7***	***Genus-wide***	
-1	1536	1508	1187	1184	1359	1300	1345	1267	1385	1411	1244	1212	∑	15938
+1	1187	1185	1402	1435	1071	1270	1156	1164	1116	1127	1422	1473	∑	15008
P_{adj}	<0.001	<0.001	0.002	<0.001	<0.001	1.00	0.014	1.00	<0.001	<0.001	1.00	<0.001	P_{TOT}	1.00
Sig.	***	***	**	***	***	*ns*	*	*ns*	***	***	*ns*	***	*Sig.*	*ns*

Table 14 Results of chi-squared test of independence of the sex comparative index among branches. P-values adjusted based on the Bonferroni correction (N=7).

Transition type	df	Chi-squared	P_{adj}
NF	11	345.34	<0.001
NM	11	339.86	<0.001
FN	11	248.75	<0.001
MN	11	65.69	<0.001
FF	11	57.29	<0.001
MM	11	188.00	<0.001
NN	11	249.80	<0.001

Table 15 Expression in the unbiased sex upon a gain of sex-biased gene expression A gain of female-biased gene expression event is more likely to involve loss of expression in the male sex than vice versa (two-tailed Fisher Exact Test: Odds ratio = 2.42, $p_{adj} < 0.001$).

Genus-wide	Loss of expression in unbiased sex	Gain of expression in biased sex
NF	218	124
NM	168	122

Table 16 Expression in the unbiased sex upon a loss of sex-biased gene expression: A loss of female-biased gene expression is more likely to involve a gain of expression in the male sex than vice versa (two-tailed Fisher Exact Test: Odds ratio = 1.21, $p_{adj} = 0.006$).

Genus-wide	Gain of expression in unbiased sex	Loss of expression in biased sex
FN	1008	1085
MN	610	794

Table 17 Expression in the unbiased sex upon no change in sex-biased gene expression: Female-biased gene are more likely to undergo disproportionate divergence in the male sex than vice versa (two-tailed Fisher Exact Test: Odds ratio = 0.89, $p_{adj} = 0.01$).

Genus-wide	Divergence in the unbiased sex	Divergence in the biased sex
FF	1532	1740
MM	1636	2094

Table 18 Tests of association between the sign of the sex comparative index and sex-linkage for branches exhibiting male-biased gene expression: Analysis restricted to species for which chromosomal location is known. Tables show $(\text{Odds ratio})_{X/A}$ comparing odds of positive index for X-linked genes to odds of positive index for autosomal genes. P-values were adjusted for multiple testing across rows and columns based on false discovery rate adjustment (N=15). No significant X-effects on the sign of the sex comparative index were observed.

Trans. Type	Dsim	Dmel	Dyak	Dana	Dpse
MM	1.78 (1.00)	1.46 (1.00)	0.73 (1.00)	0.90 (1.00)	1.18 (1.00)
NM	-	0.00 (1.00)	0.00 (1.00)	0.00 (1.00)	4.07 (1.00)
MN	0.84 (1.00)	0.54 (1.00)	1.08 (1.00)	3.00 (0.14)	2.61 (1.00)

Figure 9 The relationship between the maximal tissue of expression and the comparative-sex index. Plot of means and 95% confidence intervals of the comparative sex index for genes grouped by tissue of maximal expression (acc and tag indicate maximal expression in the accessory gland and thoracic abdomino ganglion respectively). Analysis based on inferred values on the sex comparative index for the branch connecting *D. melanogaster* to the common ancestor of *D. melanogaster* and *D. simulans*.

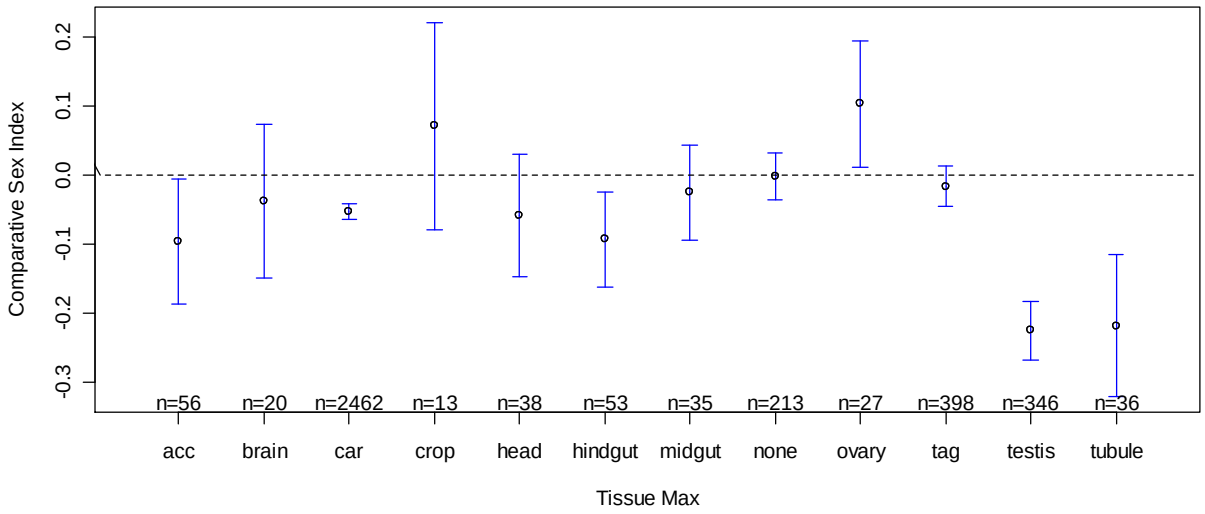


Figure 10 Spatio-temporal gene expression for conserved and non-conserved genes. Boxplots showing distributions of tissue-bias for genes with conserved female-biased gene expression (red, N_F = 41), conserved male-biased gene expression (blue, N_M = 78), conserved nonsex-biased gene expression (gray, N_N = 2400), and for genes that are inferred to have experienced at least one transition in sex-biased gene expression throughout the phylogeny (white, N_T = 1577).

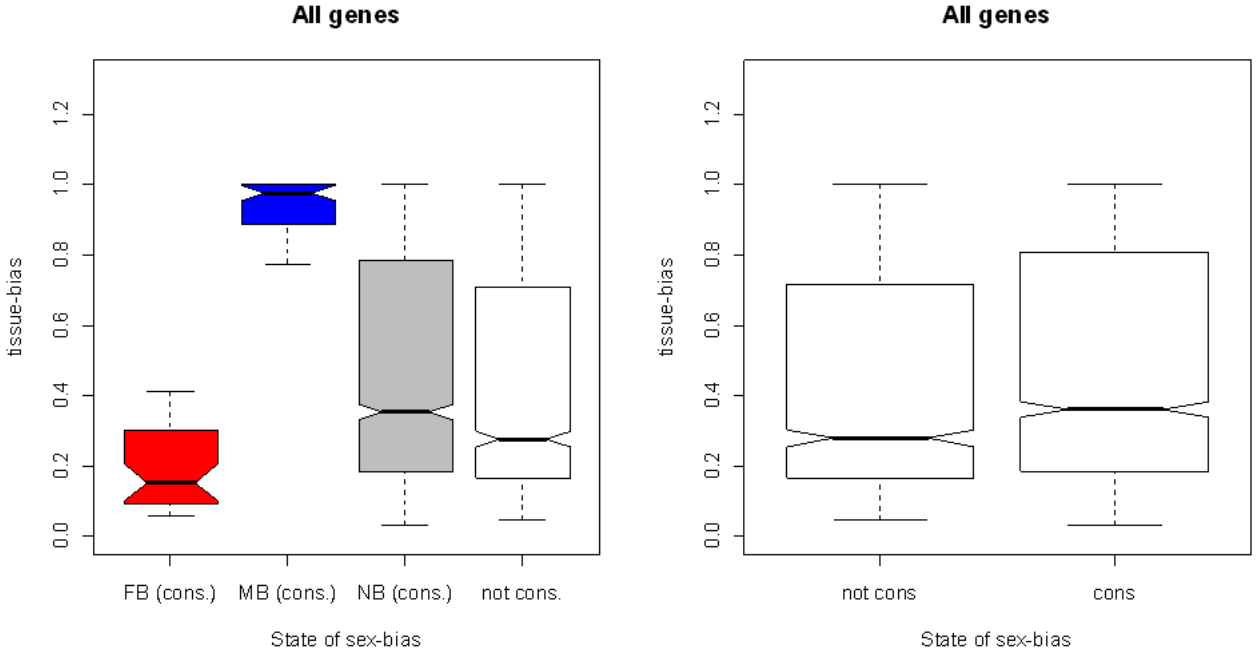


Fig 11 Gain of sex-biased gene expression and spatio-temporal gene expression. Boxplots showing the distribution of tissue-bias (left) and stage-bias (right) for genes grouped by transition type based on inferences derived for the branch connecting *D. melanogaster* to the common ancestor of *D. melanogaster* and *D.simulans* (left).

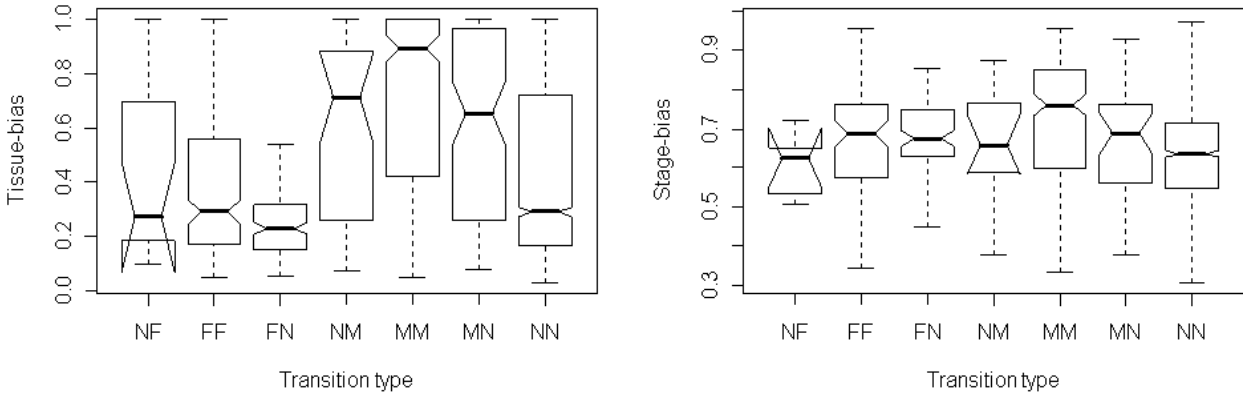


Figure 12 Boxplots showing spatio-temporal expression patterns for reproductive and non-reproductive gene classes. Boxplots on left shows distribution for groupings based on tissue of maximal expression, where repro is indicated if expression is maximal in the accessory glands, testis, ovary, or tubule. For analysis based on tissue-specific expression, repro class includes genes expressed specifically in the accessory gland, testis, ovary, tubule, or spermatheca.

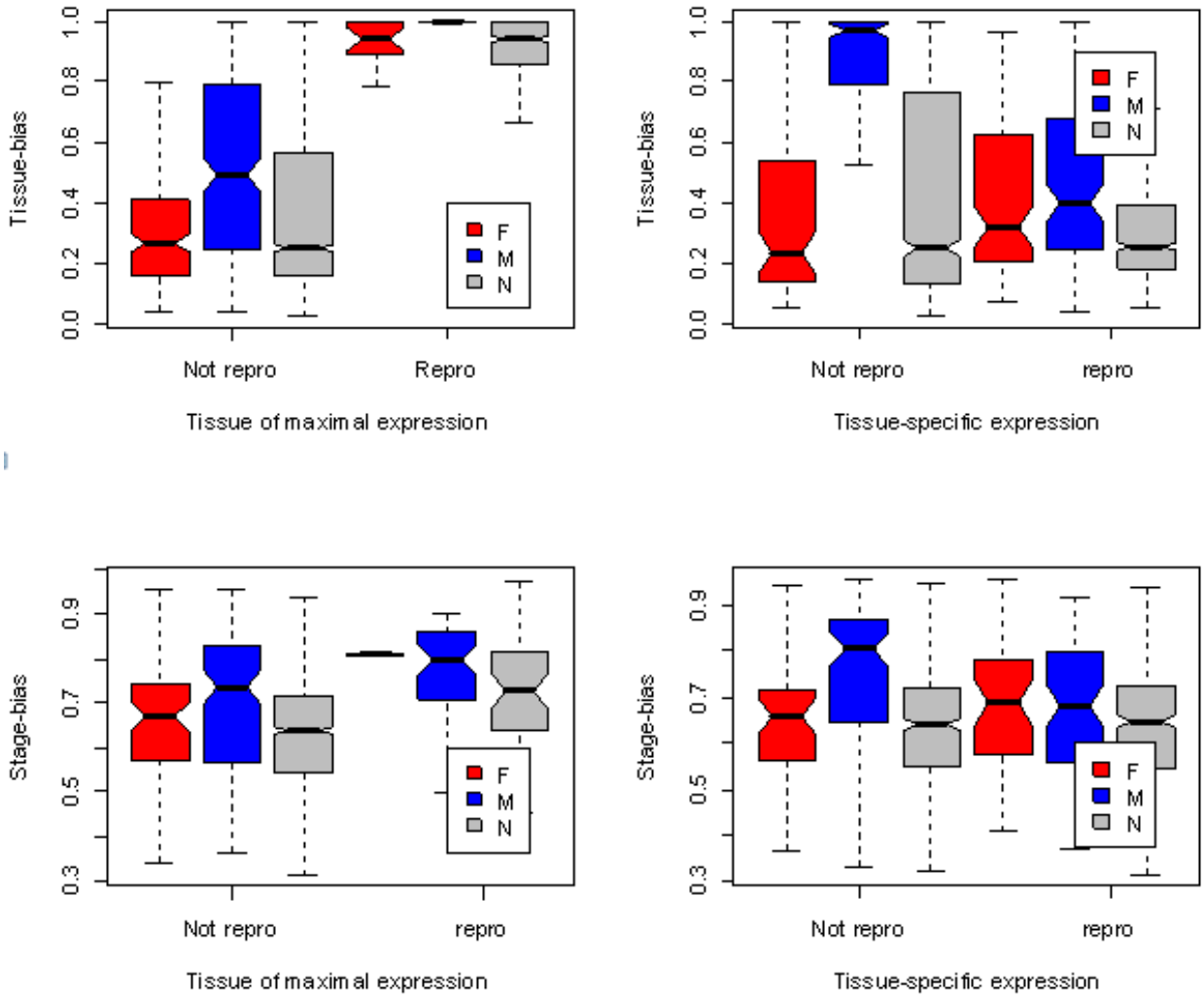


Figure 13 Boxplots showing spatio-temporal expression for labile and non-labile genes (1 = labile/candidate genes, 0 = not-labile/not candidate gene).

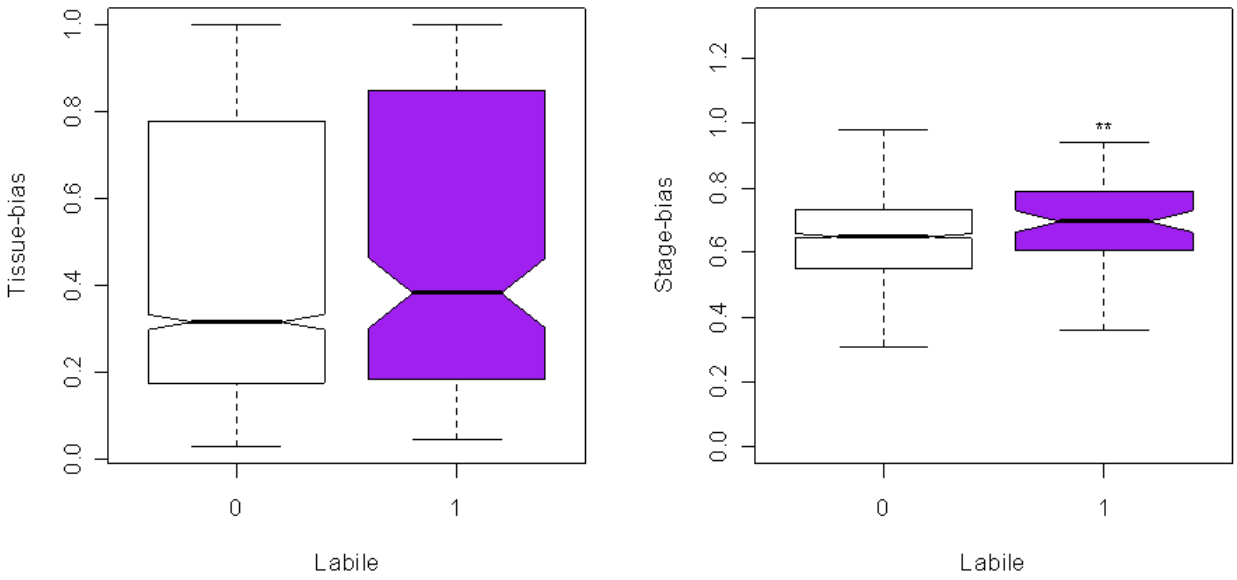


Figure 14 Histograms of the lability index for X-linked and autosomal genes. X-linked. The dashed line indicates the expected value of the lability index under the null model ($\lambda_{exp} = 0.0005$). Significance testing based on one-tailed Mann-Whitney test: (H: $\lambda_X > \lambda_A$).

