Two forms of sexual dimorphism in gene expression in *Drosophila melanogaster*: their
coincidence and evolutionary genetics

Amardeep Singh and Aneil F. Agrawal[1]

Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Canada, M5S
3B2
[1]*E-mail: a.agrawal@utoronto.ca*

Impact Summary

Differences in how genes are expressed between the sexes allow males and females to differ, on average, in a variety of morphological, physiological, and behavioural traits. Using fruit fly data, we examine sexual dimorphism in two different forms of gene expression: (i) sex-biased gene expression (SBGE) quantifies sex differences in the total amount of expression from a gene and (ii) sexual dimorphic isoform usage (SDIU) indicates whether the sexes differ in their relative usage of different isoforms (structurally different proteins coded by the same gene). Based on whole body expression data, we find that these two forms of expression dimorphism co-occur at the same genes less often than expected by chance, possibly indicating that they represent alternative solutions to sexually antagonistic selection. A separate analysis indicates that patterns of quantitative genetic variation are more conducive to genes with SDIU evolving sex differences in total expression than those without SDIU. The observation that SBGE and SDIU are negatively associated suggests that genes with SDIU are not under sexually antagonistic selection for expression or that the variants that would allow for greater dimorphism have other negative effects. An analysis of patterns of DNA sequence variation at the population level suggests that genes with SDIU experience more adaptive evolution than those without. There is little evidence that SBGE is associated with adaptation, except for strongly male-biased genes.

Abstract

Sex differences in gene regulation is a proximate mechanism by which evolution can resolve
sexually antagonistic phenotypic selection. One form of differential gene regulation is sex-biased
gene expression (SBGE) in which the sexes differ in the amount a gene is expressed. Less
attention has been given to sexual dimorphism in isoform usage (SDIU), resulting from sex
differences in alternative splicing. Here, we use RNA-seq data from two tissue types (heads and
bodies) from 18 genotypes of adult *Drosophila melanogaster* to investigate SDIU and SBGE.
Both forms of expression dimorphism are both much more prevalent in the body than the head.
SDIU is less common among sex-biased than unbiased genes in the body, though the opposite
pattern occurs in the head. With respect to body expression, we examine associations of SDIU
and SBGE with several evolutionary genetic properties. The intersexual genetic correlation in
expression (r_{mf}) increases from female-biased to unbiased to male-biased genes but is drastically
reduced for genes with extreme male bias. r_{mf} is lower for genes with versus without SDIU,
suggesting that genes with SDIU should be able to evolve greater levels of SBGE holding all else
equal. SDIU has no detectable effect on relative non-synonymous diversity. In contrast, non-
synonymous diversity is elevated for genes with extreme sex-bias (particularly male-bias) but
other effects of SBGE are more subtle. SDIU is associated with reduced values of Tajima's *D*.
After accounting for correlated variables, sex-bias tends to reduce Tajima's *D*. Finally, *DoS* is
elevated for genes with SDIU, suggestive of higher rates of adaptive evolution. In contrast,
SBGE is not associated with *DoS* after excluding genes with extreme male-bias (for which *DoS*
is highly elevated). Together, these results show that both SDIU and SBGE are independently
and distinctly associated with evolutionary genetic statistics, suggesting that selection acts
somewhat differently on these two dimensions of expression dimorphism.

Introduction

In sexual populations, the sexes employ different reproductive strategies to maximize fitness and as result, shared phenotypes often exhibit divergent, sex-specific trait optima. Because the sexes largely share a genome, the genetic basis of such traits is often the same between the sexes. Thus, genetic variation underpinning such shared traits will often experience divergent, sexually antagonistic selection resulting in an evolutionary conflict between the sexes termed 'intra-locus sexual conflict' (Arnqvist and Rowe 2005; Bonduriansky and Chenoweth 2009). Intra-locus sexual conflict has been well documented in a variety of plant (e.g., Hersh et al. 2015; Zemp et al. 2016; Delph et al.2022) and animal (e.g., Arnqvist and Rowe 2005; Cox and Calsbeek 2009) species and may be an important mechanism in maintaining genetic variation (Kidwell et al. 1977; Rice 1984). Ultimately, intra-locus sexual conflict is thought to be resolved through the evolution of sexual dimorphism (Lande 1980; Rice 1984).

Gene expression can be viewed as a trait in its own right or as a proximate mechanism affecting 'traditional' traits (e.g., morphological, behavioural, or physiological). Over the last ~15 years, there has been a strong interest in investigating sexual dimorphism in gene expression, typically called 'sex-biased gene expression', SBGE (Ellegren and Parsch 2007; Parsch and Ellegren 2013; Mank 2017). SBGE refers to sex differences in the abundance of all transcripts mapping to a given gene. In addition to dimorphism in total expression of a gene, the sexes may differ in other aspects of the transcriptome. Here we examine sexual dimorphism in isoform usage (SDIU) in addition to SBGE.

Differential isoform usage due to alternative splicing of mRNA transcripts results in the production of different isoforms from the expression of a single gene. Differential isoform usage is a widespread phenomenon among multicellular eukaryotic organisms (Ast 2004). For example, in humans, it is estimated that ~95% of multi-exon genes exhibit alternative splicing (Pan et al. 2008). Because the function of proteins and other gene products depends on the specific exons retained in mRNA transcripts after transcription, differential isoform usage has the potential to greatly expand the functional repertoire of the genome. Indeed, in humans, the majority of genes exhibiting alternative splicing are thought to cause functionally significant alterations to gene products (Modrek et al. 2001; Modrek and Lee 2002; Sorek et al. 2004). Thus, differential isoform usage can allow a single gene to alter or take on different functional roles

across developmental stages (e.g., Gibilisco et al. 2016), tissue types (e.g., Barberan-Soler and

Zahler 2008; Telonis-Scott et al. 2009; Revil et al. 2010; Ramani et al. 2011; Tan et al. 2013),

and abiotic environments (Long et al. 2013; Jakšić and Schlötterer 2016). Across the sexes,

differential isoform usage could thus allow the sexes to express their optimal trait values through

the production of different gene products despite sharing a genome.

Widespread sexual dimorphism in isoform usage has been documented in humans

(Trabzuni et al. 2013) and other primates (Blekhman et al. 2010) as well as in birds (Rogers et al.

2020). Several microarray and RNA-seq studies across *Drosophila* species have also investigated

sexual dimorphism in isoform abundance and exon usage. These studies have generally supported

a view that sexual dimorphism in isoform usage is widespread (e.g., McIntyre et al. 2006; Telonis-

Scott et al. 2009; Chang et al. 2011), tissue-specific (Telonis-Scott et al. 2009), and evolutionarily

conserved (Gibilisco et al. 2016).

Here, we further explore SDIU using publicly available RNA-seq data from two tissue

types (head and body tissue) of 18 genotypes of adult *Drosophila melanogaster.* As SBGE and

SDIU can represent different forms of sex-differential gene regulation, we are interested in their

relationship. Resolution or mitigation of intra-locus sexual conflict could require the evolution of

both SBGE and SDIU; on the other hand, SDIU and SBGE may serve as alternative routes to

mitigating sexual conflict so that the evolution of one form of sex-differential regulation obviates

selection for the other. We investigate the co-occurrence of SDIU and SBGE, finding SDIU is

less common among sex-biased than unbiased genes in the body, though the opposite pattern

occurs in the head.

Our other main avenue of investigation focuses on evolutionary genetic properties of

genes exhibiting sex-differential regulation. This follows from several previous studies,

including in flies (Griffin et al. 2013, Fraïsse et al. 2019), seed beetles (Sayadi et al. 2019) and

several bird species (Wright et al. 2019), that have examined quantitative and population

genomic properties in an attempt to infer evolutionary constraint or the form of selection acting

on sex-biased versus unbiased genes. We take a similar approach to examine how two different

forms of expression dimorphism—SBGE and SDIU—are associated with four evolutionary

genetic properties: intrasexual genetic correlation (r_{mf}) in expression, relative levels of

nonsynonymous diversity, Tajima's D, and the direction of selection (*DoS*).

Methods

Obtaining RNA-seq Data and Read Mapping

We obtained publicly available raw RNA-seq data from the DNA Data Bank of Japan
Sequence Read Archive (Accession number: DRA002265). Detailed methods on animal
husbandry, RNA extraction and sequencing can be found in Osada et al. (2017). Briefly, males
from 18 randomly chosen inbred lines of the *Drosophila* Genomic Resource Panel (DGRP;
MacKay et al. 2012; Huang et al. 2014) were crossed to females from a single inbred strain from
a West African population. RNA was extracted from heads and bodies (i.e., whole body minus
heads) of 100 virgin F1 males and females from each cross. RNA from female heads and bodies
was obtained and sequenced from two replicates of each of the 18 DGRP lines while only a
single replicate was available for males. We therefore arbitrarily retained sequence data only for
female samples labelled as 'replicate 1'. Reads were inspected for quality using *FastQC*
(v0.11.9; Andrews 2010) and *MultiQC* (v0.8.3; Ewels et al. 2016). Reads were trimmed for
adapter content using *Trimmomatic* (v0.39; Bolger et al. 2014). RNA-seq reads were then
mapped to the *Drosophila melanogaster* reference genome (release 6.32; Dos Santos et al. 2015)
using *STAR* (v2.7.1a; Dobin et al. 2013) with default settings.

