Sex-Biased Expression Is Associated With Chromatin State in Drosophila *melanogaster* and *Drosophila simulans*

Adalena V. Nanni,[1,2] Natalie Martinez,[1] Rita Graze,[3] Alison Morse,[1,2] Jeremy R.B. Newman,[2] Vaibhav Jain,[1] Srna Vlaho,[4] Sarah Signor,[5] Sergey V. Nuzhdin,[4] Rolf Renne,[1,2] and Lauren M. McIntyre *[,1,2]

[1]Department of Molecular Genetics and Microbiology, University of Florida, Gainesville, FL
[2]University of Florida Genetics Institute, University of Florida, Gainesville, FL
[3]Department of Biological Sciences, Auburn University, Auburn, AL
[4]Department of Biological Sciences, University of Southern California, Los Angeles, CA
[5]Department of Biological Sciences, North Dakota State University, Fargo, ND

***Corresponding author:** E-mail: lauren.m.mcintyre@gmail.com.

Associate editor: Patricia Wittkopp

Abstract

In *Drosophila melanogaster* and *D. simulans* head tissue, 60% of orthologous genes show evidence of sex-biased expression in at least one species. Of these, ~39% (2,192) are conserved in direction. We hypothesize enrichment of open chromatin in the sex where we see expression bias and closed chromatin in the opposite sex. Male-biased orthologs are significantly enriched for H3K4me3 marks in males of both species (~89% of male-biased orthologs vs. ~76% of unbiased orthologs). Similarly, female-biased orthologs are significantly enriched for H3K4me3 marks in females of both species (~90% of female-biased orthologs vs. ~73% of unbiased orthologs). The sex-bias ratio in female-biased orthologs was similar in magnitude between the two species, regardless of the closed chromatin (H3K27me2me3) marks in males. However, in male-biased orthologs, the presence of H3K27me2me3 in both species significantly reduced the correlation between *D. melanogaster* sex-bias ratio and the *D. simulans* sex-bias ratio. Male-biased orthologs are enriched for evidence of positive selection in the *D. melanogaster* group. There are more male-biased genes than female-biased genes in both species. For orthologs with gains/losses of sex-bias between the two species, there is an excess of male-bias compared to female-bias, but there is no consistent pattern in the relationship between H3K4me3 or H3K27me2me3 chromatin marks and expression. These data suggest chromatin state is a component of the maintenance of sex-biased expression and divergence of sex-bias between species is reflected in the complexity of the chromatin status.

Introduction

Chromatin accessibility is important for multiple levels of gene regulation, as well as in large scale modifications of expression such as in dosage compensation of sex chromosomes. Chromatin remodeling genes have roles in sex determination. For example, the sex determination gene *fru* aids in recruiting of histone deacetylase *Rpd3* and heterochromatin protein 1A encoded by *Su(var)205* (Ito, et al. 2012) and the expression of *fru* decreases with mutation of a histone demethylase *kdm4* (Lorbeck, et al. 2010). Sexually dimorphic chromatin modifications such as H3K9me2 (associated with closed chromatin) and H4K16ac (associated with open chromatin) have been reported (Brown and Bachtrog 2014). Therefore, we hypothesized that sex-bias in gene expression may be influenced by chromatin.

Sexual conflicts arise when the optima for a specific trait differ between the sexes and therefore selection differs between the sexes. These conflicts can come in two forms: interlocus and intralocus conflict (reviewed in Rice and Holland 1997; Chapman et al. 2003; Tregenza, et al. 2006; Bonduriansky and Chenoweth 2009; Cox and Calsbeek 2009; Schenkel, et al. 2018). Intralocus conflict occurs when the optimal fitness of a shared trait/locus is different between males and females, with different alleles favored in males and females. It has been argued that the degree of observed sexual dimorphism can signify the extent to which intralocus sexual conflict has been fully or partially resolved (Cox and Calsbeek 2009). In the whole fly, a small proportion (8.5%) of sex-biased genes have evidence of current intralocus conflict, or sexually antagonistic selection (Innocenti and Morrow 2010), indicating that in the majority of cases, any sex-biased expression observed in this study that is associated with intralocus conflict resolution would be expected to result from a history of partially or fully resolved intralocus conflict, rather than ongoing intralocus conflict. Little is known about possible mechanisms associated with changes in sex bias and

Article

putative conflict resolution. The approach taken in this study provides one entry point to addressing potential molecular mechanisms by examining the relationship between sex bias, divergence in sex bias, and chromatin state.

Sexually dimorphic gene expression tends to be rapidly evolving (e.g., Ellegren and Parsch 2007; Zhang, et al. 2007; Harrison, et al. 2015). Sex dimorphism of gene expression in Drosophila brain, eye, and antennal genes has been shown to be associated with sexually dimorphic behavior and sensory perception (Landry, et al. 2007; Kopp, et al. 2008; Shiao, et al. 2015; reviewed in Anholt, et al. 2020). Head tissue has also been observed to have an association between chromatin regulation and allelic imbalance in *D. melanogaster*-*D. simulans* interspecific hybrid expression (Graze, et al. 2012). Specifically, *D. simulans*-biased alleles were enriched for genes associated with "H3K4 methyltransferase activity," and *D. melanogaster*-biased alleles were enriched for genes associated with "H3K9 methyltransferase activity" (Graze, et al. 2012). This finding led us to hypothesize that there may be divergence in chromatin patterns between the species. We note that the potential increase of H3K4 methylation in *D. simulans* (Graze, et al. 2012) may not directly lead to more H3K4me3 marks but instead more alternative H3K4 methylation marks (me1 or me2) which were not evaluated in the study presented here. Similarly, H3K9 methylation, correlated with closed chromatin and silenced expression (reviewed in Boros 2012; Kimura 2013), may not necessarily be related to H3K27me3 presence.