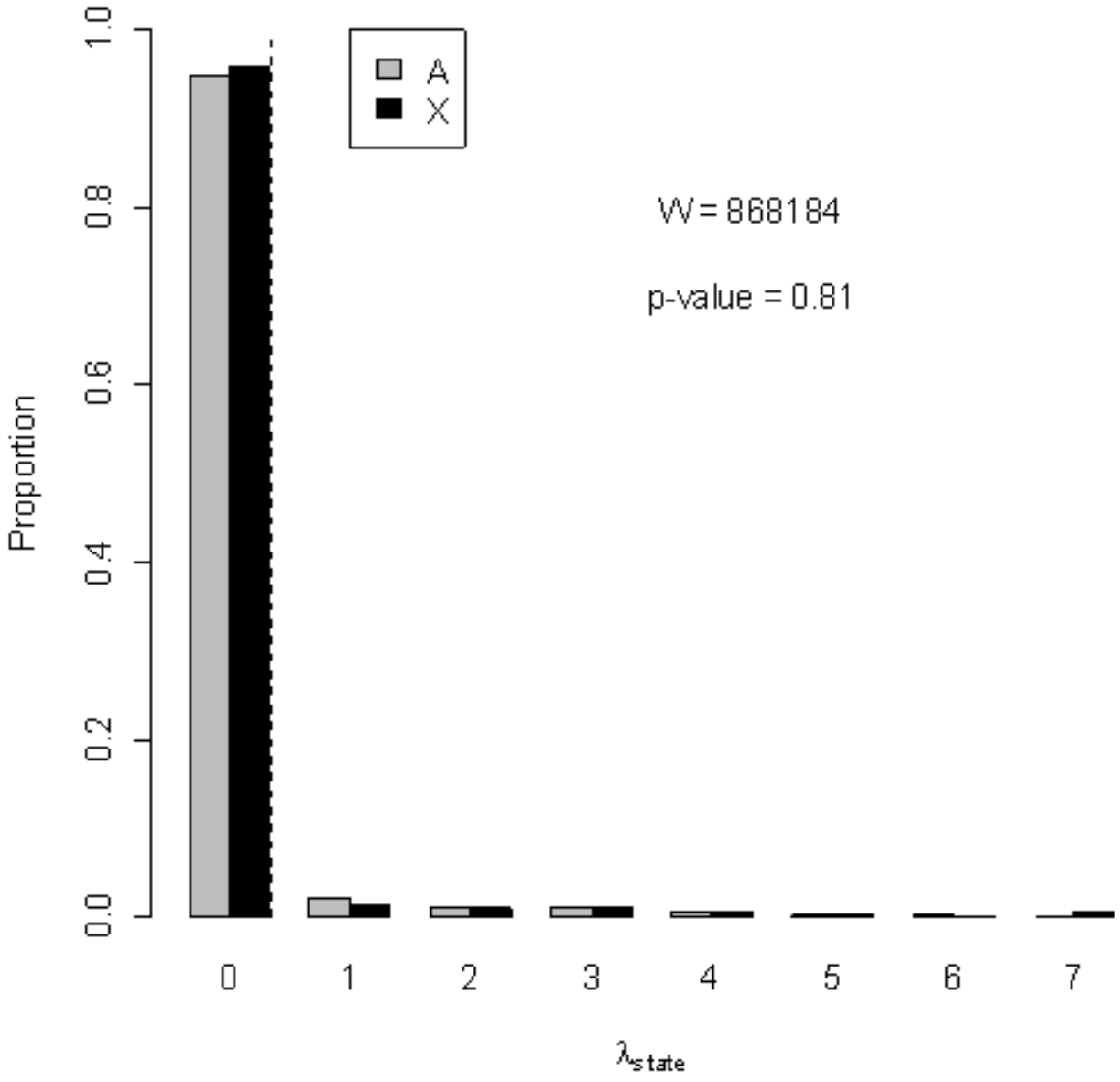


Table 19 Counts of the number of X-linked and autosomal genes for each value of the lability index. Genes that are labile ($\lambda_{state} > 0.0005$) are shaded.

λ_{state}	X	A
0	567	2771
1	8	58
2	5	29
3	5	32
4	3	13
5	1	8
6	0	2
7	2	1
Tot. labile	24	143
Tot. not labile	567	2771

Table 20 Fisher exact tests of over and under-representation of candidate genes. From left to right are names of each indicator variable, sample size, odds ratio based on two-tailed Fisher Exact Test where the ratio has the log odds of presence in the numerator, and of absence in the denominator. P-values are false discovery rate adjusted for multiple testing (N = 30).

Indicator Variable	N	Odds Ratio	P_{adj}	Sig.	Overrep.	Underrep.
Brain	3048	0.90	0.003	**		√
Head	3048	0.48	0.91	*ns*		
Crop	3048	1.25	0.11	*ns*		
Midgut	3048	0.68	0.59	*ns*		
Hindgut	3048	1.05	0.55	*ns*		
Tubule	3048	1.05	0.95	*ns*		
Ovary	3048	0.84	0.75	*ns*		
Testis	3048	1.31	0.61	*ns*		
Accessory Glands	3048	0.99	1.00	*ns*		
Larval Tubule	3048	1.06	0.95	*ns*		
Larval Fat Body	3048	1.32	0.53	*ns*		
TAG	3048	0.24	0.003	**		√
Carcass	3048	1.90	0.04	*	√	
Larval Salivary Gland	3048	1.02	0.95	*ns*		
Larval Midgut	3048	1.11	0.90	*ns*		
Larval Hindgut	3048	0.47	0.12	*ns*		
Virgin Spermatheca	3048	1.43	0.32	*ns*		
Mated Spermatheca	3048	1.43	0.32	*ns*		
Larval CNS	3048	0.40	0.008	**		√
Adult Fat Body	3048	1.64	0.12	*ns*		
Larval Carcass	3048	0.89	0.97	*ns*		
Eye	3048	0.59	0.18	*ns*		
Heart	3048	1.66	0.12	*ns*		
Trachea	3048	0.66	0.53	*ns*		
S2 Cells	3048	0.60	0.20	*ns*		
Embryonic	1633	0.88	0.004	**		√
Larval	1633	0.94	1.00	*ns*		
Metamorphic	1633	1.17	0.328	*ns*		
Adult	1633	2.72	0.001	***	√	
Germline	354	0.72	0.60	*ns*		

Figure 15 Relationship between spatial expression breadth and evolutionary rate for labile and non-labile genes. Spatial breadth classes based on subclasses of tissue-biased gene expression (N_L= 192; N_{NL} = 3904).

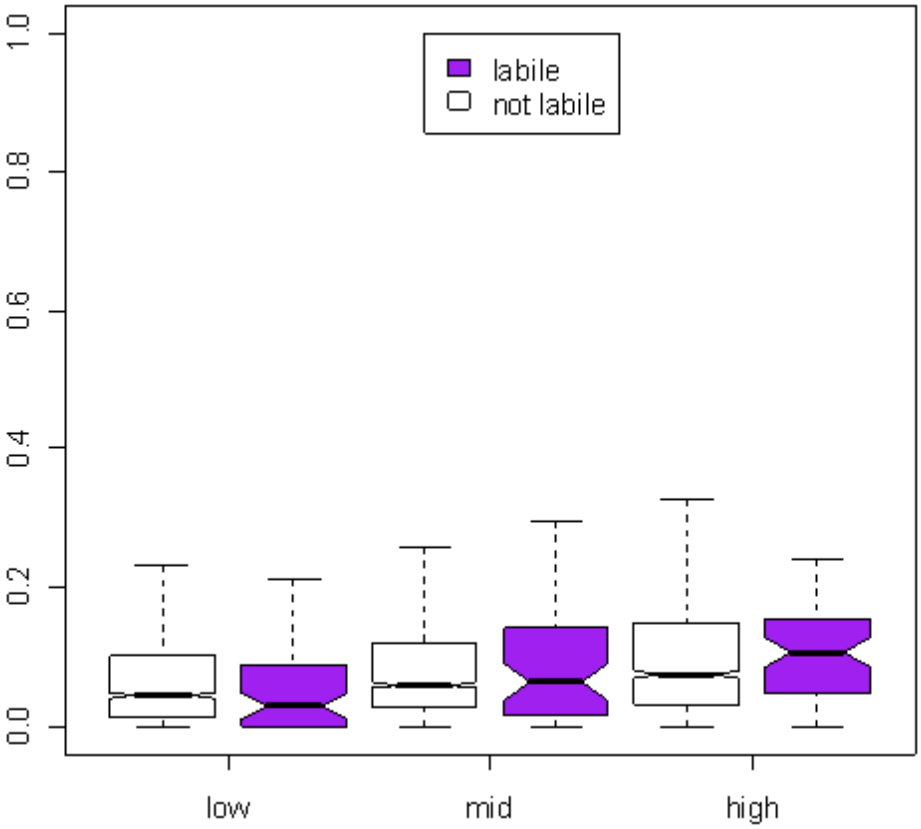


Fig 16 Relationship between lability and tissue-bias for nonsex-biased genes. Boxplots comparing tissue-bias for labile and non-labile genes (gray) for gene groupings based on the neutral cut-off derived for the state of the comparative sex index.

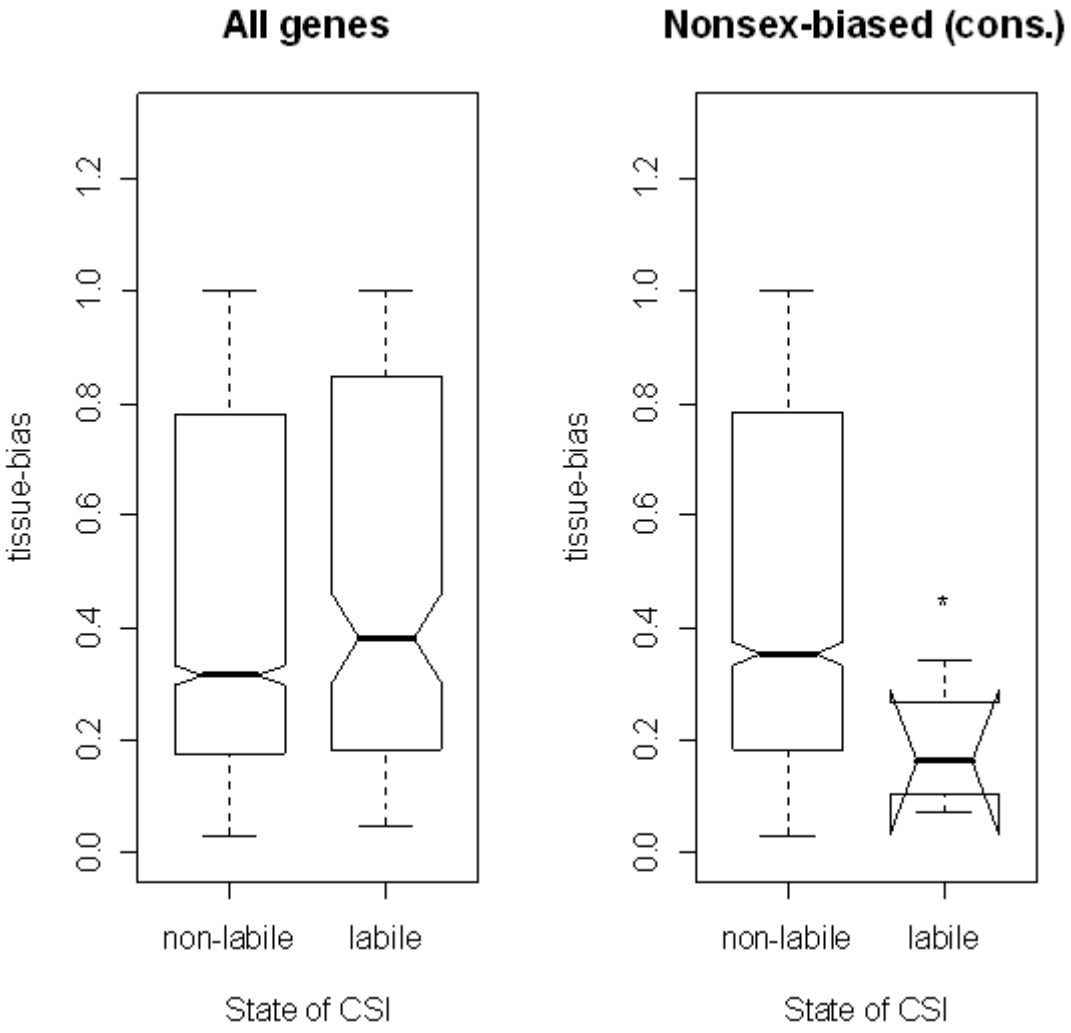


Fig 17 Histogram comparing lability index for genes that show conserved male-biased and female biased gene expression throughout the genus. Significance testing based on one-tailed Mann-Whitney test: (H_0: $\lambda_F = \lambda_M$).

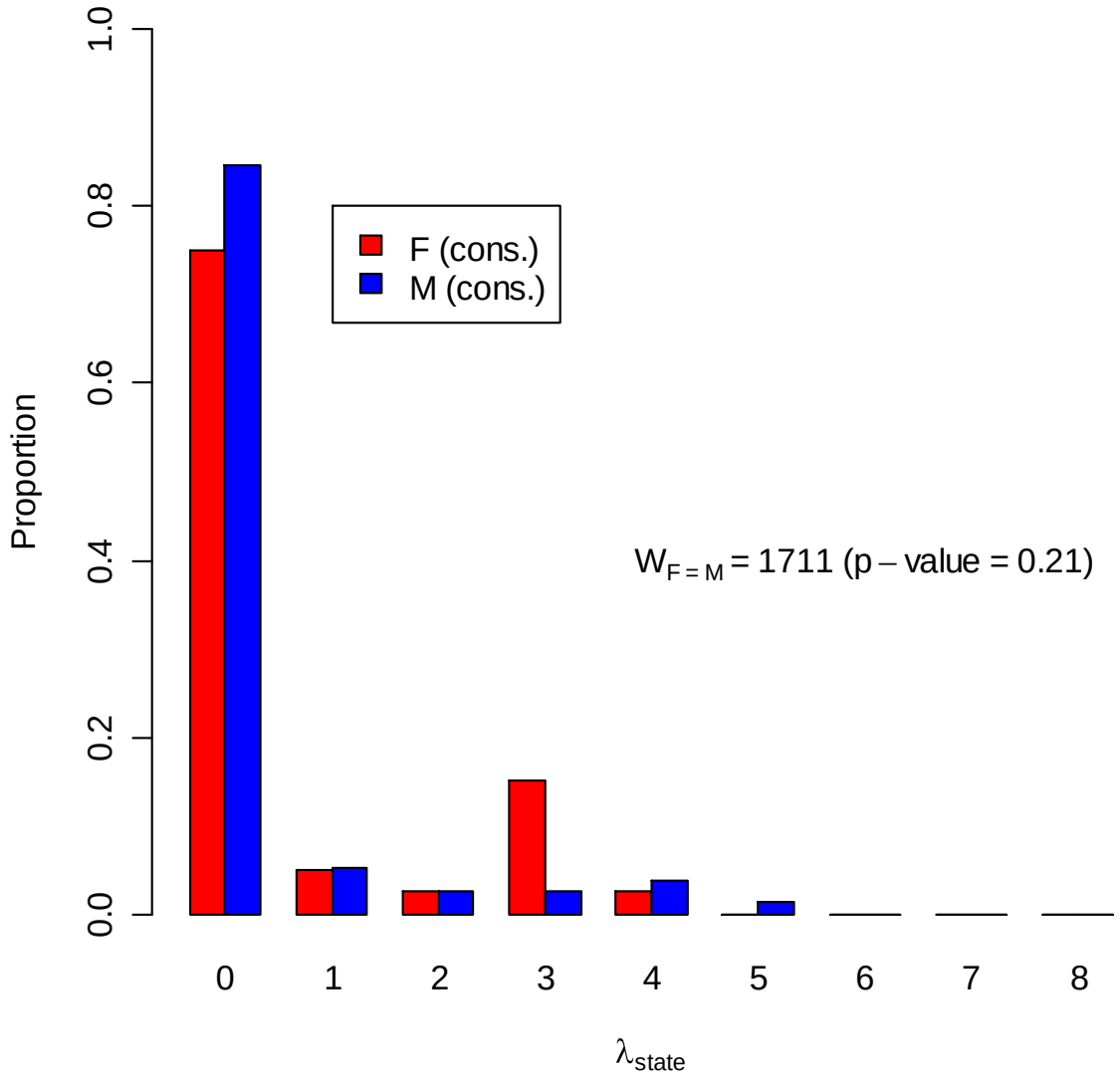


Table 21 Counts of the number of conserved female-biased and conserved male-biased genes for each value of the lability index. Genes that are labile ($\lambda_{state} > 0.0005$) are shaded.