Differential Exon Usage Analysis

To identify a list of genes exhibiting SDIU, we performed differential exon usage
analysis using the R package *JunctionSeq* (v1.17; Hartley and Mullikin 2016). Briefly,
JunctionSeq splits genes into annotated exons and splice junction 'counting bins' and counts the
number of reads that overlap each counting bin within an annotated gene. To test for differential
exon usage between the sexes, a generalized linear model (GLM) with logarithmic link function
was fit at each gene feature (i.e., exon or splice junction) to assess sex differences in the number
of reads that map to a particular gene feature relative to the total number of reads that map to the
gene to which the feature belongs. Specifically, for each gene, *JunctionSeq* fits the model

$$log(y_{ij}) = \mu + W_i + X_j + Z_{ij}$$

where y_{ij} is the normalized count of reads from the ith sex mapping to the jth gene

feature/counting bin. μ is the experiment wide, mean normalized gene expression, W_i reflects

the expected number of reads mapping to the focal gene in sex i, X_j reflects the proportion of

reads mapping to counting bin j, and Z_{ij} is an interaction term that reflects differences between

the sexes in the proportion of reads mapping to counting bin j. A significant effect of the

interaction term was taken as evidence of SDIU. To test the significance of the interaction term,

the above model was compared to a null model that was identical, but which omits the

interaction term.

JunctionSeq can test genes with at least two annotated exons. It applies a multiple-testing

correction method based on the Benjamini and Hochberg procedure (Benjamini and Hochberg

1995) to generate adjusted *P*-values. Genes were considered to show evidence for sexual

dimorphism in isoform usage if the sexes exhibited statistically significant differential usage of a

single exon (or more) within a gene using $p_{adj} < 0.05$.

Ideally, genotype would be included in the model as a random effect but *JunctionSeq*

does not allow this. This is unlikely to be a major issue here. Our dataset is perfectly balanced

with each genotype being represented by only a single sample for each sex for each tissue (head

and body data were analyzed separately). More importantly, our goal is to identify a list of genes

enriched for SDIU for further analysis; our focus is not on the accuracy of *p*-values of individual

genes. As a simple check of the *JunctionSeq* pipeline, we analyzed 10 datasets in which we

switched the male/female labels for a randomly chosen half of the genotypes. In these permuted

data sets, no genes were identified as having SDIU at $p_{adj} < 0.01$ whereas thousands of genes are

found in the true (unpermuted) data under the same criterion.

<u>Sex-biased gene expression</u>

We estimated differential gene expression between the sexes using the R package

DESeq2 (v1.28.1; Love et al. 2014). Briefly, *DESeq2* estimates normalized read counts mapping

to genes on a $\log_2$ scale and uses a GLM framework to test for differential gene expression

between experimental conditions. Before performing differential gene expression analysis, we

performed a 'pre-filtering' procedure to remove genes that had very low numbers of reads

mapping. Pre-filtering RNA-seq data has been shown to substantially increase power in detecting

differential gene expression by reducing the burden of multiple testing but can also result in an

increase in Type I error (i.e., increased false-positive rate) (Bourgon et al. 2010). To balance

these competing issues, we removed any genes that had fewer than an average of 50 reads

mapping from all 36 samples (i.e., 18 male and 18 female genotypes). We then estimated the $\log_2$

fold change ($\log_2$FC) in male to female gene expression for each annotated gene and retained

only genes that were tested for differential exon usage. In downstream analyses, we assigned

each gene to bins based on either their average expression of the gene or based on the $\log_2$FC in

male to female gene expression.

For visual inspection of patterns, we assigned each gene to one of seven bins of

predefined ranges of $\log_2$FC in gene expression: extreme female bias ($\log_2$FC < -5), strong

female bias ($-5 \leq \log_2$FC < -2), moderate female bias ($-2 \leq \log_2$FC $<$ -0.5), unbiased

expression ($0.5 \leq \log_2$FC $<$ 0.5), moderate male bias ($0.5 \leq \log_2$FC $<$ 2), strong male bias (2

$\leq \log_2$FC $<$ 5), and extreme male bias ($\log_2$FC $\geq$ 5). Binning was conducted separately with

respect to expression in heads and bodies. A small number of genes for which the standard error

in $\log_2$FC was high (> 0.5) were excluded. The bins were used for graphical presentation of

results; as described below, our main analyses treat $\log_2$FC as a continuous variable. In heads,

there were few genes with $\log_2$FC > |2| so only the three central bins are shown in the figures.

Distributions of $\log_2$FC for each tissue are shown in Fig. S1.

<u>Sex-averaged expression</u>

For the 18 samples for each sex, we used *featureCounts* (Liao et al. 2014) to generate read

counts for each sample. Using *edgeR* (Robinson et al. 2010), we applied TMM normalization

before obtaining RPKM measures. RPKM is a normalized measure of expression that

incorporates read counts per gene and gene length. Gene lengths were extracted by

featureCounts from the GFF (release 6.32; Dos Santos et al. 2015). For each gene, we averaged

the RPKM value across the 18 samples within a sex, then used the average of male and female

values as a measure of the sex-averaged expression.

<u>Tissue-specificity in expression</u>

We examined how our two measures of expression dimorphism, SDIU and SBGE, are related to

another important aspect of expression, tissue specificity (or, conversely, "expression

pleiotropy"). We used the "tissue-by-stage" estimates of tissue specificity calculated by Fraïsse

et al. (2019) applying the method of Yanai et al. (2005) to another expression dataset

(modENCODE Consortium 2010). Their measure quantifies tissue specificity jointly across four

larval tissues (carcass, digestive system, central nervous system, imaginal discs), two pupal

tissues (central nervous system, fat tissue), and three adult tissues (carcass, digestive system,

heads).

<u>Intersexual genetic correlation r_{mf}</u>

Following Fraïsse et al. (2019), we calculated the intersexual genetic correlation in gene

expression from the data of Huang et al. (2015) who measured whole body gene expression of

185 genotypes from the Drosophila Reference Genome Panel. Using *MCMCglmm* (Hadfield

2010), for each gene we fit the model $z_{ijk} = \mu + \phi_{jk} + e_{ijk}$ where z_{ijk} is the normalized

expression of the ith replicate of sex j of genotype k. μ is the mean expression. ϕ_{jk} is the

random genetic effect of sex j of genotype k and follows a bivariate Gaussian distribution

characterized by sex-specific variances and the intersexual covariance. e_{ijk} is the residual.

Models were fit using uninformative priors and run for a million iterations after a burn-in period

of 10,000 iterations with a thinning interval of 1000. From the resulting MCMC posterior of sex-

specific variances and intersexual covariances, a posterior of r_{mf} was created and the mean value

used as the estimate of r_{mf}. We only considered r_{mf} values for genes for which both sexes had

genetic variances > 0.001 and statistically significant ($p < 0.05$) in at least one sex based on

results of Huang et al. (2015; their Table S5). In the Supplementary Material we show another

version of the main figure using only genes for which the genetic variance was statistically

significant in both sexes; the patterns are very similar.

 Population Genomic Data

We obtained whole genome sequences for 205 haploid genotypes of the DPGP3

Drosophila population which were originally collected from a wild population in Siavonga,

Zambia (Lack et al. 2015, 2016). FASTA files were obtained from the *Drosophila* Genome

Nexus (DGN; Lack et al. 2015, 2016) website (https://www.johnpool.net/genomes.html) and

were filtered to remove sites with evidence of high relatedness and recent admixture using Perl

scripts provided on the DGN website. Following Sohail et al. (2017), we removed five genotypes

that had extremely high or low numbers of variants. Additionally, we removed a single genotype

(ZI28) for which the length of the X chromosome in the FASTA file differed from the other 204

genotypes. We then created a VCF file using the remaining 199 sequences with *SNP-sites*

(v2.5.1; Page et al. 2016). Coordinates for the VCF file were converted to release 6 of the

Drosophila reference genome using *UCSC liftOver* (Kent et al. 2002).

Estimating nucleotide diversity and Tajima's *D*

We obtained a GFF file for the *D. melanogaster* genome from Ensembl

(https://useast.ensembl.org/Drosophila_melanogaster/) and scored each coding site in the

genome as 4-fold (i.e., synonymous sites) or 0-fold (i.e., non-synonymous sites) degenerate

across the *Drosophila* genome using a Python script (script available at

https://github.com/tvkent/Degeneracy). We then parsed our VCF containing variant and invariant

sites in the DGRP into two files that contained either only 4-fold sites or only 0-fold degenerate

sites. After excluding sites with data from fewer than 100 lines and those with more than two

alleles, we measured per site nucleotide diversity $\pi = 2p(1-p)$, where p is the estimated

frequency of the reference base. Nonsynonymous diversity π_N within a gene was measured as

the mean nucleotide diversity at 0-fold degenerate sites within that gene and synonymous

diversity π_S within a gene as the mean nucleotide diversity at 4-fold degenerate sites within each

gene. Our analysis focuses on a measure of *relative* nonsynonymous diversity, $\tilde{\pi}_{rN} \equiv$

$\pi_N/(\pi_N + \pi_S)$. We use this rather than the more familiar ratio of π_N/π_S because the latter has

undesirable statistical properties, being highly sensitive to sampling variation in π_S and is

undefined if $\pi_S = 0$. Our measure $\tilde{\pi}_{rN}$ is bounded between 0 and 1 and will be similar to π_N/π_S

when $\pi_S \gg \pi_N$, as is expected for typical genes.