Sex determination among Drosophila is conserved, with the splicing mechanisms and functions of the terminal transcription factors *fruitless* (*fru*) and *doublesex* (*dsx*) conserved across a multitude of insect species (Salvemini, et al. 2010; Shukla and Nagaraju 2010). Fru^M is the male-specific transcription factor encoded by *fru* and is a primary regulator of sex dimorphism in the Drosophila brain (Ito, et al. 1996; Ryner, et al. 1996; Kido and Ito 2002; Demir and Dickson 2005; Manoli, et al. 2005; Stockinger, et al. 2005; Rideout, et al. 2007; Kimura, et al. 2008; von Philipsborn, et al. 2011). Drosophila brains contain of a type of neuron that expresses specific isoforms of *fru*, called *fru-P1*-expressing neurons. These neurons have been implicated in male-specific courtship behaviors (von Philipsborn, et al. 2011; Tanaka, et al. 2017; Ding, et al. 2019; Sato, et al. 2020; Liang, et al. 2021) with an excess of male-biased expression previously observed in the *fru-P1*-expressing neurons (Newell, et al. 2016). In *fru-P1*-expressing neurons of 1-day-old *D. melanogaster* adults, genes with H3K4me3 present in the TSS are enriched in males compared to females, while the reverse is true in 10- to 12-day-old adults. These sex differences at the TSS support the role of *fru-P1*-expressing neurons in directing sex-specific behaviors (Ito, et al. 1996; Ryner, et al. 1996; Demir and Dickson 2005; Manoli, et al. 2005; Stockinger, et al. 2005; Goldman and Arbeitman 2007; reviewed in Yamamoto and Koganezawa 2013) and suggest sexually dimorphic chromatin accessibility is stage and cell-type specific (Palmateer, et al. 2021).

The extent to which sex-bias in expression is conserved in head tissue and the relationship between sex-bias and chromatin in heads have not been well explored. In combination with the wealth of resources available for Drosophila as a model organism, the comparison of *D. melanogaster* and *D. simulans* provides an exceptionally tractable model in which to explore the relationship between chromatin marks and sex-bias in expression in an evolutionary context. To this end, we assess the relationship between sex-biased expression and chromatin accessibility, specifically using the H3K4me3 and H3K27me2me3 marks, within and between these two closely related Drosophila species.

Results

We assayed males and females in the sister species, *D. melanogaster* and *D. simulans*, for gene expression ($n = 48$; 2 sexes × 2 genotypes × 2 species × 6 replicates), and chromatin ($n = 24$; 2 sexes × 1 genotypes × 2 species × 6 replicates). For each sample ChIP for the open chromatin mark, H3K4me3, and closed chromatin marks, H3K27me2me3 and input were collected. We compared the two sexes within each species, trends of sex-bias between species, and one-to-one orthologous loci for gene expression and chromatin. We evaluated the relationship between sex-bias in gene expression and chromatin status. In *D. melanogaster*, 2,556 genes on the X chromosome and 14,114 genes on the autosomes, and in *D. simulans*, 2,305 genes on the X and 12,504 genes on the autosomes had evidence of gene expression and/or chromatin marks. We performed extensive quality control of the data (supplementary sections 4 through 7, Supplementary Material online). For example, to evaluate whether genome quality affected the results, all analyses were also performed with both species mapped to *D. melanogaster* (FlyBase r6.17) and both species mapped to *D. simulans* (FlyBase r2.02). While there were a few genes with map bias, there was no evidence that genome quality impacted mapping (mapping rates were similar between species and slightly higher in *D. simulans*) and no trends reported were affected by the choice to map each species to its own genome rather than mapping both to one of the two genomes (supplementary section 5.3, Supplementary Material online).

Exonic regions were separated into nonoverlapping exonic features where alternative donor/acceptor sites were quantified separately from shared exonic regions, in order to capture the potential sex-specific exonic features in each gene (Newman, et al. 2018). Nonoverlapping exonic features were quantified as $C_{is} = \left(\sum (d_{ijs})/N_i\right) \times (Q/U_s)$, where d is the depth of reads at nucleotide j of feature i, N is the length of the feature, U_s is the upper quartile of $\left(\sum (d_{ijs})/N_i\right)$ values in sample s, and Q is the median of all U_s values within the given species (Bullard, et al. 2010; Dillies, et al. 2013).

If all exonic features in a gene were detected in only one sex, the gene was labeled as sex-limited. There were 764 genes (~6% of expressed genes) determined to be sex-limited in *D. melanogaster* and 530 genes (~4% of expressed genes)

in *D. simulans* (supplementary file 1, Supplementary Material online, supplementary file 2, Supplementary Material online, *flag_sex_limited*==1). Differential expression analyses were performed separately for each exonic feature detected in both sexes of each species. Genes were considered sex-biased in expression if at least one exonic feature was statistically significantly differentially expressed between sexes. Genes with both significantly male- and female-biased exonic features were designated "male-biased and female-biased" and are expected in some genes that are sex-specifically alternatively spliced, such as the sex determination gene *dsx* (supplementary fig. S1, Supplementary Material online).