λ_{state}	F	M
0	30	66
1	2	4
2	1	2
3	6	2
4	1	3
5	0	1
Tot. labile	10	12
Tot not-labile	30	66

Fig 18 Histogram comparing the lability index for sometimes male-biased and sometimes female-biased genes throughout the genus. Significance testing based on one-tailed Mann-Whitney test (H_0: $\lambda_M = \lambda_F$).

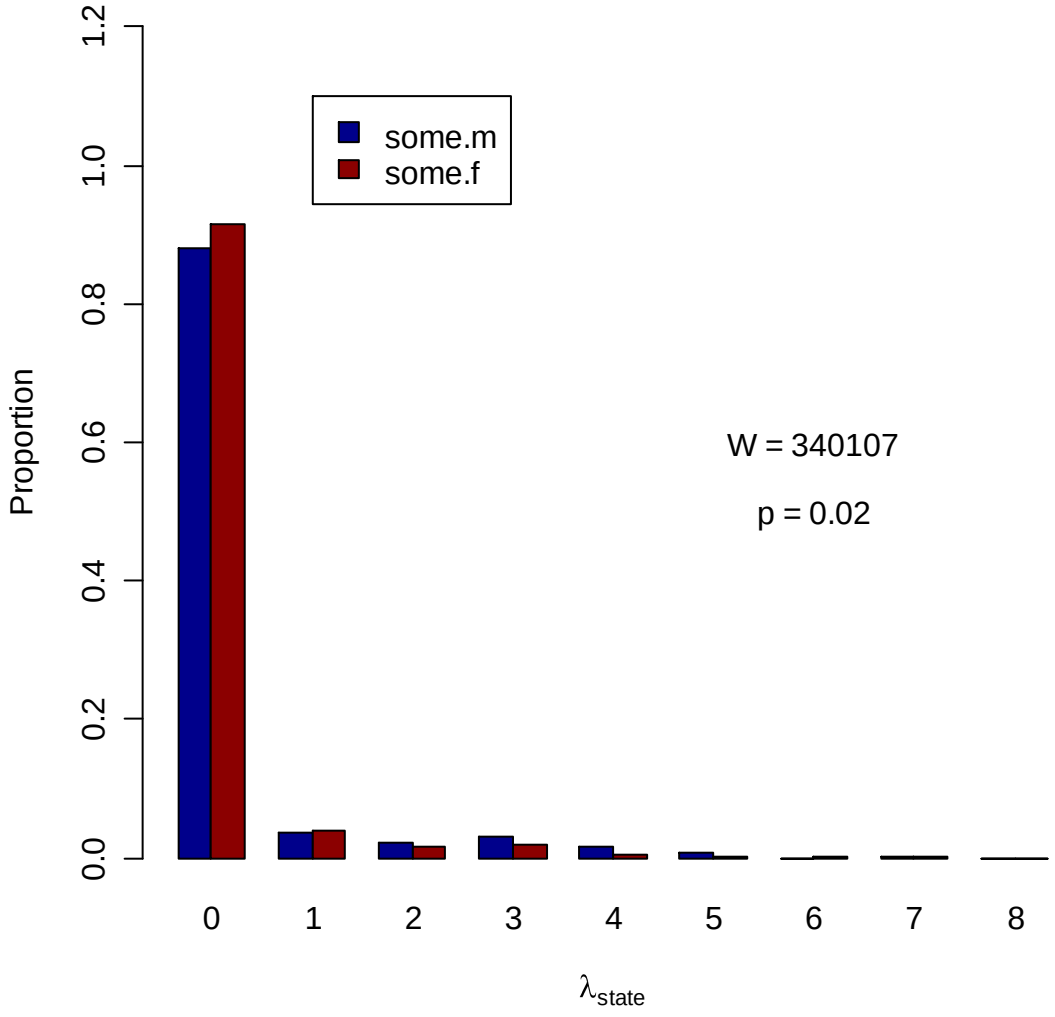


Table 22 Counts of the number of sometimes female and sometimes male genes for each class for each value of the lability index. Genes that are labile ($\lambda_{state} > 0.0005$) are shaded.

λ_{state}	Sometimes female	Sometimes male
0	810	654
1	34	28
2	15	16
3	18	23
4	4	13
5	3	13
6	1	6
7	1	2
8	0	0
Tot. labile	76	88
Tot not-labile	810	654

Figure 19 Boxplots showing the relationship between evolutionary rate, lability and adult-stage expression. Genes that are sometimes male but never female are shown to the right, while genes that are sometimes female and never male are shown to the left.

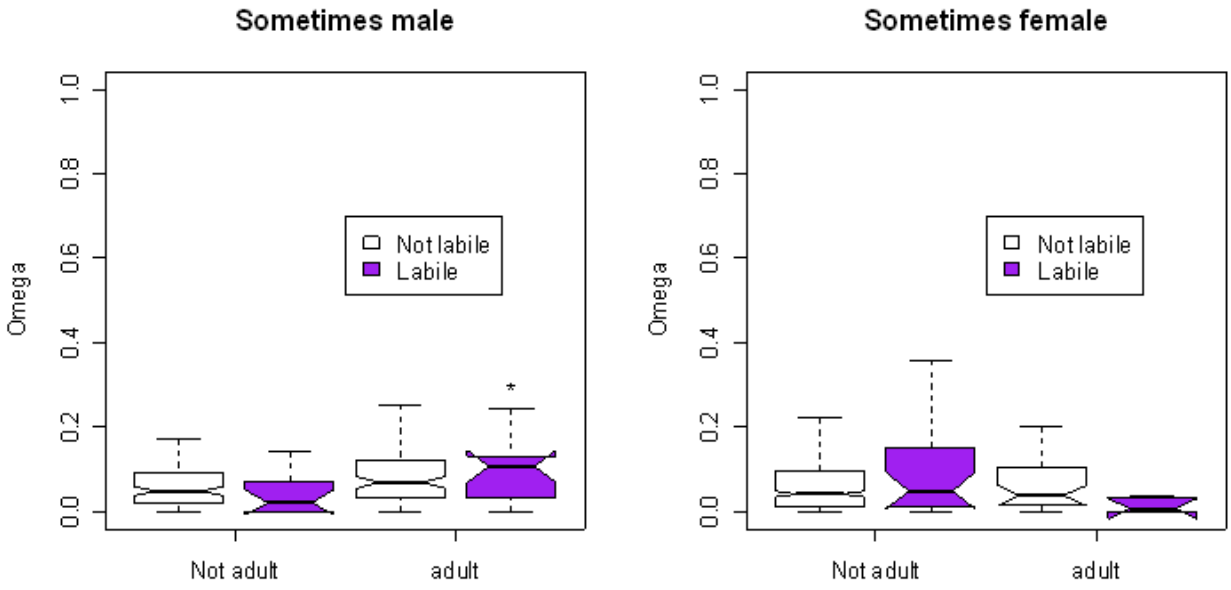


Table 23 Expression evolution as a dynamic process. Probability matrices for X-linked and autosomal genes (left) were derived from transition counts based on states of sex-biased gene expression inferred at the root of the tree and observed at each tip. Stationary distributions (middle) correspond to the left eigenvector of each probability matrix. Expected difference based on difference in estimated proportions at the stationary distribution, while actual differences in proportions are based on the observed proportions of sex-biased genes in each species.

Species	Probability Matrix for X-linked genes	Probability Matrix for Autosomal genes	Stationary Distribution for X-linked genes	Stationary Distribution for Autosomal genes	Expected Difference in Proportions	Observed difference in Proportions
	$\begin{bmatrix} P_{FF} & P_{FM} & P_{FN} \\ P_{MF} & P_{MM} & P_{MN} \\ P_{NF} & P_{NM} & P_{NN} \end{bmatrix}_X$	$\begin{bmatrix} P_{FF} & P_{FM} & P_{FN} \\ P_{MF} & P_{MM} & P_{MN} \\ P_{NF} & P_{NM} & P_{NN} \end{bmatrix}_A$	$\begin{bmatrix} \pi_X(F) \\ \pi_X(M) \\ \pi_X(N) \end{bmatrix}$	$\begin{bmatrix} \pi_A(F) \\ \pi_A(M) \\ \pi_A(N) \end{bmatrix}$	$\begin{bmatrix} \pi_X(F) - \pi_A(F) \\ \pi_X(M) - \pi_A(M) \\ \pi_X(N) - \pi_A(N) \end{bmatrix}$	$\begin{bmatrix} P_X(F) - P_A(F) \\ P_X(M) - P_A(M) \\ P_X(N) - P_A(N) \end{bmatrix}$
Dvir	$\begin{bmatrix} 0.18 & 0.00 & 0.82 \\ 0.00 & 0.27 & 0.73 \\ 0.00 & 0.00 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.31 & 0.00 & 0.69 \\ 0.00 & 0.49 & 0.51 \\ 0.01 & 0.01 & 0.99 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ 0.00 \\ 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.01 \\ 0.01 \\ 0.98 \end{bmatrix}$	$\begin{bmatrix} -0.01 \\ -0.01 \\ 0.02 \end{bmatrix}$	$\begin{bmatrix} 0.01 \\ 0.00 \\ -0.01 \end{bmatrix}$
Dmoj	$\begin{bmatrix} 0.37 & 0.00 & 0.63 \\ 0.00 & 0.41 & 0.59 \\ 0.01 & 0.01 & 0.98 \end{bmatrix}$	$\begin{bmatrix} 0.48 & 0.00 & 0.52 \\ 0.00 & 0.56 & 0.44 \\ 0.01 & 0.01 & 0.98 \end{bmatrix}$	$\begin{bmatrix} 0.02 \\ 0.01 \\ 0.97 \end{bmatrix}$	$\begin{bmatrix} 0.03 \\ 0.02 \\ 0.96 \end{bmatrix}$	$\begin{bmatrix} -0.01 \\ -0.01 \\ 0.01 \end{bmatrix}$	$\begin{bmatrix} -0.01 \\ -0.03 \\ 0.03 \end{bmatrix}$
Dpse	$\begin{bmatrix} 0.56 & 0.00 & 0.44 \\ 0.00 & 0.78 & 0.22 \\ 0.03 & 0.01 & 0.96 \end{bmatrix}$	$\begin{bmatrix} 0.58 & 0.00 & 0.42 \\ 0.00 & 0.77 & 0.23 \\ 0.03 & 0.03 & 0.94 \end{bmatrix}$	$\begin{bmatrix} 0.07 \\ 0.05 \\ 0.89 \end{bmatrix}$	$\begin{bmatrix} 0.06 \\ 0.11 \\ 0.84 \end{bmatrix}$	$\begin{bmatrix} 0.01 \\ -0.06 \\ 0.05 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ -0.02 \\ 0.02 \end{bmatrix}$
Dana	$\begin{bmatrix} 0.11 & 0.00 & 0.89 \\ 0.02 & 0.24 & 0.73 \\ 0.01 & 0.01 & 0.98 \end{bmatrix}$	$\begin{bmatrix} 0.19 & 0.00 & 0.81 \\ 0.02 & 0.38 & 0.61 \\ 0.02 & 0.00 & 0.98 \end{bmatrix}$	$\begin{bmatrix} 0.02 \\ 0.01 \\ 0.97 \end{bmatrix}$	$\begin{bmatrix} 0.02 \\ 0.00 \\ 0.97 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ 0.01 \\ 0.00 \end{bmatrix}$	$\begin{bmatrix} 0.01 \\ -0.03 \\ 0.02 \end{bmatrix}$
Dyak	$\begin{bmatrix} 0.27 & 0.00 & 0.73 \\ 0.00 & 0.47 & 0.53 \\ 0.01 & 0.01 & 0.98 \end{bmatrix}$	$\begin{bmatrix} 0.34 & 0.00 & 0.66 \\ 0.00 & 0.52 & 0.47 \\ 0.01 & 0.02 & 0.97 \end{bmatrix}$	$\begin{bmatrix} 0.02 \\ 0.02 \\ 0.96 \end{bmatrix}$	$\begin{bmatrix} 0.01 \\ 0.04 \\ 0.95 \end{bmatrix}$	$\begin{bmatrix} 0.01 \\ -0.02 \\ 0.01 \end{bmatrix}$	$\begin{bmatrix} -0.01 \\ 0.00 \\ 0.02 \end{bmatrix}$
Dsim	$\begin{bmatrix} 0.27 & 0.00 & 0.73 \\ 0.00 & 0.35 & 0.65 \\ 0.01 & 0.01 & 0.98 \end{bmatrix}$	$\begin{bmatrix} 0.30 & 0.00 & 0.70 \\ 0.00 & 0.42 & 0.58 \\ 0.01 & 0.02 & 0.97 \end{bmatrix}$	$\begin{bmatrix} 0.02 \\ 0.01 \\ 0.97 \end{bmatrix}$	$\begin{bmatrix} 0.01 \\ 0.03 \\ 0.96 \end{bmatrix}$	$\begin{bmatrix} 0.01 \\ -0.02 \\ 0.01 \end{bmatrix}$	$\begin{bmatrix} 0.01 \\ -0.01 \\ 0.00 \end{bmatrix}$
Dmel	$\begin{bmatrix} 0.21 & 0.00 & 0.79 \\ 0.00 & 0.55 & 0.45 \\ 0.02 & 0.03 & 0.94 \end{bmatrix}$	$\begin{bmatrix} 0.31 & 0.01 & 0.68 \\ 0.00 & 0.65 & 0.35 \\ 0.02 & 0.05 & 0.93 \end{bmatrix}$	$\begin{bmatrix} 0.03 \\ 0.07 \\ 0.91 \end{bmatrix}$	$\begin{bmatrix} 0.03 \\ 0.12 \\ 0.85 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ -0.05 \\ 0.06 \end{bmatrix}$	$\begin{bmatrix} 0.02 \\ -0.01 \\ -0.01 \end{bmatrix}$

Table 24 Contingency tables for hypothetical counts at the stationary distribution. Calculations based on the stationary distributions ($\pi_X^{\alpha=0.01}$ and $\pi_A^{\alpha=0.01}$) derived for -value cut-offs corresponding to α = 0.01 Hypothetical counts were calculated on the basis of the sampled numbers of X-linked (N = 591) and autosomal genes (N = 2914).

dvir	M	not M
A	34	2880
X	0	591

**Male-biased genes underrepresented at p-value < 0.01.

dmoj	M	not M
A	49	2865
X	7	585

(ns) No underrepresentation of male-biased genes.

dpse	M	not M
A	309	2605
X	37	564

***Male-biased genes underrepresented at p-value < 0.001.

dana	M	not M
A	7	2907
X	7	584

** Male-biased genes overrepresented at p-value < 0.01.

dyak	M	not M
A	104	2810
X	10	581

* Male-biased genes underrepresented at p-value < 0.05.

dsim	M	not M
A	77	2838
X	8	583

(ns) No underrepresentation of male-biased genes.

dmel	M	not M
A	363	2552
X	39	552

*** Male-biased genes underrepresented at p-value < 0.05.

Table 25 Comparing expression evolution for candidate genes. Probability matrices for X-linked and autosomal genes (left) were derived from transition counts based on states of sex-biased gene expression inferred at the root of the tree and observed at each tip. Stationary distributions (middle) correspond to the left eigenvector of each probability matrix. Expected difference at equilibrium based on difference in estimated proportions at the stationary distribution, where $P(L_X) = 0.05$ and $P(L_A) = 0.04$ are the observed proportions of labile genes on the X chromosome and autosomes respectively.