Tajima's *D* (Tajima 1989) was estimated using 4-fold degenerate sites, accounting for the

sample size and number of 4-fold sites (both invariant and variant) considered for each gene

(Walsh and Lynch 2020, p. 303); sites with data from a fewer than 100 lines were not

considered. When sites varied in sample size, we individually down sampled variant sites to the

level of the site with the lowest sample size, then calculated Tajima's *D*. Downsampling was

performed 100 times per gene and the median value of Tajima's *D* was used.

<u>Direction of Selection (*DoS*)</u>

DoS (Stoletzki and Eyre-Walker 2011) is a metric that compares relative nonsynonymous

divergence to relative nonsynonymous polymorphism; in both cases, it is relative to the sum of

synonymous and nonsynonymous changes. We used the *DoS* values reported by Fraïsse et al.

(2019). They calculated divergence with respect to *D. simulans*. They calculated polymorphism

from 197 haploid genomes of *D. melanogaster* genomes of the DPGP3. They excluded rare

variants (minor allele frequency < 5%) to reduce the influence of deleterious variants on *DoS*.

<u>Statistical Analysis</u>

For each of four evolutionary genetic properties of interest—r_{mf}, $\tilde{\pi}_{rN}$, Tajima's *D*, and *DoS*—we

used linear models to examine whether among-gene variation in the metric was predicted by four

gene expression characteristics: (i) SDIU status; (ii) SBGE measured as $\log_2$FC; (iii) sex-

averaged expression and (iv) tissue specificity. Because the effect of SBGE could be similar for

both male and female bias, we included a quadratic term for SBGE. We also included

interactions terms between sex-averaged expression and both the linear and quadratic SBGE

terms. Genes with extreme sex-bias ($|\log_2$FC$| > 5$) were excluded from these analyses because

visual inspection revealed that these genes had distinctly different values for several of the

evolutionary genetic properties of interest (see Results). For the results presented in the main

text, SDIU status was coded as a numerical value: 2 (strong evidence of SDIU: $p_{adj} < 0.01$), 1

(modest evidence: $0.2 < p_{adj} \leq 0.01$), or 0 (poor evidence: $p_{adj} \geq 0.2$). Model fits are better when

SDIU status is coded as three rather than two values but very similar results are obtained if SDIU

status is coded as "significant" ($p_{adj} < 0.05$) or "non-significant" ($p_{adj} \geq 0.05$); see Supplementary

Material.

All models included an indicator variable of whether the gene was autosomal or X-

linked. In addition, the models for each of the three population genomic metrics—$\tilde{\pi}_{rN}$, Tajima's

D, and DoS—included two other terms: log(recombination rate) and log(transcript length).

Recombination is expected to affect population genetic patterns by mediating the effects of

linked selection. Transcript length is also expected to affect linked selection as a given site in a

long gene is expected to have more closely linked targets of selection than a site in a short gene.

Recombination values came from Fraïsse et al. (2019), who obtained recombination measures

from Chan et al. (2012).

The statistical support for model terms was assessed in two different ways. We obtained

permutation-based p-values using the *lmp* function from the *lmperm* package (Wheeler and

Torchiano, 2016). We used perm = "Prob", Ca = 0.001, maxIter = 10^6 and seq = FALSE for

unique rather than sequential sum of squares. Second, we used the *Boot* function from the *boot*

package (Canty and Ripley 2020) on each linear model, followed by *confint* to obtain 95%

bootstrap confidence intervals of model terms. We specified method = "case" to sample over

cases, which is considered a conservative approach. In almost all cases, the two methods gave

consistent results, i.e., when terms were significant based on permutation p-values < 0.05, the

bootstrap confidence intervals for the estimate did not overlap zero.

Results and Discussion

Prevalence of Sexually Dimorphic Exon Usage

We find evidence of extensive sexual dimorphism in isoform usage (SDIU or 'isoform

dimorphism') among multi-exon genes expressed in both head and body tissue of adult

Drosophila melanogaster. Table 1 summarizes the results of our differential exon usage analysis.

Of multi-exon genes tested, 5375 and 790 genes expressed in adult body and head tissue exhibit

statistically significant SDIU respectively which accounts for 57.5% and 10.2% of all genes

tested in bodies and heads respectively. Considering the union of genes expressed in either

tissue, we detected SDIU in at least one of the tissues for 48% of genes. These data accord well

with several previous studies that have found evidence of widespread SDIU both in this species

(e.g., McIntyre et al. 2006; Telonis-Scott et al. 2009; Chang et al. 2011) as well as other species

of *Drosophila* (Gibilisco et al. 2016). McIntyre et al. (2006) reported that 68% of all multi-

transcript genes expressed in whole body tissue exhibited significant sex-bias in the expression

of one or more alternative isoforms in a microarray study using *D. melanogaster*. Similarly, a

microarray study in this species by Telonis-Scott et al. (2009) found that ~58% of multi-exon

genes tested in whole body tissue exhibited SDIU.

Table 1: Summary of the differential exon usage analysis between the sexes.

	Genes Expressed	Genes Tested	Genes with Sexually Dimorphic Isoform Usage (% of total genes tested)
Head	8879	7772	790 (10.2%)
Body	11221	9346	5375 (57.5%)
[1]Head or Body	11769	9788	5386 (47.8%)

[1]These data refer to the union of genes expressed, tested and exhibiting SDIU found in either
head, body or both.

Heads and bodies differ substantially in the frequency of SDIU. Although the majority of

genes exhibiting SDIU in head tissue also exhibited SDIU in the body, there are many genes

exhibiting SDIU in body tissue where we did not detect dimorphism in head tissue. 7330 genes

were tested in both tissues. 4187 are SDIU in the body but only 615 (16%) of these are also

SDIU in the head. Of the 7330, a total of 779 genes are SDIU in the head and 87% of these are

SDIU in the body. These results comport well with previously published data. For example,

Telonis-Scott et al. (2009) reported reduced levels of SDIU in genes when expressed in head

tissue compared to carcass (i.e., head- and gonad-less bodies) and gonadal tissue.

 <u>Coincidence of two forms of sexually dimorphism in expression</u>

The majority of expression studies have examined variation in total expression for a gene.

Studies of sex-biased gene expression (SBGE) have shown repeatedly that SBGE is highly tissue

specific (e.g., Parisi et al. 2004; Yang et al. 2006; Dutoit et al. 2018); both our data and other

studies (e.g., Telonis-Scott et al. 2009; Chang et al. 2011) indicate that the same is true for sex

differences in isoform usage. Because both of these forms of sexual dimorphism are thought to

be related creating phenotypic dimorphism and resolving intra-locus sexual conflict, we are

interested in understanding the relationship between SBGE and SDIU.

Figure 1. Proportion of genes exhibiting sexual dimorphic isoform usage (SDIU) (**A**) among all genes tested; and (**B**) stratified with respect to sex-biased gene expression (SBGE). Points in mauve and orange represent reflect characterizations based on expression in body and head samples, respectively. Only three levels of SBGE are shown with respect to head because very few genes fall into the other levels. Error bars represent bootstrap 95% confidence intervals. In some cases, error bars are too small to be visible.

SBGE is widespread both in heads and bodies, though there is much greater variation

among genes in SBGE in body tissue relative to head tissue (e.g., Innocenti and Morrow 2010;

Baker et al. 2011; Dutoit et al. 2018). To visualize the relationship between SDIU and SBGE, we

divided genes into bins representing different levels of sex-bias using defined ranges of $\log_2$FC

in gene expression (see Methods). Within each bin, we calculated the proportion of genes

exhibiting SDIU. With respect to expression in the body, SDIU is more common amongst

unbiased genes than sex-biased genes, except for those with extreme sex-bias (Figure 1B).

These fly body results are in line with Rogers et al. (2020) who reported non-significant trends

towards a negative relationship between SDIU and SBGE in gonadal tissue of three bird species.

The reduction in isoform dimorphism at highly sex-biased genes expressed in the body

could be because SBGE and SDIU provide alternative routes to the resolution of sexual conflict

and thus the evolution of one reduces the selective pressure for the evolution of the other. On the

other hand, the pattern goes in the opposite direction with respect to expression in the head;

SDIU is less common among unbiased genes than sex-biased genes (at least those that are

female-biased). One hypothesis for this latter pattern is that the resolution of intra-locus conflict

may often require both forms of expression dimorphism whereas those genes experiencing

sexually concordant selection will evolve neither, creating a positive association between SDIU

and SBGE.

An alternative, more mundane, explanation is that the positive association of SBGE and

SDIU in heads exists because of how we quantify gene/isoform expression (i.e., counting reads

mapping to genes); we may be biased towards finding a positive association between SBGE and

SIDU because we are observing the same phenomenon in different ways. For example, if both

males and females express a single isoform of a gene at the same level but females also express

an additional isoform that is not expressed by males, we would detect both SBGE and SDIU at

this gene. However, this methodological bias must not be very strong as there are many unbiased

genes that exhibit SDIU (670 in head and 1007 in body). Moreover, the association between

SBGE and SDIU is negative in the body.