ChIP samples were compared to input controls for genomic features (transcription start sites, 5′, 3′ UTR's, exonic features and introns). Genomic features were considered detected above the input control (DAI) in H3K4me3 if $C_{K4,\,is} > C_{Input,\,is}$, in more than 50% of the replicates for that species-sex combination, and as $C_{K27,\,is} > C_{Input,\,is}$, for H3K27me3me4. ChIP data were found to be high quality and conform with general expectations for detection of the marks (supplementary sections 7.1–7.2, Supplementary Material online). A gene was considered as having a mark if at least one exonic feature in the gene was DAI. A gene was considered sex-limited (male/female) when marks were detected in only one sex.

Consistent Sex Bias in Gene Expression Between Orthologs

Gene expression in head tissues was measured in independent replicates of males and females for each species ($n = 48$, 2 species × 2 sexes × 2 genotypes × 6 replicates). In this experiment of head tissue, 60% of genes show evidence of sex-bias in at least one of the two species. Of these ~39% (2,192) are conserved in direction and magnitude between the two species (table 1, fig. 1). To compare the relative magnitude of sex bias between the sexes and species, we calculate a sex-bias ratio defined between 0 and 1. For genes with male-biased expression this ratio is $\left(1 - \frac{\hat{f}}{\hat{m}}\right)$ and for female-biased expression this ratio is $\left(1 - \frac{\hat{m}}{\hat{f}}\right)$; where $\hat{m}$ is the normalized male expression, and $\hat{f}$ is the normalized female expression. We compared the linear relationship between the sex-bias ratio in *D. melanogaster* to *D. simulans* for orthologs with consistent sex-biased expression between the species: female-biased orthologs (Pearson's $r_f = 0.69$; T-test $H_0{:}r_f = 0$, $P < 0.0001$) and male-biased orthologs (Pearson's $r_m = 0.49$; T-test $H_0{:}r_m = 0$, $P < 0.0001$) (fig. 1A). A linear regression of the sex-bias ratio of *D. melanogaster* onto *D. simulans* would have a regression coefficient of 1 if the sex-bias ratio was the same between the two species, and less than 1 if *D. melanogaster* sex-bias ratios were larger than *D. simulans* across orthologs. We observed greater sex-bias ratios in *D. melanogaster* compared to *D. simulans* for both female-biased orthologs (T-test $H_0{:}\beta_{1f} < 1$, $P < 0.0001$) and male-biased orthologs (T-test $H_0{:}\beta_{1m} < 1$, $P < 0.0001$).

Excess of Male-bias Compared to Female-bias in Both Species and in Gains/Losses in Sex-bias

There is substantial evidence for excess of male bias in the head tissue for both species. There are more genes expressed only in males (male-limited) than genes expressed only in females (female-limited) in both species (*D. melanogaster*, 566/198, Binomial $P < 0.0001$; *D. simulans*, 340/190, binomial $P < 0.0001$). There are more orthologs with consistent male-limited expression between the species compared to female-limited expression (31/12, binomial $P \approx 0.005$). When genes are expressed in both sexes, there is an excess of male-biased expression compared to female-biased expression observed in both *D. melanogaster* (2723/2185, binomial $P < 0.0001$) and *D. simulans* (2160/1873, Binomial $P < 0.0001$) and the same trend is observed in consistently sex-biased orthologs although it is not significant after correcting for multiple testing (1153/1038, binomial $P \approx 0.01$, table 1).

In orthologous genes expressed in both sexes, sex bias can be gained/lost between the species. In these genes there is an enrichment for male-bias gains/losses compared to female-bias gains/losses, regardless of which of the two species shows the sex-bias (table 1, rows 11 vs. 12 and 14 vs. 15; supplementary fig. S2C, Supplementary Material online). While some gains/losses of sex-bias may be due to small shifts in the expression bias, we find that most of the orthologs with significant sex-bias ratio in one species have values close to zero in the other species, demonstrated by the highest density of gene sex-bias ratios closest to the horizontal/vertical line, as indicated by the ellipses (fig. 1B and C). Male-biased gains/losses tend to have larger magnitudes of sex-bias, as well as greater variability in sex-bias ratio, compared to female-biased gains/losses (fig. 1B and C).

Male-bias in Orthologs Is Associated With Signatures of Positive Selection

We hypothesized that there may be differences in adaptive evolution for specific patterns of sex-bias in expression or chromatin marks. The comparative genomics database, flyDIVas (Clark, et al. 2007; Stanley and Kulathinal 2016) provides gene-level estimates of divergence with nonsynonymous (dN) to synonymous substitution (dS) rates (dN/dS) and tests of positive selection using PAML (Yang 1997) for the melanogaster subgroup (*D. melanogaster*, *D. simulans*, *D. sechellia*, *D. yakuba*, *D. erecta*), melanogaster group (melanogaster subgroup and *D. ananassae*), and the 12 Drosophila species (melanogaster group and *D. pseudoobscura*, *D. persimilis*, *D. willistoni*, *D. mojavensis*, *D. virilis*, and *D. grimshawi*). These different group allow for evaluation of selection across these three levels of phylogenic depth; however, the number of orthologous loci does decline as the distance from melanogaster increases and the tests at the 12 genome level are to be thought of as suggestive (Stanley and Kulathinal 2016). The null hypotheses tested are codon-based tests of positive Darwinian selection based on d_N/d_s (ω) ratios estimated

Table 1. Sex Bias in Expression.