Species	**Probability Matrix for X-linked genes**	**Probability Matrix for Autosomal genes**	**Stationary Distribution for X-linked gene**	**Stationary Distribution for Autosomal genes**	**Expected Difference in Proportions**
	$\begin{bmatrix} P_{FF} & P_{FM} & P_{FN} \\ P_{MF} & P_{MM} & P_{MN} \\ P_{NF} & P_{NM} & P_{NN} \end{bmatrix}_X^L$	$\begin{bmatrix} P_{FF} & P_{FM} & P_{FN} \\ P_{MF} & P_{MM} & P_{MN} \\ P_{NF} & P_{NM} & P_{NN} \end{bmatrix}_A^L$	$\begin{bmatrix} \pi_X^L(F) \\ \pi_X^L(M) \\ \pi_X^L(N) \end{bmatrix}$	$\begin{bmatrix} \pi_A^L(F) \\ \pi_A^L(M) \\ \pi_A^L(N) \end{bmatrix}$	$\begin{bmatrix} \pi_X^L(F)\cdot P(L_X) - \pi_A^L(F)\cdot P(L_A) \\ \pi_X^L(M)\cdot P(L_X) - \pi_A^L(M)\cdot P(L_A) \\ \pi_X^L(N)\cdot P(L_X) - \pi_A^L(N)\cdot P(L_A) \end{bmatrix}$
Dvir	$\begin{bmatrix} 0.22 & 0.00 & 0.78 \\ 0.00 & 0.14 & 0.86 \\ 0.00 & 0.00 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.48 & 0.00 & 0.52 \\ 0.00 & 0.44 & 0.56 \\ 0.11 & 0.00 & 0.89 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ 0.00 \\ 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.17 \\ 0.00 \\ 0.83 \end{bmatrix}$	$\begin{bmatrix} -0.01 \\ 0.00 \\ 0.00 \end{bmatrix}$
Dmoj	$\begin{bmatrix} 0.44 & 0.00 & 0.56 \\ 0.00 & 0.14 & 0.86 \\ 0.00 & 0.00 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.59 & 0.02 & 0.39 \\ 0.02 & 0.58 & 0.41 \\ 0.11 & 0.05 & 0.84 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ 0.00 \\ 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.19 \\ 0.10 \\ 0.72 \end{bmatrix}$	$\begin{bmatrix} -0.01 \\ 0.00 \\ 0.01 \end{bmatrix}$
Dpse	$\begin{bmatrix} 0.89 & 0.00 & 0.11 \\ 0.00 & 0.57 & 0.43 \\ 0.12 & 0.12 & 0.75 \end{bmatrix}$	$\begin{bmatrix} 0.59 & 0.02 & 0.39 \\ 0.00 & 0.75 & 0.25 \\ 0.11 & 0.29 & 0.61 \end{bmatrix}$	$\begin{bmatrix} 0.47 \\ 0.12 \\ 0.41 \end{bmatrix}$	$\begin{bmatrix} 0.11 \\ 0.48 \\ 0.41 \end{bmatrix}$	$\begin{bmatrix} 0.01 \\ -0.02 \\ 0.00 \end{bmatrix}$
Dana	$\begin{bmatrix} 0.33 & 0.00 & 0.67 \\ 0.14 & 0.14 & 0.71 \\ 0.38 & 0.00 & 0.62 \end{bmatrix}$	$\begin{bmatrix} 0.37 & 0.00 & 0.63 \\ 0.05 & 0.42 & 0.53 \\ 0.21 & 0.03 & 0.76 \end{bmatrix}$	$\begin{bmatrix} 0.36 \\ 0.00 \\ 0.64 \end{bmatrix}$	$\begin{bmatrix} 0.24 \\ 0.03 \\ 0.72 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ 0.00 \\ -0.01 \end{bmatrix}$
Dyak	$\begin{bmatrix} 0.44 & 0.00 & 0.56 \\ 0.00 & 0.57 & 0.43 \\ 0.00 & 0.00 & 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.57 & 0.02 & 0.41 \\ 0.00 & 0.64 & 0.36 \\ 0.00 & 0.21 & 0.79 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ 0.00 \\ 1.00 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ 0.37 \\ 0.63 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ -0.02 \\ 0.01 \end{bmatrix}$
Dsim	$\begin{bmatrix} 0.56 & 0.00 & 0.44 \\ 0.00 & 0.43 & 0.57 \\ 0.00 & 0.12 & 0.88 \end{bmatrix}$	$\begin{bmatrix} 0.41 & 0.00 & 0.59 \\ 0.00 & 0.63 & 0.37 \\ 0.01 & 0.13 & 0.87 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ 0.18 \\ 0.82 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ 0.26 \\ 0.74 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ -0.01 \\ 0.00 \end{bmatrix}$
Dmel	$\begin{bmatrix} 0.56 & 0.00 & 0.44 \\ 0.00 & 0.71 & 0.29 \\ 0.00 & 0.50 & 0.50 \end{bmatrix}$	$\begin{bmatrix} 0.50 & 0.07 & 0.43 \\ 0.00 & 0.76 & 0.24 \\ 0.08 & 0.32 & 0.61 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ 0.64 \\ 0.36 \end{bmatrix}$	$\begin{bmatrix} 0.06 \\ 0.54 \\ 0.39 \end{bmatrix}$	$\begin{bmatrix} 0.00 \\ 0.00 \\ 0.00 \end{bmatrix}$

Table 26 Comparing expression evolution (X vs. A) for non-candidate genes. Probability matrices for X-linked and autosomal genes (left) were derived from transition counts based on states of sex-biased gene expression inferred at the root of the tree and observed at each tip for non-candidate genes. Stationary distributions (middle) correspond to the left eigenvector of each probability matrix. Expected difference at equilibrium based on difference in estimated proportions at the stationary distribution, where $P(L_X) = 0.05$ and $P(L_A) = 0.04$ are the observed proportions of labile genes on the X chromosome and autosomes respectively.

Species	Probability Matrix for X-linked genes	Probability Matrix for Autosomal genes	Stationary Distribution for X-linked gene	Stationary Distribution for Autosomal genes	Expected Difference in Proportions at Equilibirum
Dvir	$\begin{bmatrix}0.18 & 0.00 & 0.82\\ 0.00 & 0.29 & 0.71\\ 0.00 & 0.00 & 1.00\end{bmatrix}$	$\begin{bmatrix}0.29 & 0.00 & 0.71\\ 0.00 & 0.50 & 0.50\\ 0.00 & 0.01 & 0.99\end{bmatrix}$	$\begin{bmatrix}0.00\\ 0.00\\ 1.00\end{bmatrix}$	$\begin{bmatrix}0.01\\ 0.01\\ 0.98\end{bmatrix}$	$\begin{bmatrix}-0.01\\ -0.01\\ 0.03\end{bmatrix}$
Dmoj	$\begin{bmatrix}0.36 & 0.00 & 0.64\\ 0.00 & 0.45 & 0.55\\ 0.01 & 0.01 & 0.98\end{bmatrix}$	$\begin{bmatrix}0.47 & 0.02 & 0.53\\ 0.00 & 0.56 & 0.44\\ 0.01 & 0.01 & 0.98\end{bmatrix}$	$\begin{bmatrix}0.02\\ 0.01\\ 0.97\end{bmatrix}$	$\begin{bmatrix}0.02\\ 0.01\\ 0.96\end{bmatrix}$	$\begin{bmatrix}0.00\\ 0.00\\ 0.02\end{bmatrix}$
Dpse	$\begin{bmatrix}0.53 & 0.00 & 0.47\\ 0.00 & 0.81 & 0.19\\ 0.03 & 0.01 & 0.96\end{bmatrix}$	$\begin{bmatrix}0.58 & 0.00 & 0.42\\ 0.00 & 0.77 & 0.23\\ 0.03 & 0.02 & 0.95\end{bmatrix}$	$\begin{bmatrix}0.06\\ 0.04\\ 0.90\end{bmatrix}$	$\begin{bmatrix}0.06\\ 0.09\\ 0.85\end{bmatrix}$	$\begin{bmatrix}0.00\\ -0.04\\ 0.05\end{bmatrix}$
Dana	$\begin{bmatrix}0.09 & 0.00 & 0.91\\ 0.00 & 0.26 & 0.74\\ 0.01 & 0.01 & 0.98\end{bmatrix}$	$\begin{bmatrix}0.17 & 0.00 & 0.83\\ 0.01 & 0.37 & 0.62\\ 0.02 & 0.00 & 0.98\end{bmatrix}$	$\begin{bmatrix}0.01\\ 0.01\\ 0.98\end{bmatrix}$	$\begin{bmatrix}0.02\\ 0.00\\ 0.98\end{bmatrix}$	$\begin{bmatrix}-0.01\\ 0.01\\ 0.01\end{bmatrix}$
Dyak	$\begin{bmatrix}0.25 & 0.00 & 0.75\\ 0.00 & 0.45 & 0.55\\ 0.01 & 0.01 & 0.98\end{bmatrix}$	$\begin{bmatrix}0.31 & 0.02 & 0.69\\ 0.00 & 0.51 & 0.49\\ 0.01 & 0.01 & 0.98\end{bmatrix}$	$\begin{bmatrix}0.02\\ 0.02\\ 0.96\end{bmatrix}$	$\begin{bmatrix}0.01\\ 0.03\\ 0.96\end{bmatrix}$	$\begin{bmatrix}0.01\\ -0.01\\ 0.01\end{bmatrix}$
Dsim	$\begin{bmatrix}0.25 & 0.00 & 0.75\\ 0.00 & 0.33 & 0.67\\ 0.01 & 0.01 & 0.98\end{bmatrix}$	$\begin{bmatrix}0.29 & 0.00 & 0.71\\ 0.00 & 0.39 & 0.61\\ 0.01 & 0.01 & 0.98\end{bmatrix}$	$\begin{bmatrix}0.02\\ 0.01\\ 0.97\end{bmatrix}$	$\begin{bmatrix}0.01\\ 0.02\\ 0.97\end{bmatrix}$	$\begin{bmatrix}0.01\\ -0.01\\ 0.01\end{bmatrix}$
Dmel	$\begin{bmatrix}0.18 & 0.00 & 0.82\\ 0.00 & 0.52 & 0.48\\ 0.02 & 0.02 & 0.95\end{bmatrix}$	$\begin{bmatrix}0.29 & 0.01 & 0.71\\ 0.00 & 0.63 & 0.37\\ 0.02 & 0.05 & 0.93\end{bmatrix}$	$\begin{bmatrix}0.03\\ 0.05\\ 0.93\end{bmatrix}$	$\begin{bmatrix}0.02\\ 0.11\\ 0.87\end{bmatrix}$	$\begin{bmatrix}0.00\\ -0.06\\ 0.06\end{bmatrix}$

Probability Matrix for X-linked genes:

$$\begin{bmatrix}P_{FF} & P_{FM} & P_{FN}\\ P_{MF} & P_{MM} & P_{MN}\\ P_{NF} & P_{NM} & P_{NN}\end{bmatrix}_X^{\neg L}$$

Probability Matrix for Autosomal genes:

$$\begin{bmatrix}P_{FF} & P_{FM} & P_{FN}\\ P_{MF} & P_{MM} & P_{MN}\\ P_{NF} & P_{NM} & P_{NN}\end{bmatrix}_A^{\neg L}$$

Stationary Distribution for X-linked gene:

$$\begin{bmatrix}\pi_X^{\neg L}(F)\\ \pi_X^{\neg L}(M)\\ \pi_X^{\neg L}(N)\end{bmatrix}$$

Stationary Distribution for Autosomal genes:

$$\begin{bmatrix}\pi_A^{\neg L}(F)\\ \pi_A^{\neg L}(M)\\ \pi_A^{\neg L}(N)\end{bmatrix}$$

Expected Difference in Proportions at Equilibirium:

$$\begin{bmatrix}\pi_X^{\neg L}(F)\cdot[1-P(L_X)]-\pi_A^{\neg L}(F)\cdot[1-P(L_A)]\\ \pi_X^{\neg L}(M)\cdot[1-P(L_X)]-\pi_A^{\neg L}(M)\cdot[1-P(L_A)]\\ \pi_X^{\neg L}(N)\cdot[1-P(L_X)]-\pi_A^{\neg L}(N)\cdot[1-P(L_A)]\end{bmatrix}$$

Table 27 Summary of the calculations used to derive the contribution of candidate genes to 'demasuclinization'

Contribution to total difference from labile genes	$\pi^{L}_{X-A}(M) = \pi^{L}_{X}(M) \cdot P(L_X) - \pi^{L}_{A}(M) \cdot P(L_A)$
Contribution to total difference from non-labile genes	$\pi^{\neg L}_{X-A}(M) = \pi^{\neg L}_{X}(M) \cdot [1 - P(L_X)] - \pi^{\neg L}_{A}(M) \cdot [1 - P(L_A)]$
Total difference in proportions	$\pi_{X-A}(M) = \pi^{L}_{X-A}(M) + \pi^{\neg L}_{X-A}(M)$

Table 28 Decomposition of the differences in the proportion of male-biased genes at the stationary due to candidate and non-candidate genes. Total difference in the proportion of male-biased genes at equilibrium, the % contributions from labile genes, and the % contribution from non-labile genes are highlighted (gray).

Species	$\pi_X^L(M)$	$\pi_A^L(M)$	$\pi_X^{\neg L}(M)$	$\pi_A^{\neg L}(M)$	$P(L_X)$	$P(L_A)$	$\pi_{X-A}^L(M)$	$\pi_{X-A}^{\neg L}(M)$	$\pi_{X-A}(M)$	$\%\, L$	$\%\, \neg L$
Dvir	0.00	0.00	0.00	0.01	0.04	0.05	0.00	-0.01	-0.01	0.00	1.00
Dmoj	0.00	0.10	0.01	0.01	0.04	0.05	0.00	0.00	-0.01	0.73	0.27
Dpse	0.12	0.48	0.04	0.09	0.04	0.05	-0.02	-0.04	-0.06	0.31	0.69
Dana	0.00	0.03	0.01	0.00	0.04	0.05	0.00	0.01	0.01	0.13	0.87
Dyak	0.00	0.37	0.02	0.03	0.04	0.05	-0.02	-0.01	-0.03	0.65	0.35
Dsim	0.18	0.26	0.01	0.02	0.04	0.05	-0.01	-0.01	-0.02	0.35	0.65
Dmel	0.64	0.54	0.05	0.11	0.04	0.05	0.00	-0.06	-0.06	0.01	0.99

Table 29 Effects of increasing the relative proportion of labile genes on the X chromosome.

Species	$\pi_X^L(M)$	$\pi_A^L(M)$	$\pi_X^{\neg L}(M)$	$\pi_A^{\neg L}(M)$	$P(L_X)$	$P(L_A)$	$\pi_{X-A}^L(M)$	$\pi_{X-A}^{\neg L}(M)$	$\pi_{X-A}(M)$	$\%\, L$	$\%\, \neg L$
Dvir	0.00	0.00	0.00	0.01	0.09	0.00	0.00	-0.01	-0.01	0.00	1.00
Dmoj	0.00	0.10	0.01	0.01	0.09	0.00	0.00	0.00	0.00	0.00	1.00
Dpse	0.12	0.48	0.04	0.09	0.09	0.00	0.01	-0.05	-0.04	0.18	0.82
Dana	0.00	0.03	0.01	0.00	0.09	0.00	0.00	0.01	0.01	0.00	1.00
Dyak	0.00	0.37	0.02	0.03	0.09	0.00	0.00	-0.01	-0.01	0.00	1.00
Dsim	0.18	0.26	0.01	0.02	0.09	0.00	0.02	-0.01	0.00	0.57	0.43
Dmel	0.64	0.54	0.05	0.11	0.09	0.00	0.06	-0.07	-0.01	0.46	0.54

Table 30 Effects of increasing the relative proportion of labile genes on the autosome.

Species	$\pi_X^L(M)$	$\pi_A^L(M)$	$\pi_X^{\neg L}(M)$	$\pi_A^{\neg L}(M)$	$P(L_X)$	$P(L_A)$	$\pi_{X-A}^L(M)$	$\pi_{X-A}^{\neg L}(M)$	$\pi_{X-A}(M)$	$\%\, L$	$\%\, \neg L$
Dvir	0.00	0.00	0.00	0.01	0.00	0.09	0.00	-0.01	-0.01	0.00	1.00
Dmoj	0.00	0.10	0.01	0.01	0.00	0.09	-0.01	0.00	-0.01	0.93	0.07
Dpse	0.12	0.48	0.04	0.09	0.00	0.09	-0.04	-0.04	-0.08	0.54	0.46
Dana	0.00	0.03	0.01	0.00	0.00	0.09	0.00	0.01	0.01	0.21	0.79
Dyak	0.00	0.37	0.02	0.03	0.00	0.09	-0.03	-0.01	-0.04	0.81	0.19
Dsim	0.18	0.26	0.01	0.02	0.00	0.09	-0.02	-0.01	-0.03	0.72	0.28
Dmel	0.64	0.54	0.05	0.11	0.00	0.09	-0.05	-0.05	-0.10	0.48	0.52

Table 31 Effects of increasing the overall proportion of labile genes across the genome (5-fold).