It is worth noting that average gene length varies across different levels of SBGE (Fig.

S2). SDIU genes are, on average, substantially longer than non-SDIU genes with respect to total

exon length (Fig. S2). To some extent this association is expected to be at least partially due the

greater power to detect SDIU in longer genes because there will be more reads for a given level

of expression of a long gene than a short one. However, the association of SDIU and SBGE does

not appear to be an artifact driven by gene length. The proportion of SDIU genes among levels

of SBGE varies in a manner that is not consistent with variation in average gene length among

levels of SBGE (compare Fig. 1 and Fig. S2). Moreover, the patterns in the proportion of SDIU

genes among levels of SBGE remain similar even when genes are stratified by gene length (Fig.

S3). Nonetheless, these associations of gene length with both SDIU and SBGE are noteworthy.

Gene length is expected to affect the amount of linked selection and thereby influence population

genetic properties (Larracuente et al. 2008; Fraïsse et al. 2019) that we examine later.

Figure 2. Proportion of X-linked genes with respect to SDIU status. (**A**) Results for body and
head shown irrespective of SBGE. Results shown stratified by SBGE for (**B**) body and (**C**) head.
Points in purple and green represent non-SDIU and SDIU genes, respectively; points in black are
all genes irrespective of SDIU status (including genes that were not tested for SDIU). Error bars
are bootstrap 95% confidence intervals. In some cases, error bars are too small to be visible.

Given that the X chromosome spends a disproportionate amount of its time in females,

and because of its hemizygous expression in males, there has been considerable interest in

whether it is enriched for genes that experience sexually antagonistic selection contributing to

intralocus conflict or contributing to sexual dimorphism (Rice 1987; Frank and Patten 2020).

This has motivated interest in examining the location of genes exhibiting expression

dimorphism. In the body data, there is no evidence of an enrichment of genes exhibiting SDIU

on the X chromosome (Figure 2). In contrast, SBGE is strongly associated with the X

chromosome, with female-biased genes being significantly overrepresented on the X

chromosome relative to unbiased genes whereas male biased genes are significantly

underrepresented on the X chromosome (Figure 2). This observation of a "demasculinization" of

the X chromosome has been well documented in previous studies of SBGE in *D. melanogaster*

as well as several other species (Parisi et al. 2003; Ranz et al. 2003; Khil et al. 2004; Reinke et

al. 2004; Sturgill et al. 2007; Meisel et al. 2012; Albritton et al. 2014).

These patterns are quite different for the head data. In contrast to the body data, there are

a significantly higher proportion of X-linked genes exhibiting SDIU in the head relative to non-

SDIU genes (Figure 2). With respect to SBGE, both male- and female-biased genes are

overrepresented on the X relative to unbiased genes, with the strongest overrepresentation among

male-biased genes (Figure 2; see Meisel et al. (2012)). The over-representation of SDIU genes

on the X chromosome occurs primarily among unbiased genes rather than sex-biased genes,

indicating that this pattern is not driven by a positive association between SDIU and sex-biased

expression.

<u>Association of SDIU and SBGE with sex-averaged expression</u>

An important aspect of expression that we have not yet discussed is the overall level of

expression, measured here as the average RPKM across males and females. Both SDIU and

SBGE are associated with the sex-averaged expression level (Fig. 3). Expression levels are

higher for genes with versus without SDIU. This may simply reflect a statistical artifact as there

is greater power to detect SDIU in genes that have higher expression. Regardless of the

underlying reason, this association will be an important consideration when we later consider

evolutionary characteristics of genes with SDIU. In bodies, sex-averaged expression is greater

for unbiased than sex-biased genes, particularly the more strongly male-biased genes. In heads,

female-biased genes have greater average expression than unbiased genes, which in turn have

greater expression than male-biased genes.

Figure 3. Mean sex-averaged expression with respect to SDIU status. (**A**) Results for body and head shown irrespective of SBGE. Results shown stratified by SBGE for (**B**) body and (**C**) head. Points in purple and green represent non-SDIU and SDIU genes, respectively; points in black are all genes irrespective of SDIU status (including genes that were not tested for SDIU). Error bars are bootstrap 95% confidence intervals. In some cases, error bars are too small to be visible.

<u>Association of SDIU and SBGE with tissue specificity</u>

The final aspect of expression we consider is tissue specificity. We use the measure of tissue specificity calculated by Fraïsse et al. (2019) based an independent dataset with expression information from multiple tissues from several developmental stages (modENCODE Consortium 2010). Tissue specificity is higher for male-biased than female-biased or unbiased genes (Fig. 4). Tissue specificity is lower for SDIU than non-SDIU genes. In bodies, this effect is driven largely by the difference between SDIU and non-SDIU among moderately and strongly male-biased genes. In heads, this effect comes largely from female-biased and unbiased genes.

Previous studies in *Drosophila* (Assis et al. 2012; Fraïsse et al. 2019) and other animals (Mank et al. 2008; Dean and Mank 2016) have also found that sex-biased genes have higher

tissue specificity than unbiased genes; this has been interpreted as sexual dimorphism evolving
more readily when there is less pleiotropic constraint on expression. Under this premise, one
might expect the same for SDIU. However, in the body data, SBGE is associated with greater
tissue specificity (often interpreted as reduced pleiotropic constraint) whereas SDIU is associated
with lower tissue specificity. Similarly, Rogers et al. (2020) found that genes exhibiting SDIU
exhibited lower tissue specificity in gene expression than non-SDIU genes in birds.

Figure 4. Mean tissue-specificity with respect to SDIU status. (**A**) Results for body and head
shown irrespective of SBGE. Results shown stratified by SBGE for (**B**) body and (**C**) head.
Points in purple and green represent non-SDIU and SDIU genes, respectively; points in black are
all genes irrespective of SDIU status (including genes that were not tested for SDIU). Error bars
are bootstrap 95% confidence intervals. In some cases, error bars are too small to be visible.

<u>Evolutionary Properties associated with SDIU and SBGE</u>

We next examine how each of four evolutionary properties—r_{mf}, $\tilde{\pi}_{rN}$, Tajima's D, and DoS—
varies across genes in relation to both SDIU and SBGE. We do this in two ways. First, we
examine simple associations of each evolutionary property with SDIU and SBGE, which are

straightforward to represent graphically. This approach ignores other factors—including sex-
averaged expression and tissue specificity—that could be important sources of variation in the
focal evolutionary property. Such factors could obscure the relationship of the focal evolutionary
property with SDIU or SBGE. Consequently, we also use linear models in which we examine
variation in an evolutionary property as a function of SDIU and SBGE as well as other relevant
covariates. These other covariates include sex-averaged expression and tissue specificity as well
as genomic location (X or autosome). For the population genetic properties—$\tilde{\pi}_{rN}$, Tajima's D,
and DoS—we also included two other covariates that could affect the strength of linked
selection: recombination rate and gene length. For simplicity, we only perform these analyses
with respect to expression characteristics in body samples as this is where sexual dimorphism is
greatest. As we shown graphically, genes with extreme sex bias (particularly extreme male-bias)
often differ dramatically from other genes. We exclude these genes from the linear models.

Figure 5. Four evolutionary genetic properties in relation to expression properties in the body. (**A, C, E, G**) Results shown irrespective of SBGE. (**B, D, F, H**) Results shown stratified by SBGE. Points in purple and green represent non-SDIU and SDIU genes, respectively; points in black are all genes irrespective of SDIU status (including genes that were not tested for SDIU). Error bars are bootstrap 95% confidence intervals. In some cases, error bars are too small to be visible.

Intersexual genetic correlation, r_{mf}

The intersexual genetic correlation, r_{mf}, plays an important role in determining the extent to which a trait can evolve independently in one sex from the other. We examine variation in the intersexual genetic correlation in gene expression from the quantitative genetic study of DGRP lines by Huang et al. (2015). Genes with SDIU have significantly lower r_{mf} values, on average, than genes without SDIU (Fig. 5A; 95% CIs do not overlap). There is a striking relationship between r_{mf} and SBGE (Fig. 5B), with female-biased genes having lower r_{mf} than unbiased genes. Moderately and strongly male-biased genes have the highest average r_{mf} values. In contrast, genes with extreme male bias have very low r_{mf} values on average. When we analyze variation in r_{mf} using a linear model to help control for correlated variables, the patterns described above are recovered (Table 2): r_{mf} is reduced for genes with SDIU; r_{mf} increases going from female-biased to unbiased to male-biased genes. The model also indicates that r_{mf} is higher for genes with higher tissue specificity.

Table 2. Linear model results for the intersexual genetic correlation in fitness, r_{mf}

Term	Estimate [95% bootstrap CI]	*p*-value
SDIU	-0.034 [-0.047, -0.022]	$< 10^{-6}$
SBGE	0.055 [0.037, 0.072]	$< 10^{-6}$
SBGE^2	-0.016 [-0.024, -0.009]	0.0005
Avg Expression	0.069 [0.057, 0.082]	$< 10^{-6}$
SBGE × Avg Expression	0.002 [-0.004, 0.007]	0.53
SBGE^2 × Avg Expression	0.0005 [-0.0018, 0.0028]	1.00
Tissue specificity	0.349 [0.289, 0.406]	$< 10^{-6}$
X-linkage	-0.013 [-0.027, 0.001]	0.069

Analysis is based on $n = 3383$ genes with complete information for all model terms. Genes with extreme sex-bias ($\log_2 FC > |5|$) are excluded from analysis. *p*-values are permutation based.