	Pattern of sex-bias	*D. melanogaster*	*D. simulans*	Orthologs		
1	Male-biased	2723	2160	1153	}	$P = 0.01$
2	Female-biased	2185	1873	1038		
3	Male- and Female-biased	142	100	10		
4	Unbiased	6666	7410	3815		
5	Reversal	Male	Female	113		
6	Reversal	Female	Male	70		
7	Gain/loss	Male	Male and female	42		
8	Gain/loss	Female	Male and Female	16		
9	Gain/loss	Male and Female	Male	31		
10	Gain/loss	Male and Female	Female	35		
11	Gain/loss	Male	Unbiased	1049	}	$P < 0.0001$
12	Gain/loss	Female	Unbiased	872		
13	Gain/loss	Male and Female	Unbiased	53		
14	Gain/loss	Unbiased	Male	657	}	$P \approx 0.0001$
15	Gain/loss	Unbiased	Female	525		
16	Gain/loss	Unbiased	Male and female	26		
17	Expressed	11,716	11,543	9505		

The pattern of sex-bias is defined based on mutually exclusive categories. In rows 1–4, the number of genes following the pattern of sex-bias are given in columns 3 (*D. melanogaster*) and 4 (*D. simulans*), and the number of orthologous is given in column 5. Binomial test probabilities are indicated to the right of the table for the comparison of male-biased versus female-biased for consistent orthologs and gains/losses. *P*-values are in black if below the threshold of $P = 0.001$ and gray if above the threshold. Reversal of sex-bias is rare, only two percent (183/9508) of orthologs. Genes on chromosome 4 and on scaffolds, as well as those that change location, were omitted. Values for the X and autosomes separately are in supplementary table S1, Supplementary Material online. Each of the categories represented are indicated in supplementary File 1, Supplementary Material online for *D. melanogaster*, supplementary File 2, Supplementary Material online for *D. simulans*, and supplementary File 8, Supplementary Material online for orthologs.

by PAML model M0 (Yang 1997). FDR corrected *P*-values for likelihood ratio tests comparing pairs of site-specific models were obtained from the flyDIVas database to determine if: 1) model M2a (positive selection) has better fit than M1a (nearly neutral), 2) model M8 (beta+ω>1) (Yang 1997) has better fit than M7 (beta-distributed), and 3) model M8a (beta+ω=1) (Swanson, et al. 2003; Wong, et al. 2004) has better fit than model M8 (beta +ω>1). Each of these comparisons can be used to infer the presence of positive selection, but vary slightly in statistical properties (Wong, et al. 2004). This resulted in a set of nine tests of association corresponding to patterns of sex-biased expression and sex-biased marks (e.g., consistent male-biased expression) for each phylogenetic group and model comparison (supplementary table S2, Supplementary Material online). We used $P < 0.001$ as the significance threshold and find that only the *D. melanogaster-D. simulans* male-biased orthologs are consistently significantly enriched for positive selection. This is a consistent inference for 8 of the 9 tests with the exception being test M8 versus M8a in the 12 species comparison (χ^2: $P = 0.10$) (supplementary table S2, Supplementary Material online). We also found evidence of positive selection in many cases for the *D. melanogaster* subgroup and *D. melanogaster* group among genes with sex-specific H3K27me2me3 in males or females.

Sex-bias Is Associated With H3K4me3

We propose a model, "Open in Same sex and/or Closed in Opposite" (OS-CO), as an expectation of chromatin accessibility states for genes with sex-biased expression (fig. 2; supplementary fig. S3, Supplementary Material online). We expect chromatin in male-biased genes to have 1) H3K4me3 chromatin marks ("open") in males, and/or 2) H3K27me2me3 chromatin marks ("closed") in females. We test this expectation by comparing the chromatin state in male-biased orthologs to unbiased orthologs using Fisher exact test (Fisher 1934). Under the null hypothesis that chromatin is independent of sex-bias there should be no difference in the proportion of genes with H3K4me3 in males in these two groups. Similarly, we compare the presence of H3K4me3 chromatin marks in females between female-biased orthologs and unbiased orthologs. In both species, H3K4me3 chromatin marks in females are more likely to occur in female-biased orthologs relative to unbiased orthologs (*D. melanogaster*, $P < 0.0001$; *D. simulans*, $P < 0.0001$; fig. 3A) and H3K4me3 chromatin marks in males are enriched in orthologs with consistent male expression bias compared to orthologs with consistent unbiased expression (*D. melanogaster*, $P < 0.0001$; *D. simulans*, $P < 0.0001$; fig. 3B). In orthologous genes with gains/losses of sex bias, *D. melanogaster* and *D. simulans* show similar trends for the H3K4me3 chromatin marks and for the female H3K27me2me3 (fig. 3C and D). However, female-biased gains/losses in *D. simulans* have an enrichment for male H3K27me2me3, a pattern not observed in *D. melanogaster* and not present in orthologs with consistent female-bias.

Sex-biased Orthologs Have Conserved Presence of H3K4me3 Marks

In both species, the vast majority sex-biased orthologs have H3K4me3 marks in the sex with greater expression (fig. 3; supplementary fig. S6, Supplementary Material online) consistent with our model (fig. 2; supplementary fig.