Species	$\pi_X^L(M)$	$\pi_A^L(M)$	$\pi_X^{\neg L}(M)$	$\pi_A^{\neg L}(M)$	$P(L_X)$	$P(L_A)$	$\pi_{X-A}^L(M)$	$\pi_{X-A}^{\neg L}(M)$	$\pi_{X-A}(M)$	% L	% $\neg L$
Dvir	0.00	0.00	0.00	0.01	0.20	0.25	0.00	-0.01	-0.01	0.00	1.00
Dmoj	0.00	0.10	0.01	0.01	0.20	0.25	-0.02	0.00	-0.03	0.96	0.04
Dpse	0.12	0.48	0.04	0.09	0.20	0.25	-0.09	-0.03	-0.13	0.74	0.26
Dana	0.00	0.03	0.01	0.00	0.20	0.25	-0.01	0.01	0.00	0.48	0.52
Dyak	0.00	0.37	0.02	0.03	0.20	0.25	-0.09	-0.01	-0.10	0.93	0.07
Dsim	0.18	0.26	0.01	0.02	0.20	0.25	-0.03	-0.01	-0.04	0.78	0.22
Dmel	0.64	0.54	0.05	0.11	0.20	0.25	0.00	-0.04	-0.05	0.08	0.92

BIBLIOGRAPHY

Arbeitman, M. N., A. A. Fleming, M. L. Siegal, B. H. Null & B. S. Baker (2004) A genomic analysis of Drosophila somatic sexual differentiation and its regulation. *Development,* 131**,** 2007 - 2021.

Arbeitman, M. N., E. E. Furlong, F. Imam, E. Johnson, B. H. Null, B. S. Baker, M. A. Krasnow, M. P. Scott, R. W. Davis & K. P. White (2002) Gene expression during the life cycle of Drosophila melanogaster. *Science,* 297**,** 2270 - 2275.

Bachtrog, D., N. R. T. Toda & S. Lockton (2010) Dosage Compensation and Demasculinization of X Chromosomes in Drosophila. *Current Biology,* 20**,** 1476-1481.

Betran, E., K. Thornton & M. Long (2002) Retroposed new genes out of the X in Drosophila. *Genome Res.,* 12**,** 1854-1859.

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

Chintapalli, V. R., J. Wang & J. A. T. Dow (2007) Using FlyAtlas to identify better Drosophila models of human disease. *Nat Genet,* 39**,** 715 - 720.

Cobb, M., K. Connolly & B. Burnet (1985) Courtship Behaviour in the Melanogaster Species Sub-Group of Drosophila. *Behaviour,* 95**,** 203-230.

Cochran, W. G. (1968) The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies. *Biometrics,* 24**,** 295-313.

Comeron, J. M., M. Kreitman & M. Aguadé (1999) Natural Selection on Synonymous Sites Is Correlated With Gene Length and Recombination in Drosophila. *Genetics,* 151**,** 239-249.

Connallon, T. & L. L. Knowles (2005) Intergenomic conflict revealed by patterns of sex-biased gene expression. *Trends Genet.,* 21**,** 495-499.

Consortium, *D.* G. (2007) Evolution of genes and genomes on the Drosophila phylogeny. *Nature,* 450**,** 203-218.

Cutter, A. D. & S. Ward (2005) Sexual and temporal dynamics of molecular evolution in C. elegans development. *Mol. Biol. Evol.,* 22**,** 178-188.

Darwin, C. 1871. *The Descent of Man, and Selection in Relation to Sex*. London: Murray.

Drummond, D. A., J. D. Bloom, C. Adami, C. O. Wilke & F. H. Arnold (2005) Why highly expressed proteins evolve slowly. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 14338-14343.

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

Ewing, A. W. (1964) The influence of wing area on the courtship behaviour of Drosophila melanogaster. *Animal Behaviour,* 12**,** 316-320.

Fraser, H. B., A. E. Hirsh, L. M. Steinmetz, C. Scharfe & M. W. Feldman (2002) Evolutionary rate in the protein interaction network. *Science,* 296**,** 750 - 752.

Gavrilets, S. (2000) Rapid evolution of reproductive barriers driven by sexual conflict. *Nature,* 403**,** 886-889.

Jonathan, R. G., K. C. Adam & R. R. William (2002) The X chromosome is a hot spot for sexually antagonistic fitness variation. *Proceedings of the Royal Society B: Biological Sciences,* 269**,** 499-505.

Jordan, I. K., Y. Wolf & E. Koonin (2003) No simple dependence between protein evolution rate and the number of protein-protein interactions: only the most prolific interactors tend to evolve slowly. *BMC Evolutionary Biology,* 3**,** 1.

Khil, P. P., B. Oliver & R. D. Camerini-Otero (2005) X for intersection: retrotransposition both on and off the X chromosome is more frequent. *Trends Genet,* 21**,** 3-7.

Kirkwood, T. B. L. & M. R. Rose (1991) Evolution of Senescence: Late Survival Sacrificed for Reproduction. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* 332**,** 15-24.

Larracuente, A. M., T. B. Sackton, A. J. Greenberg, A. Wong, N. D. Singh, D. Sturgill, Y. Zhang, B. Oliver & A. G. Clark (2008) Evolution of protein-coding genes in Drosophila. *Trends in Genetics,* 24**,** 114-123.

Lemos, B., B. R. Bettencourt, C. D. Meiklejohn & D. L. Hartl (2005a) Evolution of Proteins and Gene Expression Levels are Coupled in Drosophila and are Independently Associated with mRNA Abundance, Protein Length, and Number of Protein-Protein Interactions. *Mol Biol Evol,* 22**,** 1345-1354.

--- (2005b) Evolution of Proteins and Gene Expression Levels are Coupled in Drosophila and are Independently Associated with mRNA Abundance, Protein Length, and Number of Protein-Protein Interactions. *Molecular Biology and Evolution,* 22**,** 1345-1354.

Liao, B.-Y., N. M. Scott & J. Zhang (2006) Impacts of Gene Essentiality, Expression Pattern, and Gene Compactness on the Evolutionary Rate of Mammalian Proteins. *Molecular Biology and Evolution,* 23**,** 2072-2080.

Mank, J. E., L. Hultin-Rosenberg, M. Zwahlen & H. Ellegren (2008) Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. *Am. Nat.*

Marais, G., P. Nouvellet, P. D. Keightley & B. Charlesworth (2005) Intron Size and Exon Evolution in Drosophila. *Genetics,* 170**,** 481-485.

Meiklejohn, C. D., J. Parsch, J. M. Ranz & D. L. Hartl (2003) Rapid evolution of male-biased gene expression in Drosophila. *Proc. Natl Acad. Sci. USA,* 100**,** 9894-9899.

Meisel, R. P. (2011) Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein coding sequence evolution. *Molecular Biology and Evolution*.

Oliver, B. & M. Parisi (2004) Battle of the Xs. *Bioessays,* 26**,** 543-548.

Panhuis, T. M. & W. J. Swanson (2006) Molecular evolution and population genetic analysis of candidate female reproductive genes in Drosophila. *Genetics,* 173**,** 2039-2047.

Paradis, E., J. Claude & K. Strimmer (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics,* 20**,** 289-290.

Parisi, M., R. Nuttall, D. Naiman, G. Bouffard, J. Malley, J. Andrew, S. Eastman & B. Oliver (2003a) Paucity of genes on the Drosophila X chromosome showing male-biased expression. *Science,* 299**,** 697 - 700.

Parisi, M., R. Nuttall, D. Naiman, G. Bouffard, J. Malley, J. Andrews, S. Eastman & B. Oliver (2003b) Paucity of Genes on the Drosophila X Chromosome Showing Male-Biased Expression. *Science,* 299**,** 697-700.

Paterson, S., T. Vogwill, A. Buckling, R. Benmayor, A. J. Spiers, N. R. Thomson, M. Quail, F. Smith, D. Walker, B. Libberton, A. Fenton, N. Hall & M. A. Brockhurst (2010) Antagonistic coevolution accelerates molecular evolution. *Nature,* 464**,** 275-278.

Pearl, J. 2009. Letter to the Editor: Remarks on the Method of Propensity Score. . 1415–1424. Statistic in Medicine.

Powell, J. R. 1997. *Progress and Prospects in Evolutionary Biology: The Drosophila Model*. New York: Oxford University Press, Inc.

Ranz, J. M., C. I. Castillo-Davis, C. D. Meiklejohn & D. L. Hartl (2003) Sex-Dependent Gene Expression and Evolution of the Drosophila Transcriptome. *Science,* 300**,** 1742-1745.

Rhen, T. (2000) Sex-Limited Mutations and the Evolution of Sexual Dimorphism. *Evolution,* 54**,** 37-43.

Rice, W. R. (1984) Sex chromosomes and the evolution of sexual dimorphism. *Evolution,* 38**,** 735-742.

Rice, W. R. & A. K. Chippindale (2001) Intersexual ontogenetic conflict. *Journal of Evolutionary Biology,* 14**,** 685-693.

Rice, W. R. & B. Holland (1997) The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen. *Behav. Ecol. Sociobiol.,* 41**,** 1-10.

Rosenbaum, P. R. 2002. *Observational Studies*. Springer.

Rosenbaum, P. R. & D. B. Rubin (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika,* 70**,** 41-55.

Rubin, D. B. (2005) Causal Inference Using Potential Outcomes. *Journal of the American Statistical Association,* 100**,** 322-331.

Schluter, D., T. Price, A. O. Mooers & D. Ludwig (1997) Likelihood of ancestor states in adaptive radiation. *Evolution,* 51**,** 1699-1711.

Stuart, E. A. & D. B. Rubin. 2007. Best Practices in Quasi-Experimental Designs: Matching methods for causal inference. . In *Best practices in quantitative methods,* ed. J. Osborne, 155-176. Thousand Oaks,CA: Sage.

Sturgill, D., Y. Zhang, M. Parisi & B. Oliver (2007) Demasculinization of X chromosomes in the Drosophila genus. *Nature,* 450**,** 238-241.

Swanson, W. J. & V. D. Vacquier (2002) The rapid evolution of reproductive proteins. *Nature Rev. Genet.,* 3**,** 137-144.

Swanson, W. J., Z. Yang, M. F. Wolfner & C. F. Aquadro (2001) Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. *Proceedings of the National Academy of Sciences,* 98**,** 2509-2514.

Trivers, R. L. (1972) Parental Investment and Sexual Selection.

Vicoso, B. & B. Charlesworth (2006) Evolution on the X chromosome: unusual patterns and processes. *Nat Rev Genet,* 7**,** 645 - 653.

Zhang, L. & W.-H. Li (2004) Mammalian Housekeeping Genes Evolve More Slowly than Tissue-Specific Genes. *Molecular Biology and Evolution,* 21**,** 236-239.

Zhang, Y., D. Sturgill, M. Parisi, S. Kumar & B. Oliver (2007) Constraint and turnover in sex-biased gene expression in the genus Drosophila. *Nature,* 450**,** 233-237.

Zhang, Z., T. M. Hambuch & J. Parsch (2004) Molecular evolution of sex-biased genes in Drosophila. *Mol. Biol. Evol.,* 21**,** 2130-2139.

Zhang, Z. & J. Parsch (2005) Positive correlation between evolutionary rate and recombination rate in Drosophila genes with male-biased expression. *Mol. Biol. Evol.,* 22**,** 1945-1947.

Appendix A: Methods for inferring states of sex-biased gene expression and sex comparative divergence

In the text below, methods for assigning states of sex-biased gene expression and states of the sex comparative index are described. In the second chapter, only states of sex-biased gene expression inferred for *D. melanogaster* are used. In the third chapter, both states of sex-biased gene expression and states of the sex comparative index inferred throughout the *Drosophila* phylogeny are used.

Data on sex-dependent gene expression

Data on sex-biased gene expression comes from Zhang et al.'s multi-species study of sex-biased gene expression for 7 species of *Drosophila*. Processed microarray data were obtained from the Gene Expression Omnibus under accession GSE6640, where a full description of their experimental design can also be found.

To summarize, gene expression was assayed using a standard 60-mer Nimblegen platform design for *D. Melanogaster* and multiple custom 50-mer expression arrays designed using genomic assemblies for 6 species of *Drosophila* (*D. simulans, D. yakuba, ***D. ananassae, D. pseudoobscura, D. virilis and D. mojavensis*). Two-dye hybridizations were performed for each species, where the two channels correspond to hybridizations based on samples derived from individual male and female whole adults. For each species, at least two biological replicates were conducted with dye-swapping (that is, a minimum of 4 hybridizations were conducted per species). Each hybridization was processed according to quality control criteria (see Zhang *et al.*), and expression values for each probe were derived using variance stabilization normalization applied across the set of passed hybridizations. A probeset for a given gene comprises the set of oligonucleotides that uniquely and perfectly map to its sequence, where biological sources of variation in intensity the measured intensities among probes include among splice and among individual variation. States of sex-biased gene expression are typically determined by comparing the distribution of post-processed male and female expression as measured across probes belonging to same probeset using a non-parametric hypothesis test, typically the Mann-Whitney U-test, where p-values are false discovery rate adjusted for multiple comparison (Zhang et al. 2007).

Here, array probes were first re-mapped using the most up to date GLEANR gene prediction models and assemblies (pre-computed data obtained from FlyBase: FB2011_10). An approximate permutation test was used to assign states of sex-biased expression and sex comparative divergence according to the procedure described below.

Definitions:

Let $\boldsymbol{M_i}$ and $\boldsymbol{F_i}$ denote the complete probeset for a single gene expressed in species *i*:

$$\boldsymbol{M_i} = \left(m_{11}, m_{12}, m_{13}, \ldots, m_{in_{m_i}}\right)$$

$$\boldsymbol{F_i} = \left(f_{11}, f_{12}, f_{13}, \ldots, f_{in_{f_i}}\right)$$

m_{ij} and f_{ij}correspond to processed (log-transformed and normalized) intensity values for the jth probe in the probeset corresponding to gene i. Above, n_{m_i} and n_{mf_i} refer to the total number of probes in each probeset (these numbers may differ due to the application of quality control criteria during processing).

$\bar{\boldsymbol{M}}$ and $\bar{\boldsymbol{F}}$ contain mean female and male expression values evaluated for each of the 7 orthologs:

$$\bar{\boldsymbol{M}} = (\bar{M}_{\boldsymbol{1}}, \bar{M}_{\boldsymbol{2}}, \ldots, \bar{M}_{\boldsymbol{7}})$$

$$\bar{\boldsymbol{F}} = (\bar{F}_1, \bar{F}_2, \ldots, \bar{F}_7)$$

To assign states of sex-bias to each tip and internal node, a test statistic (a.k.a sex-bias) is defined for each internal node or tip:

$$\tau_i = \bar{M}_i - \bar{F}_i$$

A second test statistic is defined in order to assign states of the sex-comparative index to each branch, where the tilde in the equation below indicates that a difference may be taken between two ancestral inferences ($\tilde{M}_i$ or $\tilde{F}_i$, as defined below) or between an ancestral inference and the mean of the measured values for a given probeset ($\bar{M}_i$ or $\bar{F}_i$, as defined above):

$$\zeta_b = \frac{\left|\Delta\tilde{F}_{ij}\right| - \left|\Delta\tilde{M}_{ij}\right|}{\sqrt{\Delta\tilde{F}_{ij}{}^2 + \Delta\tilde{M}_{ij}{}^2}}$$

Approximate permutation test

Permutations tests can be used to assign ancestral states. Because there are so many possible ways of ordering the data, an approximate rather than an exact permutation test procedure was conducted. The approximate permutation test uses Monte Carlo simulation to derive the distribution of each of the test statistics defined above. For each tip, there are $N_i = \binom{n_{m_i} + n_{f_i}}{n_{m_i}}$ possible ways of allocating the observed expression values into male and female groupings and there are $N = \sum_i N_i = \sum_i \binom{n_{m_i} + n_{f_i}}{n_{m_i}}$ possible allocations for the phylogeny. Approximate permutation tests were conducted according to the procedure described below.