The relationship between r_{mf} and SBGE was first investigated by Griffin et al. (2013) using microarray expression data from a smaller set of genotypes (40 vs 185 here) also from the DGRP. They report a negative relationship between r_{mf} and sex-bias but they did not distinguish between male- and female-biased genes. Similarly, Fraïsse et al. (2019) calculated r_{mf} using the same RNAseq data set we used (Huang et al. 2015). They also noted a negative relationship between r_{mf} and sex-bias but did not distinguish between male- and female-bias. In their study of *D. serrata*, Allen et al. (2018) noted that r_{mf} was on average lower for female- than male-biased genes, consistent with what we observe. Allen et al. (2018) also noted that the correlation between r_{mf} and the magnitude of sex-bias was weaker for female- than male-biased genes; we find the opposite (females: ρ = -0.20 [-0.24, -0.16]; males: ρ = -0.13 [-0.17, -0.09]; calculated excluding genes in the unbiased and extreme sex-biased categories).

Relative level of nonsynonymous diversity, $\tilde{\pi}_{rN}$

A reduction in nucleotide diversity for nonsynonymous relative to synonymous variants is

typically interpreted as indicative of purifying selection on amino acid sequence variation, with

lower nonsynonymous diversity levels indicating stronger selection. A simple inspection of

relative levels of nonsynonymous diversity $\tilde{\pi}_{rN}$ reveals no clear difference between genes with

versus without SDIU (Fig. 5C, D). Genes with extreme sex-bias have elevated levels of $\tilde{\pi}_{rN}$,

consistent with the idea that such genes experience weaker net purifying selection because they

are primarily selected in only one sex (Pröschel et al. 2006; Dapper and Wade 2016, 2020).

Excluding the extreme sex-bias genes, we applied a linear model to variation in $\tilde{\pi}_{rN}$

(Table 3). This model indicated an interaction of SBGE with sex-averaged expression level. On

average, $\tilde{\pi}_{rN}$ is reduced in genes with high versus low average expression. This is consistent with

results from a variety of taxa, which is usually interpreted as indicating stronger purifying

selection on highly expressed genes (e.g., Carneiro et al. 2012; Paape et al. 2012; Williamson et

al. 2014; Ho et al. 2019). Further, our model indicates that $\tilde{\pi}_{rN}$ declines from female-biased to

unbiased to male-biased genes for highly and moderately expressed genes but does the opposite

for female-biased genes. (The interaction term indicates that $\tilde{\pi}_{rN}$ varies more with respect to sex-

averaged expression for female-biased genes than male-biased genes.)

The model also indicates that $\tilde{\pi}_{rN}$ increases with tissue specificity, consistent with the

idea of weaker purifying selection on genes with less pleiotropy (e.g., Osada 2007; Paape et al.

2013). $\tilde{\pi}_{rN}$ declines with recombination, as expected from recombination reducing selective

interference. However, $\tilde{\pi}_{rN}$ is elevated for X-linked genes and shorter genes. Neither of these

results match expectation as selection is thought to be more effective on the X than autosomes in

flies (Campos et al. 2018) and longer genes are expected to experience more selective

interference.

Table 3. Linear model results for the relative level of nonsynonymous diversity, $\tilde{\pi}_{rN}$.

Term	Estimate [95% bootstrap CI]	*p*-value
SDIU	0.0021 [0.0020, 0.0061]	1.00
SBGE	-0.011 [-0.0179, 0.0047]	$< 10^{-6}$
SBGE^2	0.0015 [-0.0011, 0.0039]	1.00
Avg Expression	-0.0013 [-0.0173, 0.0089]	$< 10^{-6}$
SBGE × Avg Expression	0.0031 [0.0007, 0.0054]	0.00029
SBGE^2 × Avg Expression	-0.000043 [-0.000935, 0.000962]	0.42
Tissue specificity	0.045 [0.024, 0.067]	$< 10^{-6}$
X-linkage	0.0069 [0.0021, 0.0119]	0.011
log(recombination)	-0.021 [-0.023, -0.018]	$< 10^{-6}$
log(gene length)	-0.025 [-0.032, -0.019]	$< 10^{-6}$

Analysis is based on *n* = 4629 genes with complete information for all model terms. Genes with extreme sex-bias (log$_2$FC > |5|) are excluded from analysis. *p*-values are permutation based.

Tajima's D

Tajima's *D* is a summary statistic that reflects information about the shape of the underlying genealogy, which can be affected by the strength and form of selection at linked sites. As such, variation in Tajima's *D* has been used to look for indirect evidence that genes with different expression properties are selected differently (Sayadi et al. 2019; Wright et al. 2019). As shown in Fig. 5E, genes with SDIU have lower values of Tajima's *D* (measured using 4-fold degenerate sites). Though there is no visually obvious pattern with respect to SBGE (Fig. 5F), the linear model indicates strongly significant effects of SBGE involving interactions with sex-averaged expression (Table 4). Tajima's *D* is reduced in genes with high sex-averaged expression.

Roughly speaking, the model indicates that Tajima's D is reduced for sex-biased genes relative

to unbiased biased genes, at least for genes with intermediate levels of sex-averaged expression.

For genes with low sex-averaged expression, the highest Tajima's D values are inferred to occur

for female-biased genes; for genes with high sex-averaged expression, the highest Tajima's D

values occur for slightly male-biased genes.

To the extent variation in Tajima's D is due to variation in selection, it could simply

reflect variation in the strength of purifying selection. We fit another version of the model

including $\tilde{\pi}_{rN}$ as a covariate in a crude attempt to control for the strength of purifying selection.

There was a strongly significant negative effect of $\tilde{\pi}_{rN}$, which could be interpreted as weaker

purifying selection causing more negative values of Tajima's D; a pattern expected by theory

over some range of selection strengths (Charlesworth et al. 1995; Cvijović et al. 2018). However,

the patterns of Tajima's D with respect to SDIU and SBGE remained qualitatively unchanged,

e.g., Tajima's D is lower for genes with SDIU. This leaves open the possibility that other

differences in selection contribute to these patterns. For example, genes with SDIU could be less

likely to experience balancing selection or more likely to experience selective sweeps than genes

without SDIU.

Table 4. Linear model results for Tajima's *D*.

Term	Estimate [95% bootstrap CI]	*p*-value
SDIU	-0.0037 [-0.0065, -0.0009]	0.0075
SBGE	-0.0066 [-0.0097, -0.0035]	$< 10^{-6}$
SBGE^2	-0.0018 [-0.0031, -0.0006]	$< 10^{-6}$
Avg Expression	-0.0057 [-0.0087, -0.0028]	$< 10^{-6}$
SBGE × Avg Expression	0.0022 [0.0020, 0.0033]	0.00022
SBGE^2 × Avg Expression	0.00067 [0.00024, 0.00110]	$< 10^{-6}$
Tissue specificity	0.037 [0.023, 0.052]	$< 10^{-6}$
X-linkage	-0.021 [-0.024, -0.017]	$< 10^{-6}$
log(recombination)	-0.0023 [-0.0036, -0.0009]	0.00031
log(gene length)	0.0067 [0.0025, 0.0110]	0.00060

Analysis is based on *n* = 4754 genes with complete information for all model terms. Genes with
extreme sex-bias (log$_2$FC > |5|) are excluded from analysis. *p*-values are permutation based.

Direction of Selection DoS

An excess of divergence to polymorphism for nonsynonymous relative to synonymous variants

can be indicative of adaptive evolution (McDonald and Kreitman 1991). Based on this idea,

Stoletzki and Eyre-Walker (2011) devised the *DoS* metric for which positive values are

indicative of adaptive evolution, with more strongly positive values reflecting more frequent

adaptive evolution. *DoS* is higher for genes with SDIU than those without (Fig. 5G; Table 5).

DoS is substantially elevated for genes with extreme male-bias (Fig. 5H). Excluding the genes

with extreme sex-bias, the linear model detects no significant effect of SBGE on *DoS*. There is
also a significant positive effect of sex-averaged expression, qualitatively consistent with the
findings of Huang (2022; but see Fraïsse et al. 2019). Consistent with other studies, the model
indicates more adaptive evolution for X-linked genes, genes with higher recombination rates,
and shorter genes (e.g., Campos et al. 2018, Fraïsse et al. 2019).

Table 5. Linear model results for *DoS*.