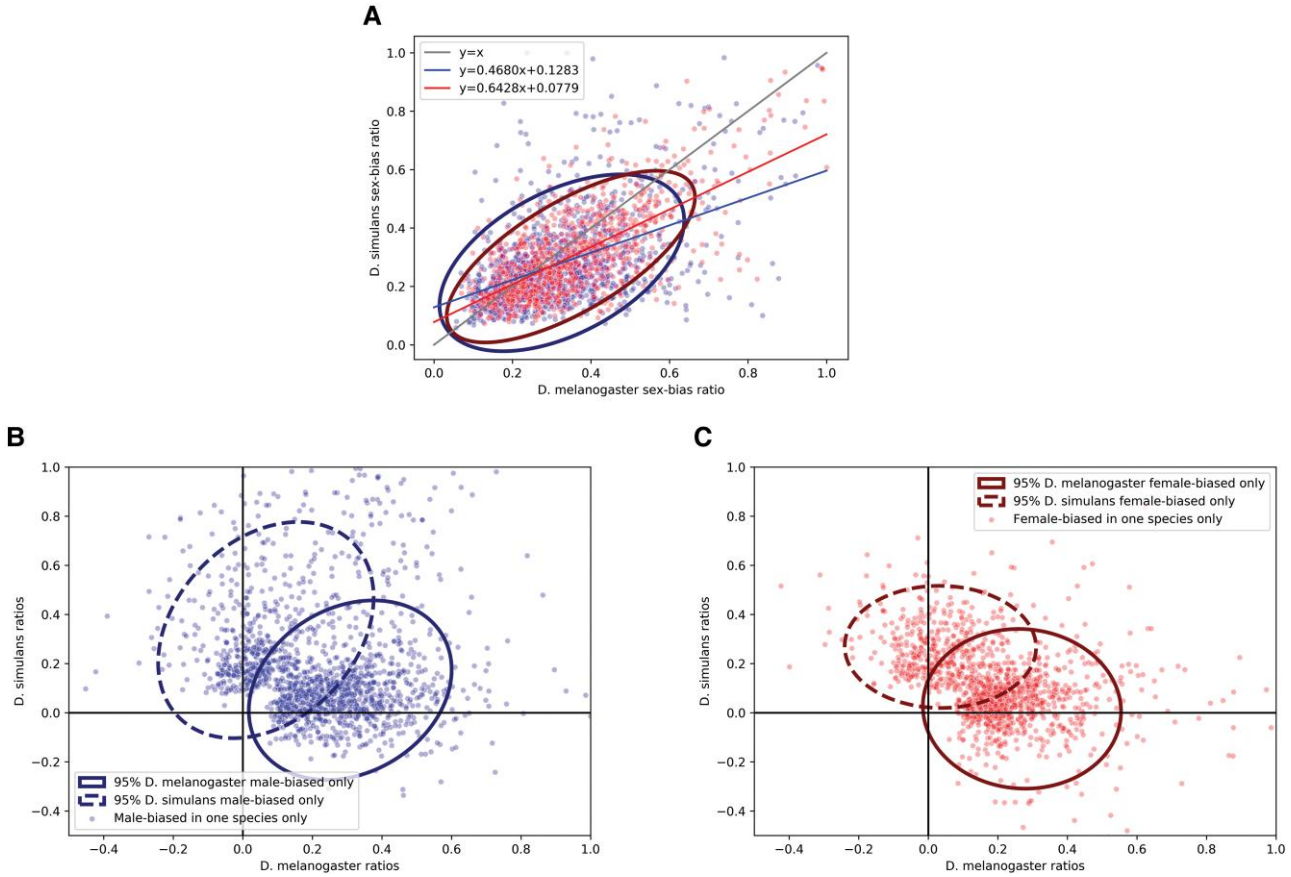

Fig. 1. Sex-bias ratios in orthologs. Sex-bias ratio in *D. melanogaster* (X-axis) and *D. simulans* (Y-axis). Female sex-bias is plotted in red, and male-bias is in blue. A value close to 1 indicates extreme sex-bias, while a value close to 0 indicates low sex-bias. (Panel A) Orthologous genes where sex-bias is in the same direction between the two species. A linear regression line with ellipses representing the 95th percentile of the observed data for males and females separately. The gray line represents the value of the regression if the sex-ratio is equal in the two species ($y = x$, $\beta = 1$). (Panel B) Gains and losses in male sex-bias. (Panel C) Gains and losses in female sex-bias. The solid ellipses represent the 95th percentile of the orthologs with significant sex-bias in *D. melanogaster*. The dashed ellipses represent 95th percentile of the orthologs with significant sex-bias in *D. simulans*.

S3, Supplementary Material online). Male-biased orthologs are significantly enriched for conserved H3K4me3 marks in males (~89% of male-biased orthologs vs. ~76% of unbiased orthologs; χ^2: $P < 0.0001$). Similarly, female-biased orthologs are significantly enriched for conserved H3K4me3 marks in females (~90% of female-biased orthologs vs. ~73% of unbiased orthologs; χ^2: $P < 0.0001$). Our model predicts that sex-bias in expression may also result from repression of expression in the opposite sex (closed in opposite sex, CO). We find that the relationship of sex-biased expression and H3K27me2me3 repressive marks in the opposite sex of the sex bias depends upon the direction of the sex-bias (fig. 3A and B), indicating the sexes may have evolved to have different mechanisms of sex-biased expression regulation. When male-biased orthologs have conserved H3K4me3 marks in males, the sex-bias ratio in the two species is more similar when there are no H3K27me2me3 mark in females (fig. 4A; $\beta_{1A} = 0.50$), than when both species have a female H3K27me2me3 mark (fig. 4C; $\beta_{1C} = 0.18$; β_{1A} vs. β_{1C}: $P < 0.0001$). When the female H3K27me2me3 mark is not conserved between the species (present in either species but not both), the sex-bias ratio is similar (although trending lower, $\beta_{1B} = 0.41$) to when both species have the female H3K27me2me3 mark (fig. 4B; β_{1A} vs. β_{1B}: $P \approx 0.16$). In contrast, the relationship in the sex-bias ratio between the species does not change in female-biased orthologs based on the presence of the male H3K27me2me3 chromatin mark (fig. 4D and F; β_{1D} vs. β_{1F}: $P \approx 0.47$, β_{1D} vs. β_{1E}: $P \approx 0.93$).