Step 1: Calculate test statistics based on observed expression values

For tips:

For each species, the test statistic measuring sex-bias is calculated by taking the difference between the mean of the female and male expression values for each corresponding probeset.

$$\tau_i^{obs} = \bar{M}_i^{obs} - \bar{F}_i^{obs}$$

For internal nodes and branches:

To calculate the observed value for the sex-comparative index and for sex-bias for internal nodes, ancestral inference must first be conducted. Ancestral values were inferred using the maximum likelihood framework described above (**eq. [1]**), where $\bar{\boldsymbol{M}}^{obs}$ and $\bar{\boldsymbol{F}}^{obs}$correspond to $\boldsymbol{x}$ in the likelihood equation above. Following ancestral inference, the test statistics are calculated as follows:

$$\tau_i^{obs} = \tilde{M}_i^{obs} - \tilde{F}_i^{obs}$$

In the above, $\tilde{M}_\alpha^{obs}$ and $\tilde{F}_\alpha^{obs}$ refer to ancestral inferences for male and female expression derived for each ancestral node *i*.

Similarly, ancestral inferences are used to derive ζ_b^{obs} where $\Delta\tilde{M}_{ij}$ in the equation below may involve a difference between ancestral inferences or between an ancestral inference and an observed value.

$$\zeta_b^{obs} = \frac{\left|\Delta\tilde{M}_{ij}\right| - \left|\Delta\tilde{F}_{ij}\right|}{\sqrt{\Delta\tilde{F}_{ij}{}^2 + \Delta\tilde{M}_{ij}{}^2}}$$

Step 2: Simulate distributions for each test statistic

1. Select 1 of the *N* possible allocations by randomly shuffling the assignments of male and female expression at each tip
2. Compute τ_i^{obs}, and ζ_b^{sim}
3. Save values to $\boldsymbol{T_i}$ and $\boldsymbol{B_b}$
4. Repeat N_{sim} times

Step 3: Conduct hypothesis testing to assign states

$\boldsymbol{T}_i$ and $\boldsymbol{B_b}$ contain the simulated distributions for each statistic. To assign ancestral states of sex-biased gene expression, hypothesis testing was conducted by comparing observed values on the observed statistics to corresponding distributions obtained from simulation. In each case, p-values were false discovery rate adjusted for an experiment-wise level of $\alpha = 0.01$ (N = 3,505). The p-value cut-offs used here are the same as those that used to classify genes into male-biased, female-biased, and nonsex-biased groups in previous analyses on the same dataset, where both 'demasculinization' and 'faster-male' evolution were observed (Zhang et al. 2007). By analogy to existing conventions, the same p-value was used to assign states of sex comparative divergence.

More than 10,000 simulations were conducted. Tip assignments derived from this technique were compared to assignments derived from Mann-Whitney tests evaluated at the tips to assure the convergence of assignments with published values on sex-biased gene expression (Zhang et al. 2007).

Appendix B: Supplement to Chapter 2

Figure 20 Graphical representation of covariate balance and overlap for M vs. F (M1)

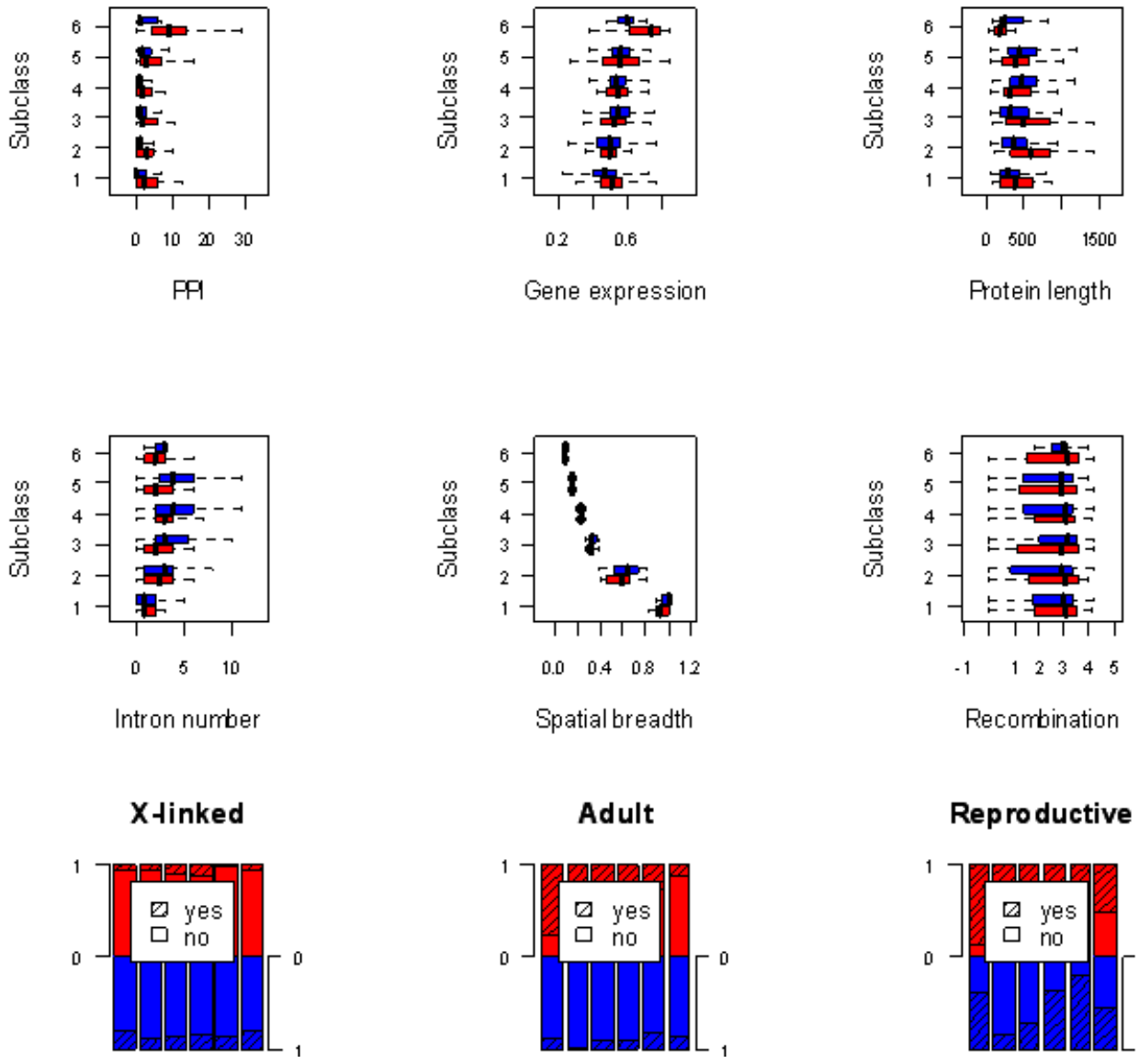


Figure 21 Graphical representation of covariate balance and overlap for M vs. N (M1)

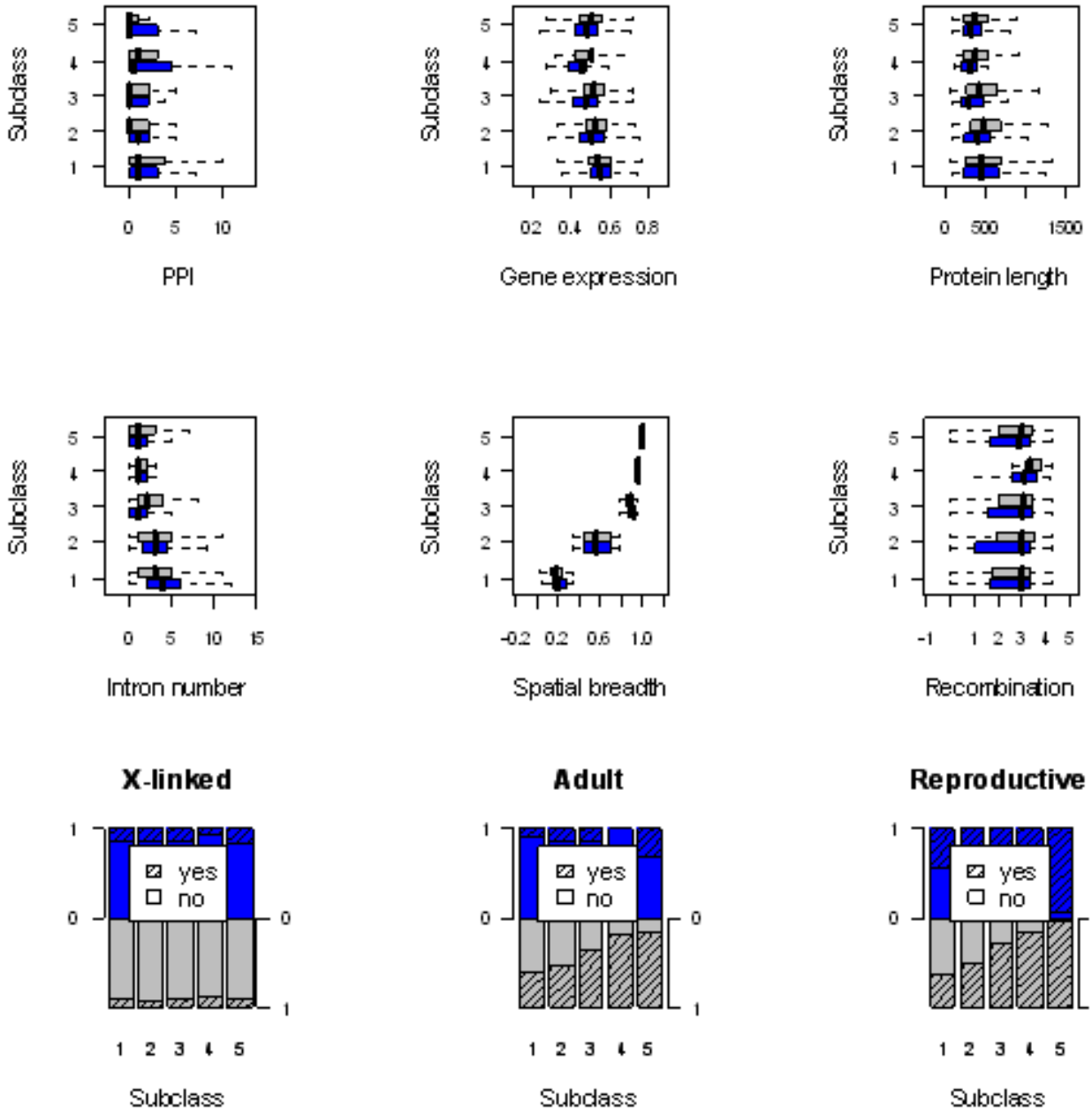


Figure 22 Graphical representation of covariate balance and overlap for F vs. N (M1)

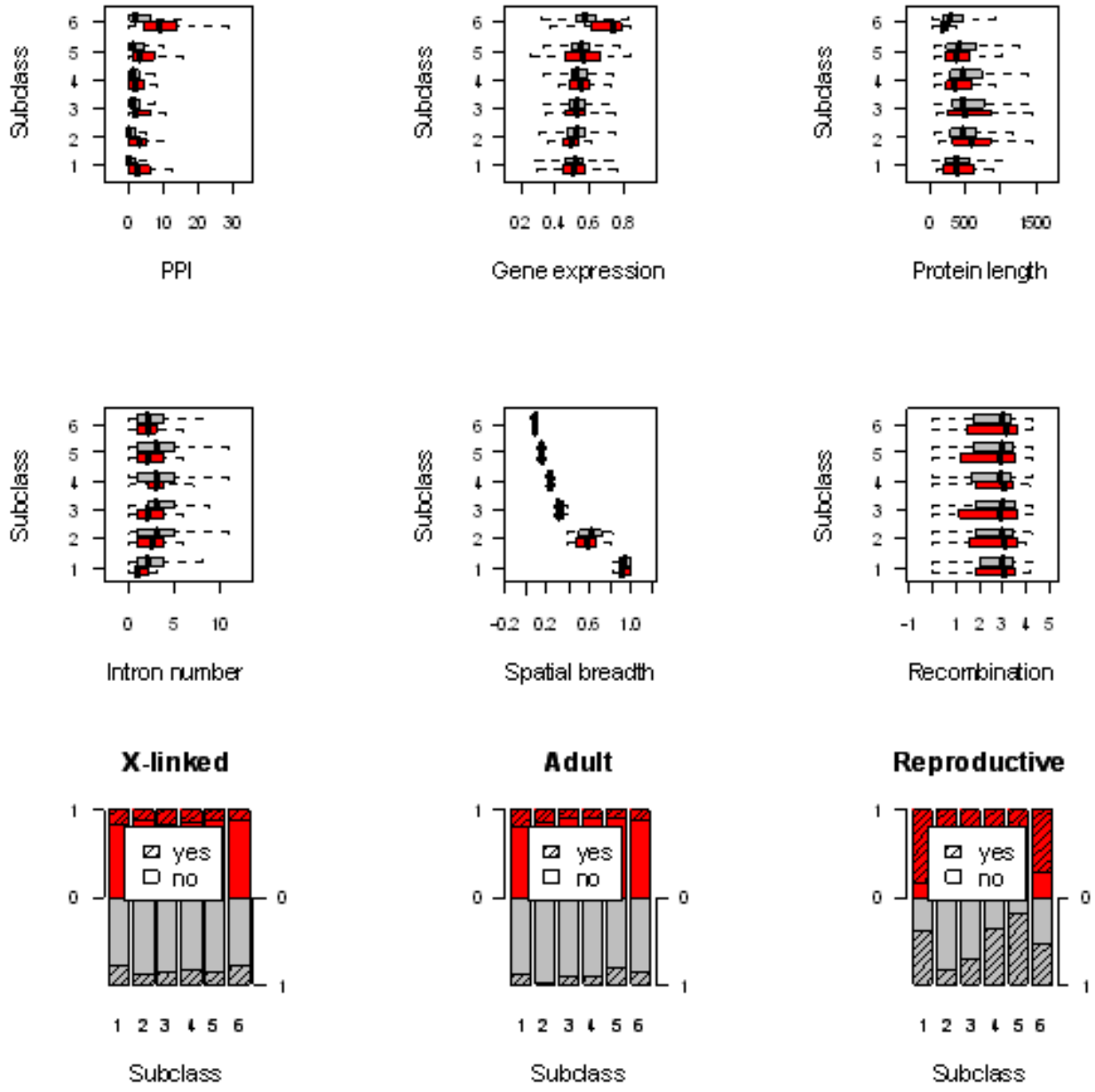


Figure 23 Graphical representation of covariate balance and overlap for M vs. F (M2)

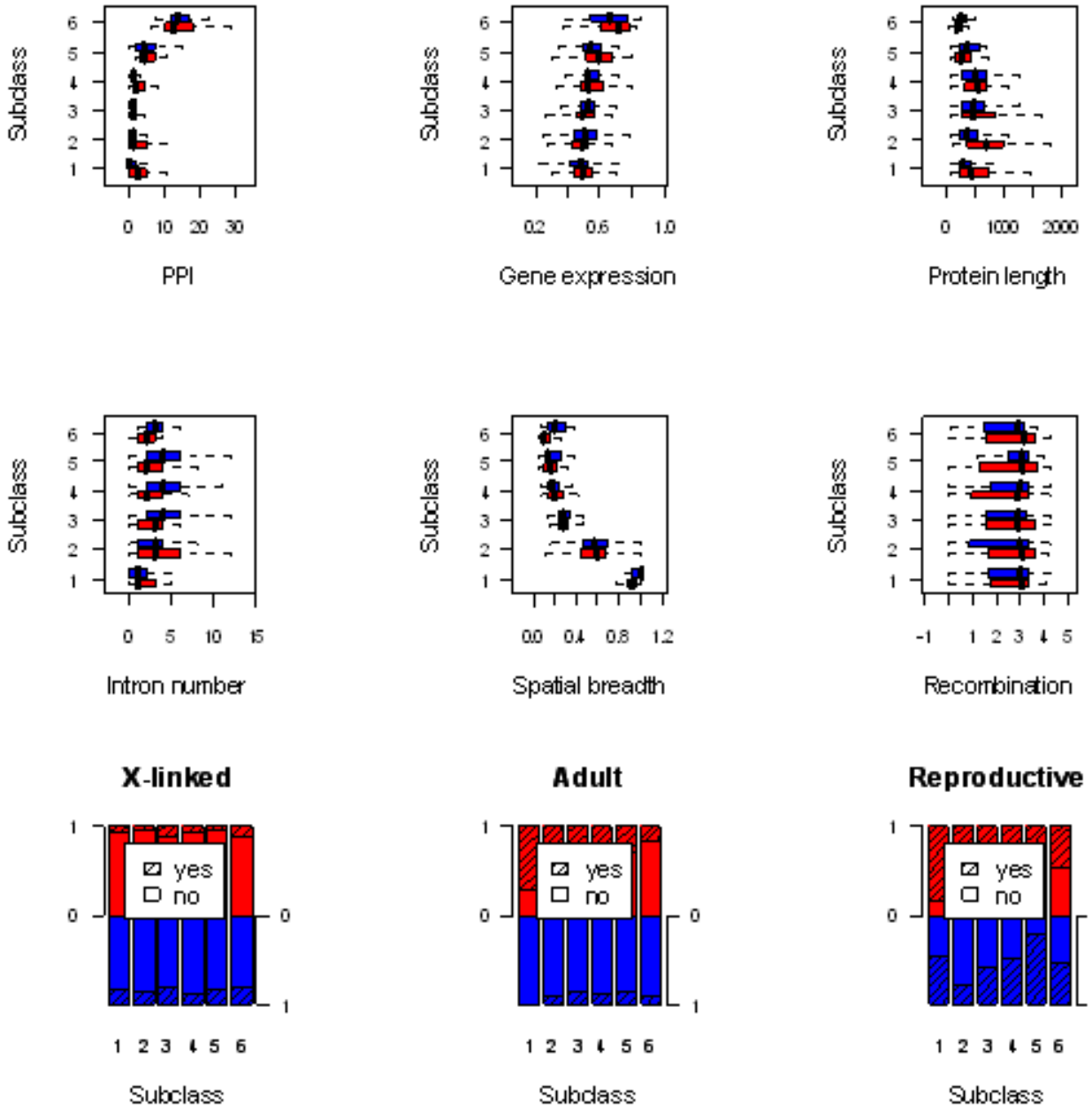


Figure 24 Graphical representation of covariate balance and overlap for M vs. N (M2)