Term	Estimate [95% bootstrap CI]	*P*-value
SDIU	0.012 [0.006, 0.018]	0.00050
SBGE	-0.0058 [-0.0135, 0.0020]	0.13
SBGE^2	0.0012 [-0.0019, 0.0042]	1.00
Avg Expression	0.013 [0.006, 0.019]	$< 10^{-6}$
SBGE × Avg Expression	0.0013 [-0.0014, 0.0039]	0.30
SBGE^2 × Avg Expression	-0.00020 [-0.00123, 0.00088]	1.00
Tissue specificity	0.0076 [-0.0222, 0.0373]	1.00
X-linkage	0.027 [0.020, 0.035]	$< 10^{-6}$
log(recombination)	0.022 [0.019, .025]	$< 10^{-6}$
log(gene length)	-0.026 [-0.035, -0.017]	$< 10^{-6}$

Analysis is based on *n* = 4765 genes with complete information for all model terms. Genes with
extreme sex-bias (log$_2$FC > |5|) are excluded from analysis. *p*-values are permutation based.

A potential concern in interpretating *DoS* is that large positive values could be primarily

due to low levels of nonsynonymous polymorphism, with divergence playing little role.

However, divergence and polymorphism components each contribute substantially to variation in

DoS in this dataset. Moreover, if we apply a linear model to just the divergence component of

DoS (i.e, $D_n/(D_n + D_s)$), we again detect a positive effect of SDIU (Table S1). In addition, the

divergence component of *DoS* declines from female-biased to unbiased to male-biased genes,

except for those genes with high sex-averaged expression (i.e., a significant negative SBGE

effect in combination with a significant positive interaction between SBGE and sex-averaged
expression).

Conclusions

SDIU and SBGE are two different ways to differentially regulate genes across the sexes and thus
could be mechanisms by which sexual conflict is resolved. Both SBGE and SDIU are
widespread in *Drosophila melanogaster*, particularly among genes expressed in body tissue.
Genes with SDIU have lower values of the intersexual genetic correlation in expression r_{mf}. This
indicates that patterns of standing genetic variation are more conducive to evolving greater sex
differences in a gene's total expression (i.e., stronger SBGE) for genes with SDIU versus
without, holding the amount of genetic variation constant. Thus, one might predict genes with
SDIU to have higher levels of SBGE, which stands at odds with the observed negative
association between these two forms of expression dimorphism in the body (Fig. 1). This could
be because when one of these mechanisms evolves, it may negate the selective pressure to
evolve the other (i.e., SDIU genes are less likely to be under antagonistic selection for total
expression than non-SDIU genes). A second possibility is that the genetic variants that allow for
a greater degree of independence in total expression—those that lower r_{mf}—have other
constraining pleiotropic effects (e.g., perhaps such variants are more likely to be due *trans*-
regulatory factors that affect many other genes) or such variants only affect the expression of
some isoforms and those are not the targets of antagonistic selection for expression level.

SDIU and SBGE are differently associated with the population genetic metrics we
examined, suggesting that selection acts differently on these two dimensions of expression
dimorphism. For example, one interpretation of the $\tilde{\pi}_{rN}$ analysis is that purifying selection tends
to get stronger going from female-biased to unbiased to male-biased genes (excluding genes with
extreme sex-bias). In contrast, purifying selection is similar on genes with versus without SDIU.
An interpretation of the *DoS* analysis is that SDIU genes experience more adaptive evolution
than genes without SDIU. Perhaps when sexes use different isoforms, it is easier for adaptation
to occur in each sex because there is less pleiotropic constraint imposed by the other sex. In
contrast, there is no clear relationship between adaptive evolution and SBGE (when genes with
extreme sex-bias are excluded).

Such interpretations should be regarded with caution for several reasons. First, population

genetic metrics can be influenced by a multitude of factors and the interpretations above are not

necessarily correct. Second, an unmeasured correlated variable could be responsible for the

reported associations, though we incorporated several plausibly important variables into our

analyses to reduce this possibility. Third, the comparison of SDIU and SBGE is complicated by

the nature of how they are defined and measured here. SBGE has both a direction (i.e., negative

to positive representing female- to male-biased) and a magnitude. SDIU, as measured here,

simply indicates whether the sexes differ in their distribution of isoform usage; there is no

direction. Magnitude only affects SDIU status via the statistical power to detect the difference,

which also depends on other factors (e.g., expression level, sequencing depth). (Quantifying a

"magnitude" for SDIU is complicated by the multitude of ways that isoform distributions can

differ; but see Rogers et al. 2020.)

Much of our analysis has focused on expression properties from whole body samples

because our main interest here is in dimorphism and whole bodies are much more dimorphic

than heads. Many related studies have also focussed on expression from whole body data (e.g.,

Innocenti & Morrow 2010, Griffiths et al. 2013, Cheng & Kirkpatrick 2016, Fraïsse et al. 2019).

On a practical level, only a subset of genes are expressed at a sufficiently high level to warrant

characterization of SBGE and SDIU in multiple tissue types. Nonetheless, expression properties

of a gene can vary across tissues, life stages, and environmental conditions. Within the first

section of our results, we highlighted differences between head and body including the

frequencies of SDIU and SBGE as well as their co-occurrence (Table 1, Fig. 1) and over/under-

representation on X (Fig. 2). Thus, using only one type of sample (e.g., whole body) is, at best, a

partial representation of a gene's expression characteristics.

Furthermore, whole body samples represent a collection of tissue types and expression

profiles differ across tissues. Thus, expression differences between sexes can be due to either

expression changes within tissues between sexes or the relative contribution of each tissue within

each sex (Montgomery & Mank 2016). We take the coarser perspective that expression in whole

body samples represents a set of phenotypes in which we can examine dimorphism, while

acknowledging the limitations on interpreting the biological source of dimorphism. Despite the

caveats of using whole body expression, these measures of SBGE and SDIU are associated with

several key evolutionary properties. These patterns are intriguing though a bigger challenge will

be to gain a clearer understanding of why they occur.

Acknowledgements: We wish to thank Tyler Kent for providing a script to identify N-fold

degeneracy and Julia Kreiner, and James Santangelo for bioinformatics advice. Funding was

provided in the form of an Ontario Graduate Scholarship to AS and an NSERC discovery grant

to AFA.

Literature Cited

Literature Cited

Albritton, S. E., A. L. Kranz, P. Rao, M. Kramer, C. Dieterich, and S. Ercan. 2014. Sex-biased
gene expression and evolution of the X chromosome in nematodes. Genetics 197:865–883.

Allen, S. L., R. Bonduriansky, and S. F. Chenoweth. 2018. Genetic constraints on
microevolutionary divergence of sex-biased gene expression. Phil. Trans. R. Soc. B 373:
20170427.

Andrews, S. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data.

Arnqvist, G., and L. Rowe. 2005. Sexual Conflict. Princeton University Press, Princeton NJ.

Assis, R., Q. Zhou, and D. Bachtrog. 2012. Sex-biased transcriptome evolution in
Drosophila. Genome Biol Evol., 4:1189-1200.

Ast, G. 2004. How did alternative splicing evolve? Nat. Rev. Genet. 5:773–782.

Baker, D. A., T. Nolan, B. Fischer, A. Pinder, A. Crisanti, and S. Russell. 2011. A
comprehensive gene expression atlas of sex- and tissue-specificity in the malaria vector,
Anopheles gambiae. BMC Genomics 12:296.

Barberan-Soler, S., and A. M. Zahler. 2008. Alternative splicing regulation during *C. elegans*
development: Splicing factors as regulated targets. PLoS Genet. 4:e1000001.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the False Discovery Rate : A Practical and
Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B 57:289–300.

Blekhman, R., J. C. Marioni, P. Zumbo, M. Stephens, and Y. Gilad. 2010. Sex-specific and
lineage-specific alternative splicing in primates. Genome Res. 20:180–189.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina
sequence data. Bioinformatics 30:2114–2120.

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

Bourgon, R., R. Gentleman, and W. Huber. 2010. Independent filtering increases detection
power for high-throughput experiments. Proc. Natl. Acad. Sci. U. S. A. 107:9546–9551.

Campos, J. L., K. J. Johnston, and B. Charlesworth. 2018. The effects of sex-biased gene

expression and X-linkage on rates of sequence evolution in Drosophila. Mol. Biol. Evol. 35:

655-665.

Canty A., and B. Ripley. 2021. *boot*: Bootstrap R (S-Plus) Functions. R package version

1.3-28.

Carneiro, M., F. W. Albert, J. Melo-Ferreira, N. Galtier, P. Gayral, J. A. Blanco-Aguiar, R.

Villafuerte, M. W. Nachman, and N. Ferrand. 2012. Evidence for widespread positive and

purifying selection across the european rabbit (*Oryctolagus cuniculus*) genome. Mol. Biol.

Evol. 29:1837–1849.

Chan A.H., P. A. Jenkins, and Y. S. Song. 2012. Genome-wide fine-scale recombi- nation rate

variation in Drosophila melanogaster. PLoS Genet. 8(12): e1003090.

Chang, P. L., J. P. Dunham, S. V. Nuzhdin, and M. N. Arbeitman. 2011. Somatic sex-specific

transcriptome differences in *Drosophila* revealed by whole transcriptome sequencing. BMC

Genomics 12.

Charlesworth, D., B. Charlesworth, and M. T. Morgan. 1995. The pattern of neutral molecular

variation under the background selection model. Genetics, 141: 1619-1632.

Cheng, C., and M. Kirkpatrick. 2016. Sex-specific selection and sex-biased gene expression in

humans and flies. PLoS Genetics 12: e1006170.