Discussion

In *D. melanogaster* and *D. simulans* head tissue, ~60% of genes show evidence of sex-bias in at least one of the two species. Of these, ~39% (2,192) are conserved in direction between the two species. The direction of sex-bias in expression agreed between the two species much more frequently than expected by chance (male-bias: $\kappa = 0.41$, $P < 0.0001$; female bias: $\kappa = 0.45$, $P < 0.0001$). This agreement in presence/absence of sex-bias between *D. melanogaster* and *D. simulans* may be due to the short evolutionary time and the maintenance of the ancestral state, where the sex-bias in the common ancestor is random. Under the null hypothesis that the direction of

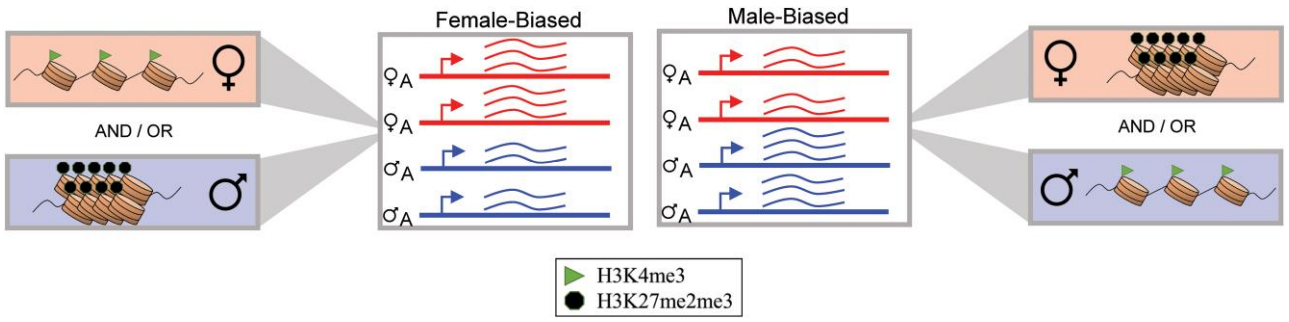

FIG. 2. "Open in Same and/or Closed in Opposite" (OS-CO): a model for chromatin accessibility patterns for sex-biased expression. Genes with female-biased expression (left) are expected to have open chromatin marks (H3K4me3) in females and/or closed chromatin marks (H3K27me2me3) in males. Genes with male-biased expression patterns are expected to have open chromatin marks (H3K4me3) in males and/or closed chromatin marks (H3K27me2me3) in females. Not all sex-biased genes are expected to have these marks as there are other chromatin marks and regulatory factors that influence chromatin state.