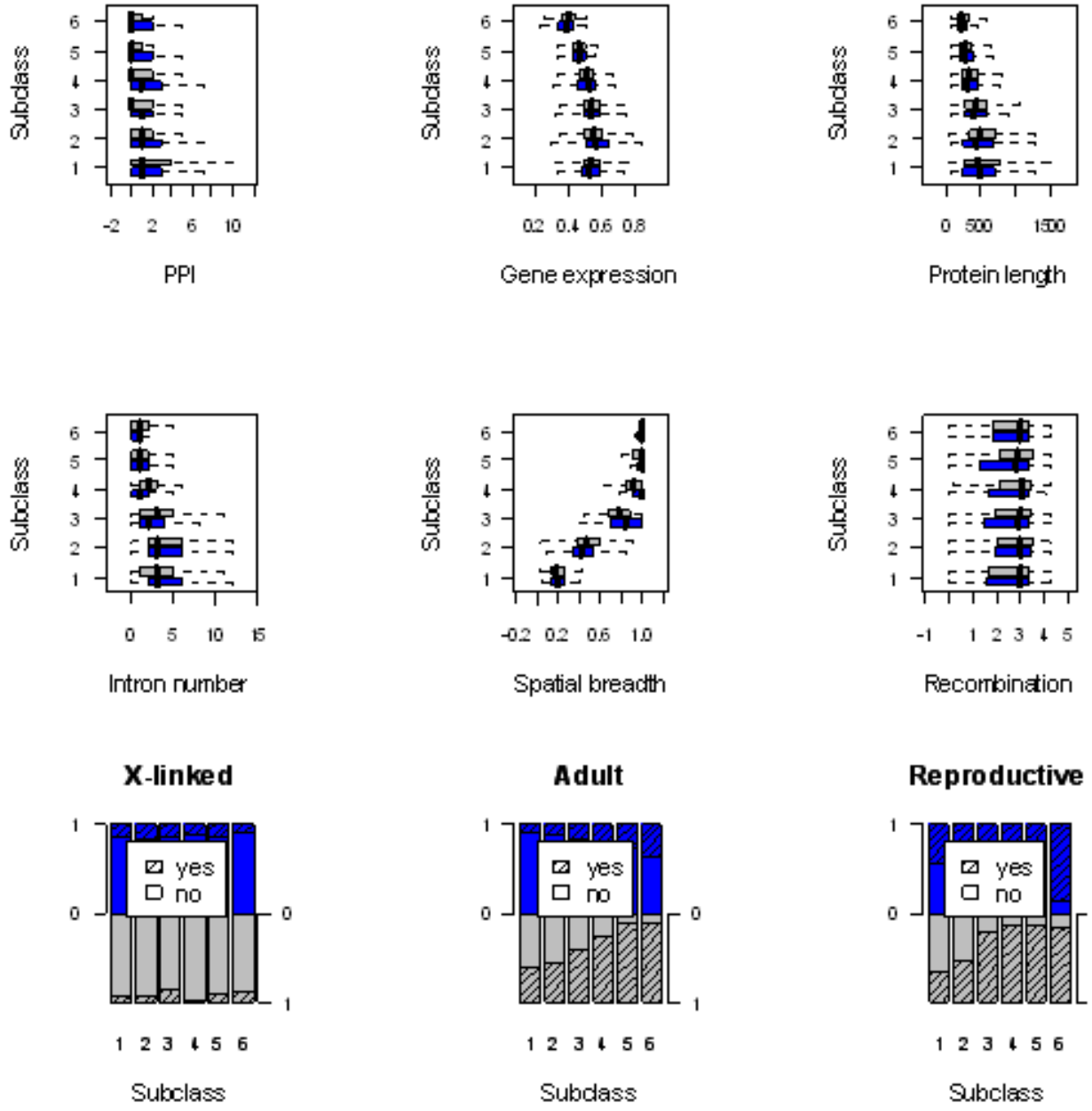


Figure 25 Graphical representation of covariate balance and overlap for F vs. N (M2)

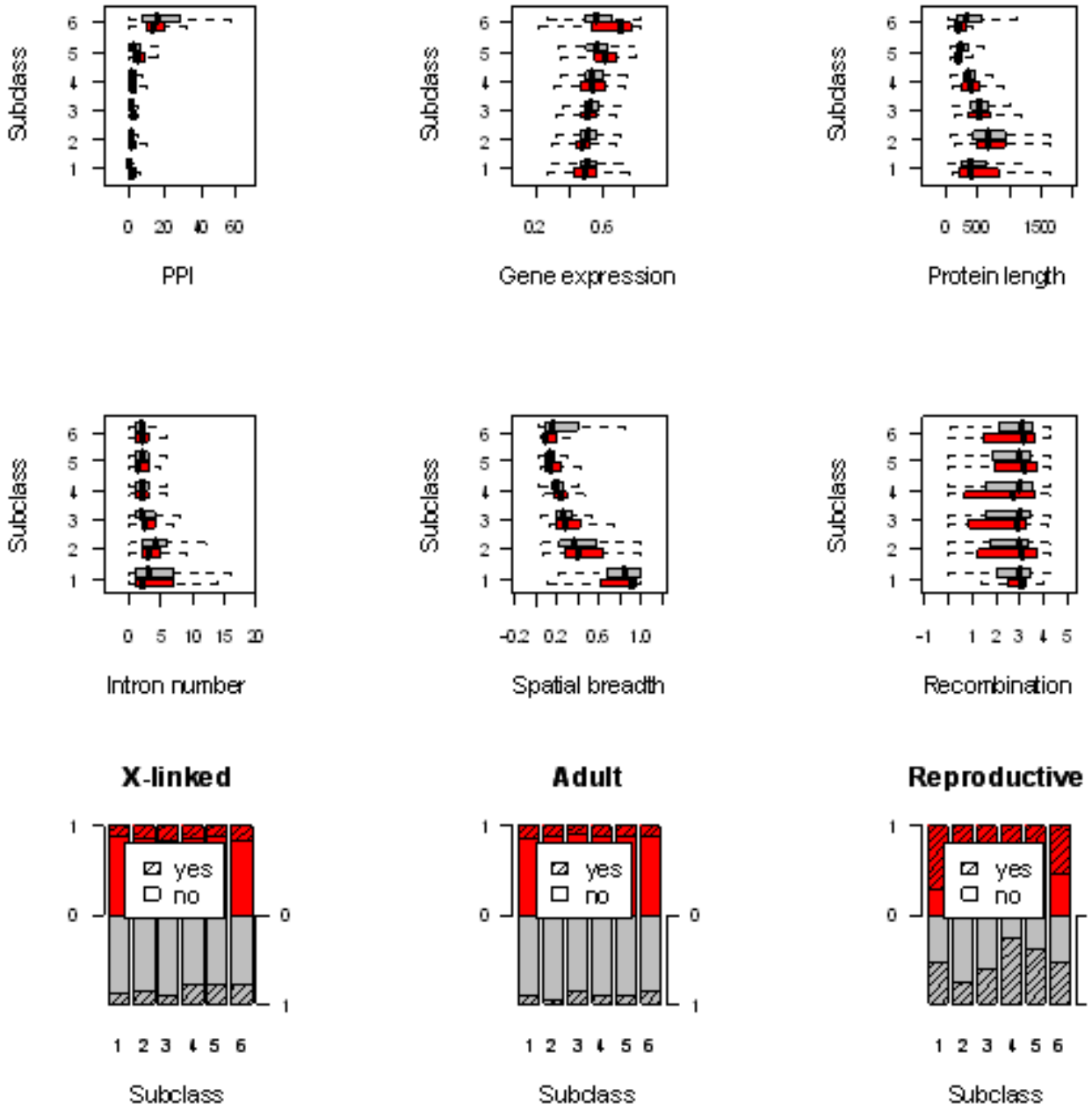


Figure 26 Graphical representation of covariate balance and overlap for M vs. F (M3)

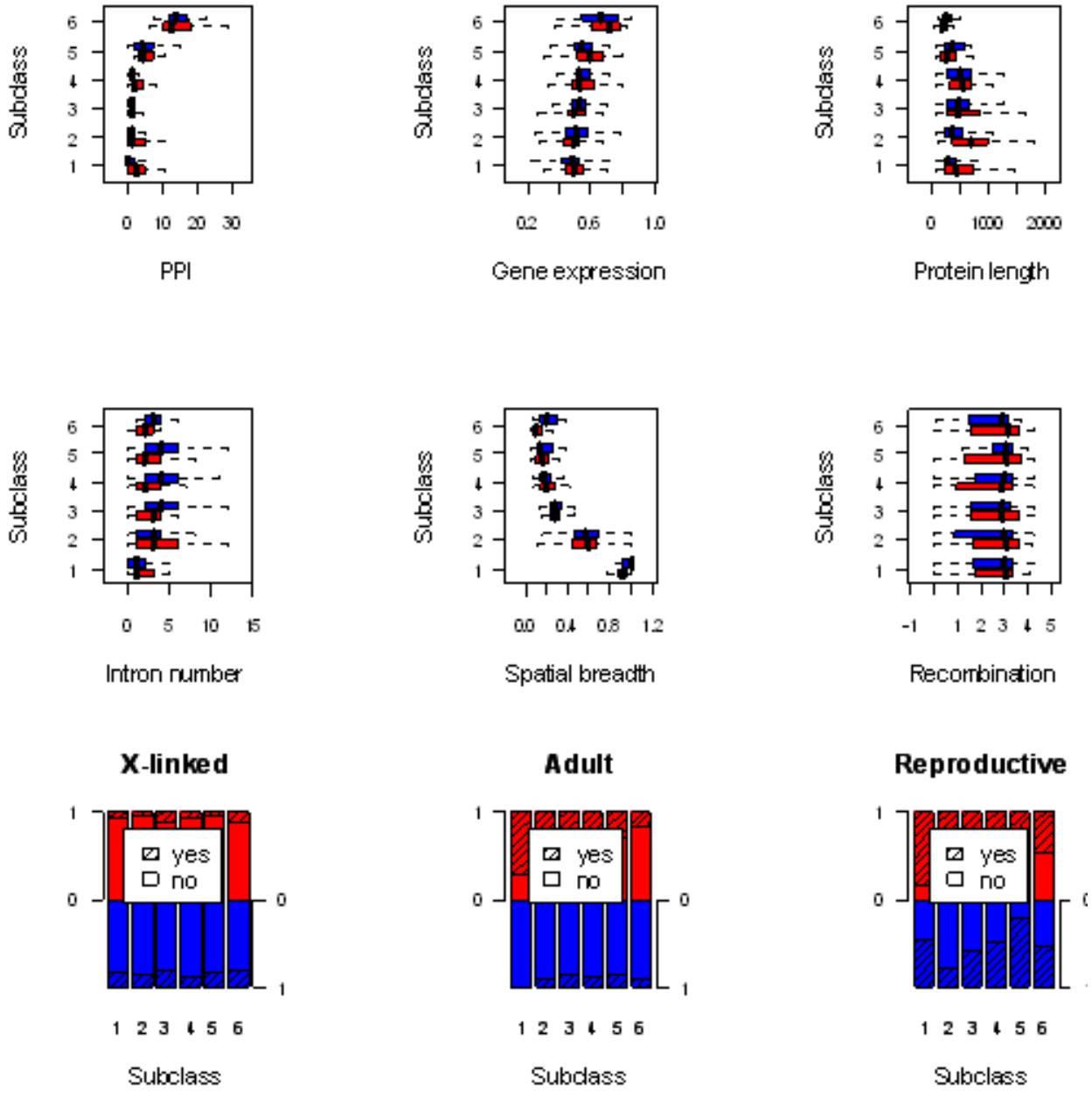


Figure 27 Graphical representation of covariate balance and overlap for M vs. N (M3)

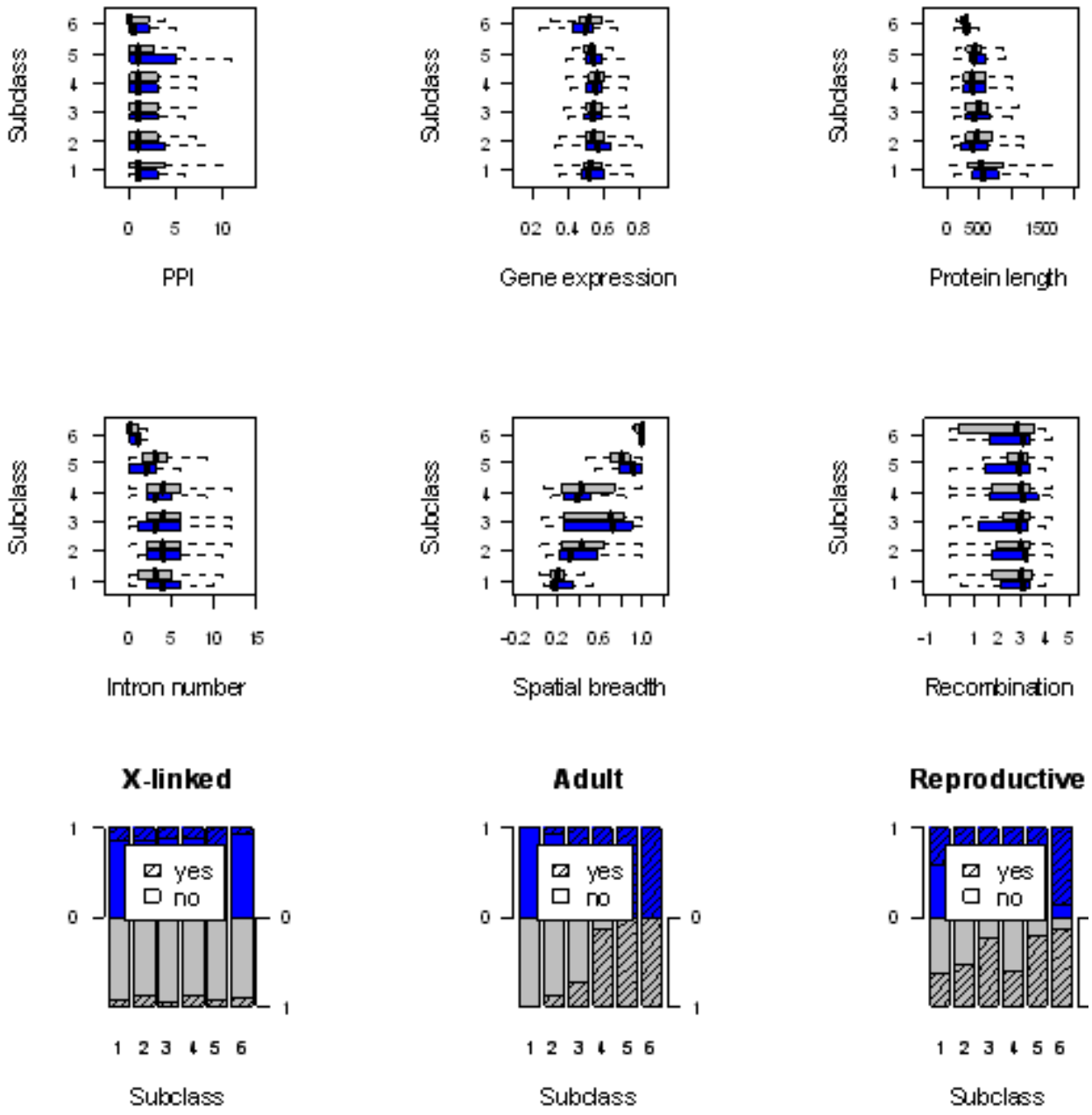


Figure 28 Graphical representation of covariate balance and overlap for F vs. N (M3)