Cox, R. M., and R. Calsbeek. 2009. Sexually Antagonistic Selection, Sexual Dimorphism, and

the Resolution of Intralocus Sexual Conflict. Am. Nat. 173:176–187.

Cvijović, I., B. H. Good, and M. M. Desai. 2018. The effect of strong purifying selection on

genetic diversity. Genetics 209: 1235-1278.

Dapper, A. L., and M. J. Wade. 2016. The evolution of sperm competition genes: the effect of

mating system on levels of genetic variation within and between species. Evolution 70:502-

511.

Dapper, A. L., and M. J. Wade. 2020. Relaxed Selection and the Rapid Evolution of

Reproductive Genes. Trends Genet. 36:640–649. Elsevier Ltd.

Dean, R., and J. E. Mank. 2016. Tissue specificity and sex-specific regulatory variation permit
the evolution of sex-biased gene expression. Am. Nat. 188:E74–E84.

Delph, L. F., K. E. Brown, L. D. Rios, and J. K. Kelly. 2022. Sex-specific natural selection on
SNPs in *Silene latifolia*. Evol. Lett. https://doi.org/10.1002/evl3.283

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson,
and T. R. Gingeras. 2013. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics
29:15–21.

Dos Santos, G., A. J. Schroeder, J. L. Goodman, V. B. Strelets, M. A. Crosby, J. Thurmond, D.
B. Emmert, W. M. Gelbart, N. H. Brown, T. Kaufman, M. Werner-Washburne, R. Cripps,
K. Broll, L. S. Gramates, K. Falls, B. B. Matthews, S. Russo, P. Zhou, M. Zytkovicz, B.
Adryan, H. Attrill, M. Costa, S. Marygold, P. McQuilton, G. Millburn, L. Ponting, R.
Stefancsik, S. Tweedie, and G. Grumbling. 2015. FlyBase: Introduction of the *Drosophila*
melanogaster Release 6 reference genome assembly and large-scale migration of genome
annotations. Nucleic Acids Res. 43:D690–D697.

Dutoit, L., C. F. Mugal, P. Bolívar, M. Wang, K. Nadachowska-Brzyska, L. Smeds, H. P. Yazdi,
L. Gustafsson, and H. Ellegren. 2018. Sex-biased gene expression, sexual antagonism and
levels of genetic diversity in the collared flycatcher (*Ficedula albicollis*) genome. Mol.
Ecol. 27:3572–3581.

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

Ewels, P., M. Magnusson, S. Lundin, and M. Käller. 2016. MultiQC: Summarize analysis results
for multiple tools and samples in a single report. Bioinformatics 32:3047–3048.

Fraïsse, C, G. Puixeu Sala, and B. Vicoso. 2019. Pleiotropy Modulates the Efficacy of Selection
in Drosophila melanogaster. Mol Biol Evol. 36:500-515.

Frank, S. A., and M. M. Patten. 2020. Sexual antagonism leads to a mosaic of X-autosome
conflict. Evolution (N. Y). 74:495–498.

Gibilisco, L., Q. Zhou, S. Mahajan, and D. Bachtrog. 2016. Alternative Splicing within and
between *Drosophila* Species, Sexes, Tissues, and Developmental Stages. PLoS Genet.

12:e1006464.

Griffin, R. M., R. Dean, J. L. Grace, P. Rydén, and U. Friberg. 2013. The shared genome is a
pervasive constraint on the evolution of sex-biased gene expression. Mol. Biol. Evol. 30:
2168-2176.

Hadfield, J. D. 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models:
The MCMCglmm R Package. J. Stat. Softw. 33:1–22.

Hartley, S. W., and J. C. Mullikin. 2016. Detection and visualization of differential splicing in
RNA-Seq data with JunctionSeq. Nucleic Acids Res. 44:e127.

Hersh, E., J. A. Madjidian, S. Andersson, M. Strandh, W. S. Armbruster, and A. Lankinen. 2015.
Sexual antagonism in the pistil varies among populations of a hermaphroditic mixed-mating
plant. J. Evol. Biol. 28:1321–1334.

Ho, E. K. H., M. Bartkowska, S. I. Wright, and A. F. Agrawal. 2019. Population genomics of the
facultatively asexual duckweed *Spirodela polyrhiza*. New Phytol. 224:1361–1371.

Huang, W., M. A. Carbone, M. M. Magwire, J. A. Peiffer, R. F. Lyman, E. A. Stone, E. A., R. R.
H. Anholt and T. F. Mackay. 2015. Genetic basis of transcriptome diversity in *Drosophila*
melanogaster. Proc. Natl. Acad. Sci. U. S. A. 112: E6010-E6019.

Huang, W., A. Massouras, Y. Inoue, J. Peiffer, M. Ràmia, A. M. Tarone, L. Turlapati, T.
Zichner, D. Zhu, R. F. Lyman, M. M. Magwire, K. Blankenburg, M. A. Carbone, K. Chang,
L. L. Ellis, S. Fernandez, Y. Han, G. Highnam, C. E. Hjelmen, J. R. Jack, M. Javaid, J.
Jayaseelan, D. Kalra, S. Lee, L. Lewis, M. Munidasa, F. Ongeri, S. Patel, L. Perales, A.
Perez, L. L. Pu, S. M. Rollmann, R. Ruth, N. Saada, C. Warner, A. Williams, Y. Q. Wu, A.
Yamamoto, Y. Zhang, Y. Zhu, R. R. H. Anholt, J. O. Korbel, D. Mittelman, D. M. Muzny,
R. A. Gibbs, A. Barbadilla, J. S. Johnston, E. A. Stone, S. Richards, B. Deplancke, and T. F.
C. Mackay. 2014. Natural variation in genome architecture among 205 *Drosophila*
melanogaster Genetic Reference Panel lines. Genome Res. 24:1193–1208.

Huang, Y. F. 2022. Dissecting genomic determinants of positive selection with an evolution-
guided regression model. Mol. Biol. Evol. 39: msab291.

Innocenti, P., and E. H. Morrow. 2010. The Sexually Antagonistic Genes of *Drosophila*

melanogaster. PLoS Biol. 8:e1000335.

Jakšić, A. M., and C. Schlötterer. 2016. The interplay of temperature and genotype on patterns of

alternative splicing in *Drosophila melanogaster*. Genetics 204:315–325.

Kent, W. J., C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, and a. D.

Haussler. 2002. The Human Genome Browser at UCSC. Genome Res. 12:996–1006.

Khil, P. P., N. A. Smirnova, P. J. Romanienko, and R. D. Camerini-Otero. 2004. The mouse X

chromosome is enriched for sex-biased genes not subject to selection by meiotic sex

chromosome inactivation. Nat. Genet. 36:642–646.

Kidwell, J. F., M. T. Clegg, F. M. Stewart, and T. Prout. 1977. Regions of stable equilibria for

models of differential selection in the two sexes under random mating. Genetics 85:171–

183.

Lack, J. B., C. M. Cardeno, M. W. Crepeau, W. Taylor, R. B. Corbett-Detig, K. A. Stevens, C.

H. Langley, and J. E. Pool. 2015. The *Drosophila* genome nexus: A population genomic

resource of 623 Drosophila melanogaster genomes, including 197 from a single ancestral

range population. Genetics 199:1229–1241.

Lack, J. B., J. D. Lange, A. D. Tang, R. B. Corbett-Detig, and J. E. Pool. 2016. A thousand fly

genomes: An expanded *Drosophila* genome nexus. Mol. Biol. Evol. 33:3308–3313.

Lande, R. 1980. Sexual Dimorphism, Sexual Selection, and Adaptation in Polygenic Characters.

Evolution. 34:292–305.

Larracuente, A. M., T. B. Sackton, A. J. Greenberg, A. Wong, N. D. Singh, D. Sturgill D, Y.

Zhang, B. Oliver, and A. G. Clark. 2008. Evolution of protein-coding genes in *Drosophila*.

Trends Genet. 24: 114–123.

Long, Y., G. Song, J. Yan, X. He, Q. Li, and Z. Cui. 2013. Transcriptomic characterization of

cold acclimation in larval zebrafish. BMC Genomics 14.

Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion

for RNA-seq data with DESeq2. Genome Biol. 15:1–21.

MacKay, T. F. C., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles, D. Zhu, S. Casillas, Y.

Han, M. M. Magwire, J. M. Cridland, M. F. Richardson, R. R. H. Anholt, M. Barrón, C.

Bess, K. P. Blankenburg, M. A. Carbone, D. Castellano, L. Chaboub, L. Duncan, Z. Harris,

M. Javaid, J. C. Jayaseelan, S. N. Jhangiani, K. W. Jordan, F. Lara, F. Lawrence, S. L. Lee,

P. Librado, R. S. Linheiro, R. F. Lyman, A. J. MacKey, M. Munidasa, D. M. Muzny, L.

Nazareth, I. Newsham, L. Perales, L. L. Pu, C. Qu, M. Ràmia, J. G. Reid, S. M. Rollmann,

J. Rozas, N. Saada, L. Turlapati, K. C. Worley, Y. Q. Wu, A. Yamamoto, Y. Zhu, C. M.

Bergman, K. R. Thornton, D. Mittelman, and R. A. Gibbs. 2012. The *Drosophila*

melanogaster Genetic Reference Panel. Nature 482:173–178.