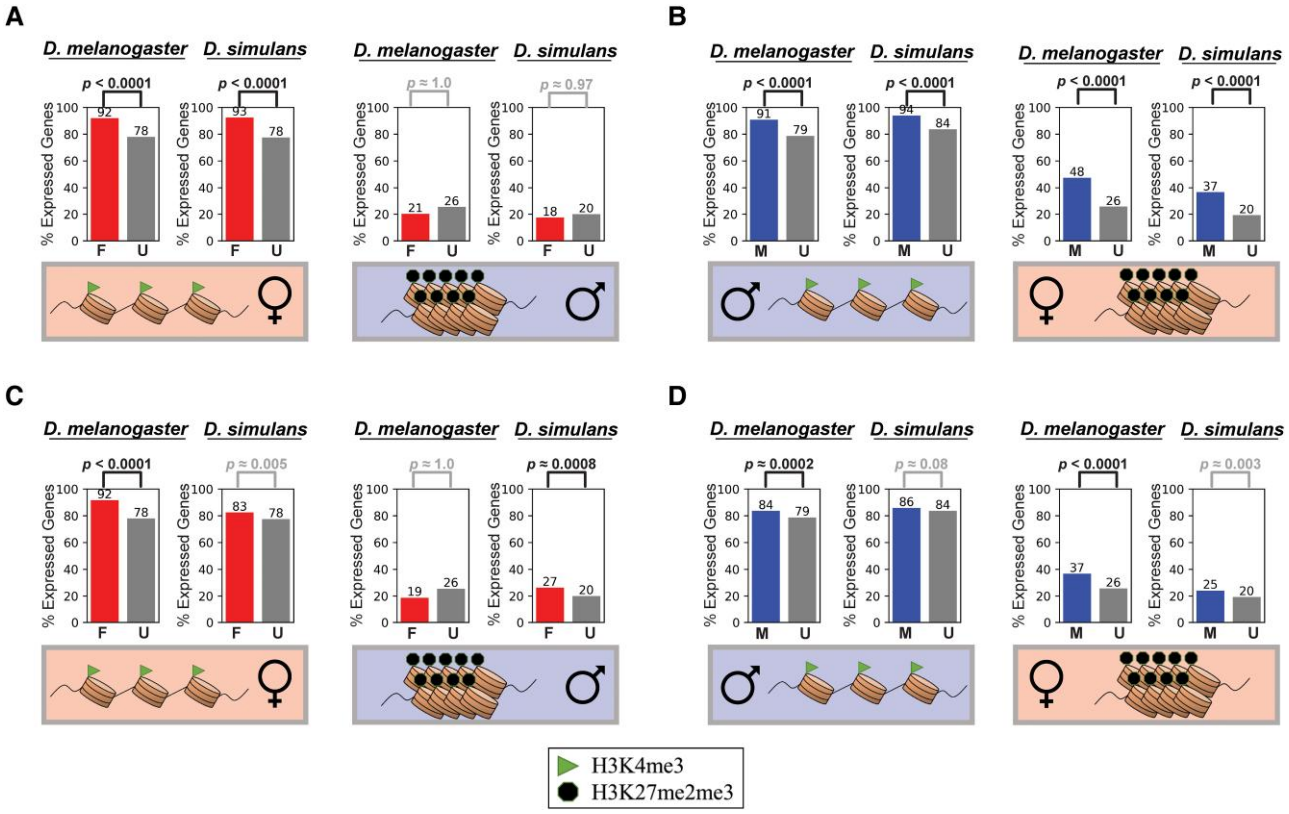

FIG. 3. Chromatin association differs between consistent and inconsistent sex-biased expression. Orthologs with consistent sex-biased expression in *D. melanogaster* and *D. simulans* (A, B) and with gains/losses of sex-bias between the species (C, D). The Y-axis of each graph represents the percent of genes with a chromatin mark shown in the cartoon below the graph. Numbers of genes in each category are given in table 1. For female-biased (F) and male-biased (M) genes and unbiased (U). (*A*) Consistent female-biased orthologs, (*B*) Consistent male-biased orthologs, (*C*) Gains/losses in female-bias (*D*) Gains/losses in male-bias. *P*-values are reported for a one-sided Fisher's exact test (Fisher 1934). Significant *P*-values ($P < 0.001$) are in black. Tests separating the X and autosomes are in supplementary figure S6, Supplementary Material online.

bias is random, we expect to see approximately an even number of gains/losses in transitions between the two species from unbiased to male- or female-biased. In a binomial test, the null hypothesis of equal probability for male/female gain/loss ($P = 0.5$) is rejected for both transitions from unbiased genes in *D. melanogaster* to sex biased genes in *D. simulans* (~55% male-biased, Binomial $P < 0.0001$) and unbiased genes in *D. simulans* to sex-biased genes in *D. melanogaster* (~56% male-biased, Binomial $P \approx 0.0001$). There is also more male-bias than female-bias in sex-limited expression ($P < 0.0001$ for both species) and