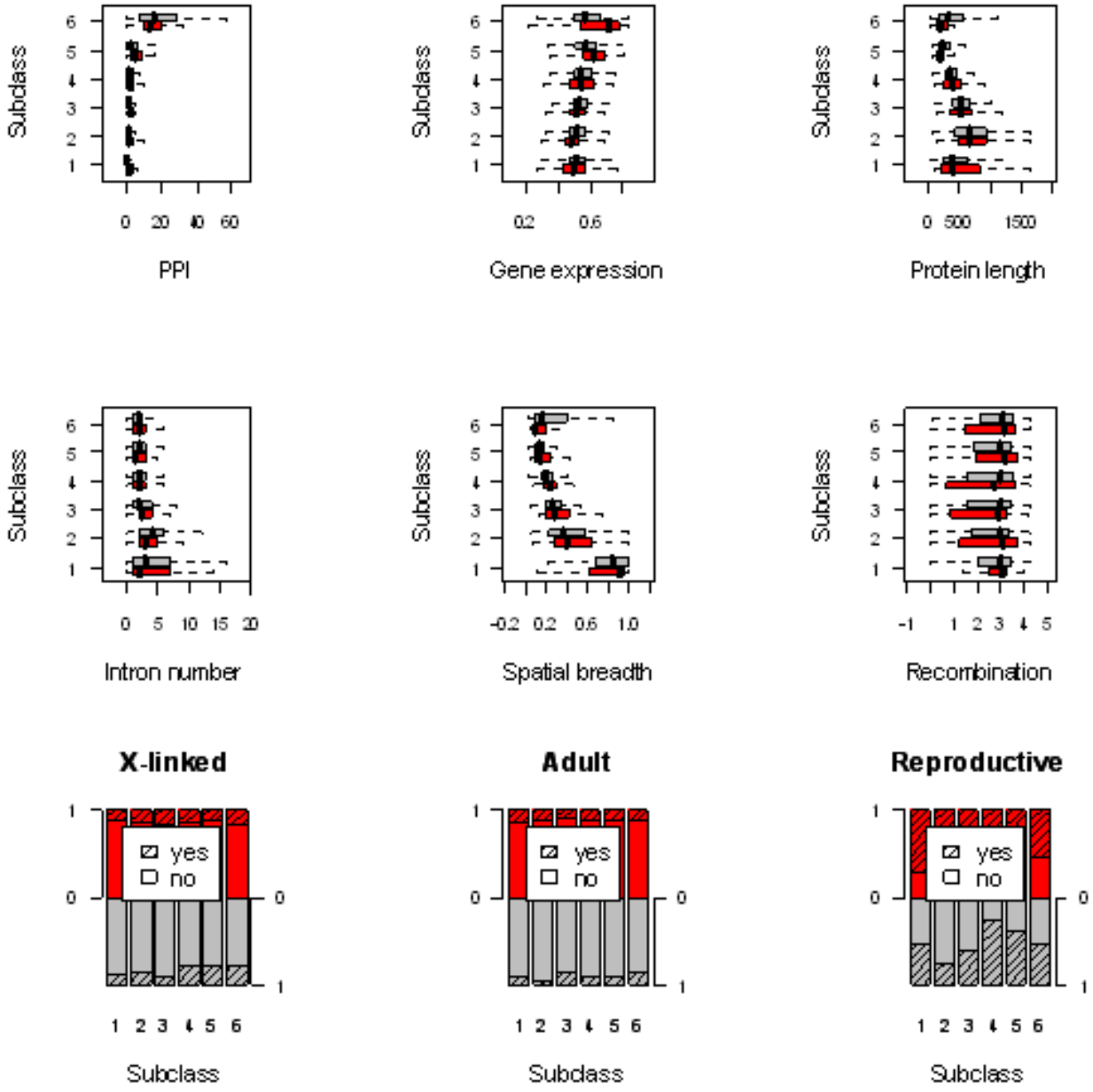


Table 32 Model criteria and best propensity score models for the M vs. F comparison

Crit.	Proposed confounders	(N_F,N_M,N_N)	Comp.	Best propensity-score model
M1	Spatial breadth, Reproductive	(324,915)	M vs. F	P-score = Spatial breadth
M2	PPI Gene expression Recombination Protein length Intron number Spatial breadth X-linkage Reproductive	(324,915)	M vs. F	P-score = Spatial breadth + PPI + Protein length + Spatial breadth x PPI + PPI x Protein length
M2	PPI Gene expression Recombination Protein length Intron number Spatial breadth X-linkage Reproductive Adult	(324,344)	M vs. F	P-score = Spatial breadth + PPI + Protein length + Tissue-bias + PPI + Protein length + Spatial breadth x PPI + PPI x Protein length

Table 33 Covariate balance for M1-M3 (F vs. M): From left to right, list of covariates, mean difference prior to adjustment, mean difference after adjustment, the percentage change in the mean difference before and after adjustment, and the balance score (β). Variables included in matching criteria for model selection are indicated (†).

Table 33a M1: F vs. M

Covariates	Diff. (no adj.)	Diff. (adj.)	Perc. Change	β
PPI	-2.46	-1.97	-20.07	0.23
Gene expression	-0.05	-0.02	-68.54	0.16
Recombination	-0.01	-0.04	381.07	0.03
Protein length	-87.58	-111.21	26.99	0.32
Intron number	-0.25	0.32	29.15	0.11
Spatial breadth†	0.35	0.03	-91.71	0.10
X-linkage	0.68	0.07	-89.01	0.08
Adult	0.79	-0.31	-60.66	0.34
Reproductive†	0.03	0.01	-62.16	0.02

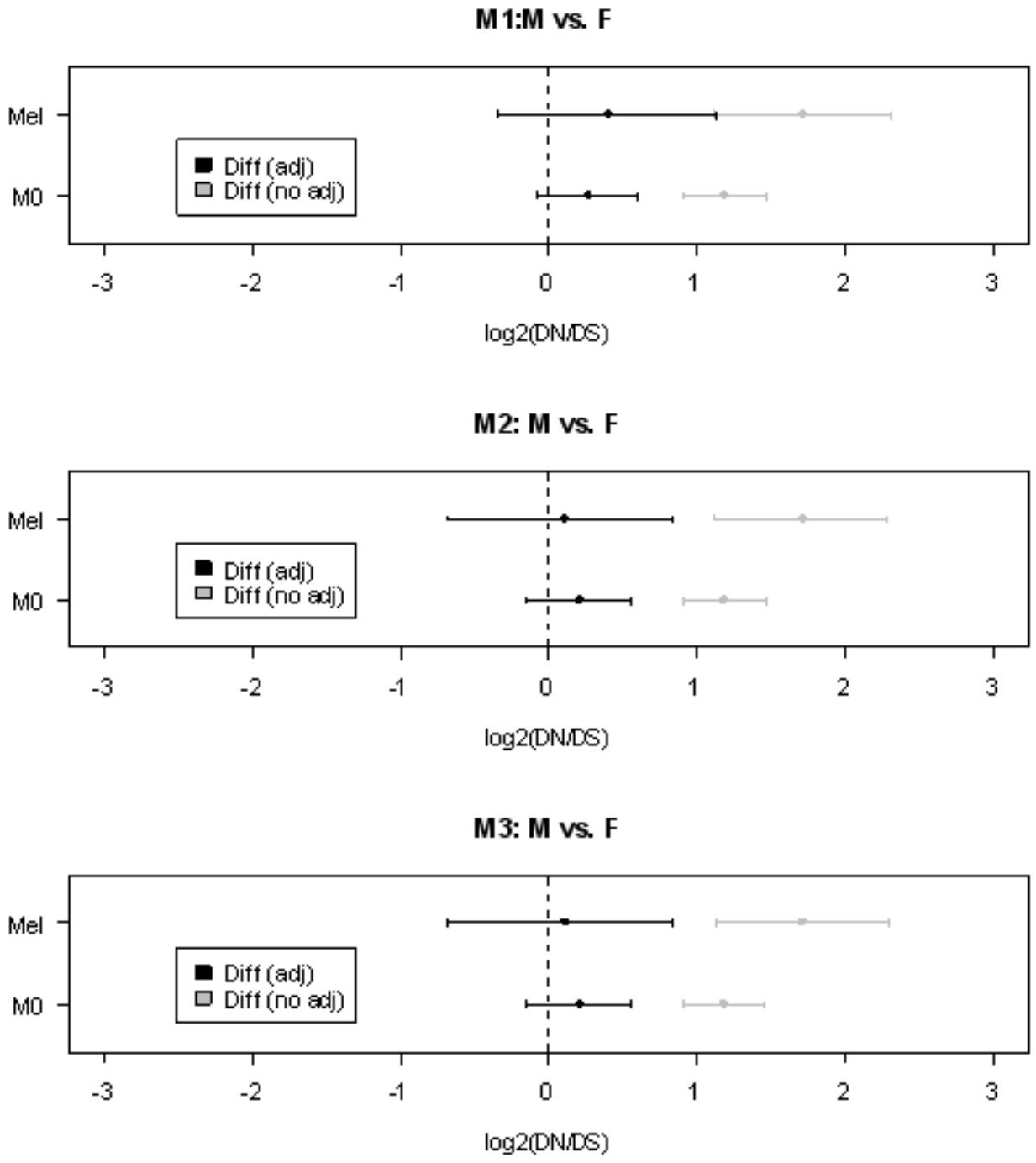
Table 33b M2: F vs. M

Covariates	Diff. (no adj.)	Diff. (adj.)	Perc. Change	β
PPI†	-2.46	-0.58	-76.37	0.07
Gene expression†	-0.05	-0.01	-78.93	0.10
Recombination†	-0.01	-0.03	224.79	0.02
Protein length†	-87.58	-126.22	44.12	0.35
Intron number†	-0.25	0.23	-5.69	0.08
Spatial breadth†	0.35	0.04	-89.54	0.12
X-linkage†	0.68	0.07	-89.01	0.07
Adult	0.79	-0.33	-58.76	0.33
Reproductive†	0.03	0.01	-70.86	0.01

Table 33c M3: F vs. M

Covariates	Diff. (no adj.)	Diff. (adj.)	Perc. Change	β
PPI†	-2.46	-1.97	-20.07	0.23
Gene expression†	-0.05	-0.02	-68.54	0.16
Recombination†	-0.01	-0.04	381.07	0.03
Protein length†	-87.58	-111.21	26.99	0.32
Intron number†	-0.25	0.32	29.15	0.11
Spatial breadth†	0.35	0.03	-91.71	0.10
X-linkage†	0.68	0.07	-89.01	0.08
Adult†	0.79	-0.31	-60.66	0.34
Reproductive†	0.03	0.01	-62.16	0.02

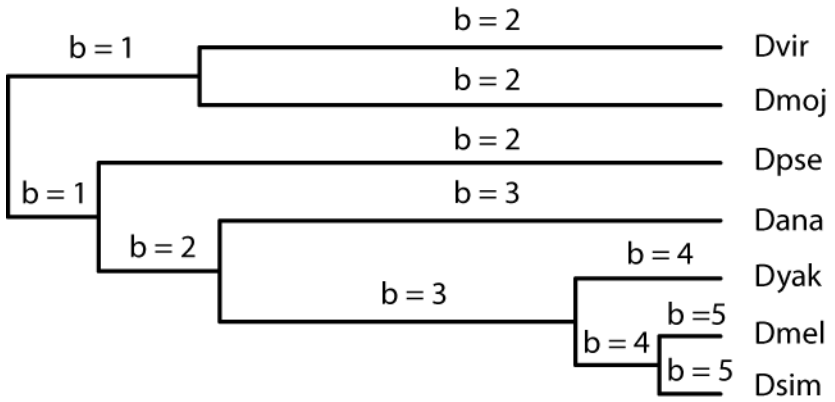
Figure 29 Effect estimates for direct comparison of F vs. M



Appendix C: Derivation of the expected value of the lability index based on states of the sex comparative index

In this appendix the derivation of the expected value for the lability index based on the state of the sex comparative index is shown. First, each branch is labeled from root to tip for each lineage. The lability index is a path-based statistic where the probability of transitions along each branch is first summed over the branches of each lineage and is then summed over all 7 species' lineages.

Figure 30 Graphic for deriving the expected value of the lability index based on the state of the sex comparative index. Each branch is labeled numerically from root to tip.



The probability that the b^{th} branch is a transition:

$P_b =$ Prob(b^{th} branch is a transition)

$=$ Prob(b^{th} is m or f) $*$ Prob(not all previous are n)*Prob(previous m or f is opposite to b^{th})

$=$ Prob(b^{th} is m or f) $*$ (1 − Prob(all previous are n)) $*$ Prob(previous m or f is opposite)

$= \alpha * \left(1 - (1-\alpha)^{b-1}\right) * 1/2$

$= \alpha/2 * (1 - (1-\alpha)^{b-1})$

The probability of transitions is evaluated for each branch along each species' lineage:

$$b = 1 \qquad P_1 = 0$$

$$b = 2 \qquad P_2 = \frac{\alpha}{2} * \left(1 - (1-\alpha)\right) = \frac{\alpha^2}{2}$$

$$b = 3 \qquad P_3 = \frac{\alpha}{2} * (1 - (1-\alpha)^2) = \frac{\alpha^2}{2} * (2 - \alpha)$$

$$b = 4 \qquad P_4 = \frac{\alpha}{2} * (1 - (1-\alpha)^3) = \frac{\alpha^2}{2} * (3 - 3\alpha + \alpha^2)$$

$$b = 5 \qquad P_5 = \frac{\alpha}{2} * (1 - (1-\alpha)^4) = \frac{\alpha^2}{2} * (4 - 6\alpha + 6\alpha^2 - \alpha^3)$$

Finally, the total number of transitions is summer over the lineages:

Expected number of transitions over tree = $\sum_{\text{branches in tree}} P_b = \frac{\alpha^2}{2} * (10 - 20\alpha + 10\alpha^2 - 2\alpha^3)$

Appendix D: Supplement to Chapter 3

A summary of results from analysis where the lability index is calculated from the sign of the raw index:

For assignments into labile and non-labile classes based on the raw sign of the comparative sex index, X-linked genes are marginally more labile than autosomal genes (H: $\lambda_X > \lambda_A$, p = 0.06, one-tailed t-test, Fig 3; Table 4), and labile genes (defined as $\lambda > \lambda_{exp} = 8$) are significantly overrepresented among X-linked genes (one-tailed Fisher Exact Test: $(\text{Odds ratio})_{X/A} = 1.39$, p = 0.003). No significant differences in tissue-biased gene expression were seen in comparisons of labile genes with non-labile genes, even among genes that are nonsex-biased throughout the phylogeny (Mann-Whitney test: W = 281250; p = 0.96, Fig), even when analysis was restricted to genes that are nonsex-biased throughout the phylogeny (Mann-Whitney test: W = 254066; p = 0.55). No under-/overrepresentations were seen for various tissue or stage specific genes, marginally underrepresented among in brain-specific genes (p_{adj}=0.052). Adjusting for tissue-bias, labile genes show no evidence of enhanced rates of gene evolution within any subclass of tissue-biased gene expression (two-tailed Mann Whitney applied within each subclasses of tissue-biased gene expression: H_{low}: $\omega_{lab} = \omega_{non\text{-}lab}$ (W = 77970, p-value = 0.55), H_{mid}: $\omega_{lab} = \omega_{non\text{-}lab}$ (W = 866670, p-value = 0.08), H_{high}: $\omega_{lab} = \omega_{non\text{-}lab}$ (W = 90479, p-value = 0.2582), nor is lability correlated with rates of sequence evolution (Kendall Tau Test: $\tau = 0.0086$, $z = 0.75$, p-value = 0.45).