Mank, J. E. 2017. Population genetics of sexual conflict in the genomic era. Nat. Rev. Genet.

18:721–730.

Mank, J. E., L. Hultin-Rosenberg, M. Zwahlen, and H. Ellegren. 2008. Pleiotropic constraint

hampers the resolution of sexual antagonism in vertebrate gene expression. Am. Nat 171:

35-43.

McDonald, J. H., and M. Kreitman. 1991. Adaptive protein evolution at the Adh locus in

Drosophila. Nature 351: 652-654.

McIntyre, L. M., L. M. Bono, A. Genissel, R. Westerman, D. Junk, M. Telonis-Scott, L.

Harshman, M. L. Wayne, A. Kopp, and S. V. Nuzhdin. 2006. Sex-specific expression of

alternative transcripts in *Drosophila*. Genome Biol. 7:R79.

modENCODE Consortium. 2010. Identification of functional elements and regulatory circuits

by *Drosophila* modENCODE. Science 330: 1787–1797.

Meisel, R. P., J. H. Malone, and A. G. Clark. 2012. Disentangling the relationship between sex-

biased gene expression and X-linkage. Genome Res. 22:1255–1265.

Modrek, B., and C. Lee. 2002. A genomic view of alternative splicing. Nat. Genet. 30:13–19.

Modrek, B., A. Resch, C. Grasso, and C. Lee. 2001. Genome-wide detection of alternative

splicing in expressed sequences of human genes. Nucleic Acids Res. 29:2850–2859.

Montgomery, S. H., and J. E. Mank. 2016. Inferring regulatory change from gene expression: the

confounding effects of tissue scaling. Mol. Ecol. 25: 5114-5128.

Osada, N. 2007. Inference of expression-dependent negative selection based on polymorphism
and divergence in the human genome. Mol. Biol. Evol. 24: 1622-1626.

Osada, N., R. Miyagi, and A. Takahashi. 2017. Cis- and trans- regulatory effects on gene
expression. Genetics 206:2139–2148.

Paape, T., T. Bataillon, P. Zhou, T. J. Y. Kono, R. Briskine, N. D. Young, and P. Tiffin. 2013.
Selection, genome-wide fitness effects and evolutionary rates in the model legume
Medicago truncatula. Mol. Ecol. 22:3525–3538.

Page, A. J., B. Taylor, A. J. Delaney, J. Soares, T. Seemann, J. A. Keane, and S. R. Harris. 2016.
SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microb.
genomics 2:e000056.

Pan Q., Shai O., Lee L.J., Frey B.J., and B. J. Blencowe. 2008. Deep surveying of alternative
splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet.
40:1413–1415.

Parisi, M., R. Nuttall, P. Edwards, J. Minor, D. Naiman, J. Lü, M. Doctolero, M. Vainer, C.
Chan, J. Malley, S. Eastman, and B. Oliver. 2004. A survey of ovary-, testis-, and soma-
biased gene expression in *Drosophila melanogaster* adults. Genome Biol. 5:R40.

Parisi, M., R. Nuttall, D. Naiman, G. Bouffard, J. Malley, J. Andrews, S. Eastman, and B. Oliver.
2003. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression.
Science 299:697–700.

Parsch, J., and H. Ellegren. 2013. The evolutionary causes and consequences of sex-biased gene
expression. Nat. Rev. Genet. 14:83–87. Nature Publishing Group.

Pröschel, M., Z. Zhang, and J. Parsch. 2006. Widespread adaptive evolution of *Drosophila* genes
with sex-biased expression. Genetics 174:893–900.

Ramani, A. K., J. A. Calarco, Q. Pan, S. Mavandadi, Y. Wang, A. C. Nelson, L. J. Lee, Q.
Morris, B. J. Blencowe, M. Zhen, and A. G. Fraser. 2011. Genome-wide analysis of
alternative splicing in *Caenorhabditis elegans*. Genome Res. 21:342–348.

Ranz, J. M., C. I. Castillo-Davis, C. D. Meiklejohn, and D. L. Hartl. 2003. Sex-dependent gene

expression and evolution of the *Drosophila* transcriptome. Science (80-.). 300:1742–1745.

Reinke, V., I. S. Gil, S. Ward, and K. Kazmer. 2004. Genome-wide germline-enriched and sex-
biased expression profiles in *Caenorhabditis elegans*. Development 131:311–323.

Revil, T., D. Gaffney, C. Dias, J. Majewski, and L. A. Jerome-Majewska. 2010. Alternative
splicing is frequent during early embryonic development in mouse. BMC Genomics 11:399.

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

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

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

Rogers, T. F., D. H. Palmer, and A. E. Wright. 2020. Sex-Specific Selection Drives the
Evolution of Alternative Splicing in Birds. Mol. Biol. Evol. 38:519–530.

Rowe, L., S. F. Chenoweth, and A. F. Agrawal. 2018. The Genomics of Sexual Conflict. Am.
Nat. 192:274–286.

Sayadi, A., A. Martinez Barrio, E. Immonen, J. Dainat, D. Berger, C. Tellgren-Roth, B. Nystedt,
and G. Arnqvist. 2019. The genomic footprint of sexual conflict. Nat. Ecol. Evol. 3:1725–
1730.

Sohail, M., O. A. Vakhrusheva, J. H., Sul, S. L., Pulit, S. L., Francioli, L. C., et al. 2017.
Negative selection in humans and fruit flies involves synergistic epistasis. Science, 35: 539-
542.

Sorek, R., R. Shamir, and G. Ast. 2004. How prevalent is functional alternative splicing in the
human genome? Trends Genet. 20:68–71.

Stoletzki, N., and A. Eyre-Walker. 2011. Estimation of the neutrality index. Mol. Biol. Evol. 28:
63-70.

Sturgill, D., Y. Zhang, M. Parisi, and B. Oliver. 2007. Demasculinization of X chromosomes in

the *Drosophila* genus. Nature 450:238–241.

Tajima, F. 1989. Statistical Method for Testing the Neutral Mutation Hypothesis by DNA
Polymorphism. Genetics 123:585–595.

Tan, M. H., K. F. Au, A. L. Yablonovitch, A. E. Wills, J. Chuang, J. C. Baker, W. H. Wong, and
J. B. Li. 2013. RNA sequencing reveals a diverse and dynamic repertoire of the *Xenopus*
tropicalis transcriptome over development. Genome Res. 23:201–216.

Telonis-Scott, M., A. Kopp, M. L. Wayne, S. V. Nuzhdin, and L. M. McIntyre. 2009. Sex-
specific splicing in *Drosophila*: Widespread occurrence, tissue specificity and evolutionary
conservation. Genetics 181:421–434.

Trabzuni, D., A. Ramasamy, S. Imran, R. Walker, C. Smith, M. E. Weale, J. Hardy, and M.
Ryten. 2013. Widespread sex differences in gene expression and splicing in the adult human
brain. Nat. Commun. 4. Nature Publishing Group.

Walsh, B., and M. Lynch. 2020. Evolution and Selection of Quantitative Traits. Oxford
University Press, New York.

Wheeler, B., and M. Torchiano. 2016. *lmPerm*: Permutation Tests for Linear Models. *R* package
version 2.1.0, <https://CRAN.R-project.org/package=lmPerm>.

Williamson, R. J., E. B. Josephs, A. E. Platts, K. M. Hazzouri, A. Haudry, M. Blanchette, and S.
I. Wright. 2014. Evidence for Widespread Positive and Negative Selection in Coding and
Conserved Noncoding Regions of Capsella grandiflora. PLoS Genet. 10:e1004622.

Wright, A. E., M. Fumagalli, C. R. Cooney, N. I. Bloch, F. G. Vieira, S. D. Buechel, N. Kolm,
and J. E. Mank. 2018. Male-biased gene expression resolves sexual conflict through the
evolution of sex-specific genetic architecture. Evol. Lett. 2:52–61.

Wright, A. E., T. F. Rogers, M. Fumagalli, C. R. Cooney, and J. E. Mank. 2019. Phenotypic
sexual dimorphism is associated with genomic signatures of resolved sexual conflict. Mol.
Ecol. 28:2860–2871.

Yanai, I., H. Benjamin, M. Shmoish, V. Chalifa-Caspi, M. Shklar, R. Ophir, A. Bar-Even, S.
Horn-Saban, M. Safran, E. Domany, et al. 2005. Genome-wide midrange transcription

profiles reveal expression level relationships in human tissue specification. Bioinformatics
21: 650–659.

Yang, X., E. E. Schadt, S. Wang, H. Wang, A. P. Arnold, L. Ingram-Drake, T. A. Drake, and A.
J. Lusis. 2006. Tissue-specific expression and regulation of sexually dimorphic genes in
mice. Genome Res. 16:995–1004.

Yang, L., G. K. Smyth, and W. Shi. 2014. *featureCounts*: an efficient general purpose program
for assigning sequence reads to genomic features. Bioinformatics 30: 923–930

Zemp, N., R. Tavares, A. Muyle, D. Charlesworth, G. A. B. Marais, and A. Widmer. 2016.
Evolution of sex-biased gene expression in a dioecious plant. Nat. Plants 2:16168.