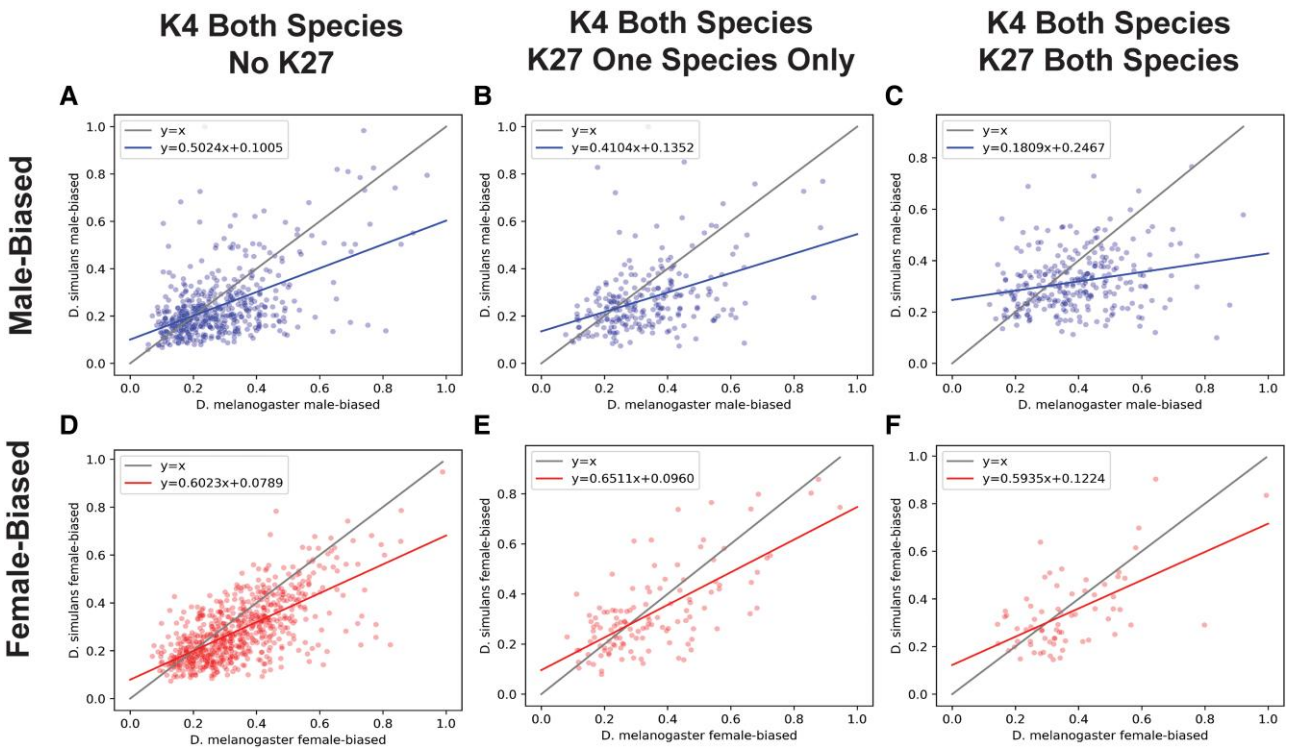

FIG. 4. Male and female biased orthologs. Orthologous genes where sex-bias is in the same direction between the two species. A linear regression line for the regression of the sex-bias ratio in *D. melanogaster* (X-axis) on *D. simulans* (Y-axis). (Panel A; $\beta_1 = 0.50$) presence of male H3K4me3 and absence of female H3K27me2me3 (K27) in both species, (Panel B; $\beta_1 = 0.41$) presence of male H3K4me3 in both species and female H3K27me2me3 in either species, and (Panel C; $\beta_1 = 0.18$) presence of male H3K4me3 and female H3K27me2me3 in both species. For female-biased orthologs: (Panel D; $\beta_1 = 0.60$) presence of female H3K4me3 and absence of male H3K27me2me3 in both species, (Panel E; $\beta_1 = 0.65$) presence of female H3K4me3 in both species and male H3K27me2me3 in either species, and (Panel F; $\beta_1 = 0.59$) presence of female H3K4me3 and male H3K27me2me3 in both species.

the number of male-biased orthologs are greater than the number of female biased orthologs.

There is conservation in the splicing regulation observed in the Drosophila sex determination pathway (Salvemini, et al. 2010; Shukla and Nagaraju 2010). In this study 19 genes in the sex-determination pathway were expressed in both sexes and species. Eleven of these were consistent in sex-bias between the two species, including the terminal transcription factors of the sex determination pathway, *dsx* and *fru*. We hypothesized that the conservation of sex-biased splicing in these transcription factors may contribute to consistent sex-bias between the two species. Both genes have male- and female-specific isoforms (supplementary fig. S1, Supplementary Material online) and *dsx* contributes to the regulation of sexual dimorphism in the brain of both sexes (Rideout, et al. 2007, 2010; Kimura, et al. 2008; Arbeitman, et al. 2016). Male-biased orthologs were enriched for genes regulated by *dsx* (χ^2: $P < 0.0001$). The other transcription factor, *fru*, is highly conserved in sex-specific splicing across insects (Salvemini, et al. 2010) and Fru^M is associated with chromatin remodeling factors (Lorbeck, et al. 2010; Ito, et al. 2012). Male-biased orthologs were enriched for genes regulated by the Fru^M protein in *D. melanogaster* males (χ^2: $P < 0.0001$) and female-biased orthologs were depleted for signatures of Fru^M (χ^2: $P = 0.002$).

Sex-bias is conserved in magnitude, as well as direction, between *D. melanogaster* and *D. simulans*. Intriguingly, sex-bias ratios for expression are more similar between the species in females than males, suggesting there may be less evolutionary constraint in males, or potentially a difference in selection between the sexes. There is evidence for positive selection in male-biased orthologs, but not in female-biased orthologs. Further, in male-biased orthologs, the magnitude of sex-bias differs between male-biased orthologs where there are female H3K27me2me3 chromatin marks in both species compared to male-biased orthologs where there are no H3K27me2me3 chromatin marks in females.

The findings that when female H3K27me2me3 marks are absent in both species the male sex-bias ratio is more similar between the species than when there is a mark in only one species, suggests that the H3K27me2me3 chromatin marks may play a role in resolving the ongoing sexual conflict in males. The divergence in the degree of male-bias is associated with female H3K27me2me3 marks which are predicted to reduce expression in females. It is possible that this reflects a mechanism of resolving cases of intralocus conflict in which expression of an allele in females has deleterious effects.

Some evidence toward this hypothesis is the observation that in genes with female bias in *D. simulans* (and unbiased in *D. melanogaster*) there is an excess of H3K27me2me3 in males of *D. simulans* compared to unbiased orthologs. This may suggest that genes with female-biased expression in *D. melanogaster* either 1) do not involve deleterious effects in *D. melanogaster* males, 2) involve genes that are important for male fitness and are incompatible with H3K27me2me3 marks and gene silencing, or 3) do not involve resolution of intralocus sexual conflict. These overall patterns reveal specific testable hypotheses regarding the role of activation and repression via chromatin modifications in the resolution of intralocus sexual conflict for future experiments. It is likely that our observations using these specific marks do not completely reflect final active or repressed states of expression resulting from the chromatin state as a whole, as we assayed only two of the many possible marks. Our study does not demonstrate a causal relationship between chromatin accessibility and sex-biased expression, nor do we claim to provide a comprehensive survey of chromatin accessibility. Rather, our findings likely reflect the role of different regulators that impact chromatin states. We have demonstrated that a surprising number of orthologs have consistent sex-bias and that H3K4me3 and H3K27me2me3 provide potential insight into the maintenance of sex-bias between the two closely related species *D. simulans* and *D. melanogaster*.

Methods

Experimental Design

Isogenic male and female *D. melanogaster* (DGRP r153 and r301) (Mackay, et al. 2012) and *D. simulans* (Winters lines sz11 and sz12) (Signor 2017) flies were raised on standard Bloomington recipe medium at 25 °C with a 12-h light/dark cycle. There were 2 sexes and 2 genotypes for each species with 6 replicates for a total of 48 samples. These samples were collected as part of a larger project that included exposure to ethanol. The analyses presented here focus on sex and species differences. Samples were flash frozen in liquid nitrogen and freeze dried (supplementary fig. S7, Supplementary Material online).

For RNA-seq, 12 heads from each sample were collected. mRNA purification, cDNA synthesis, and dual index barcoding library preparation were carried out by Rapid Genomics (Gainesville, FL, http://rapid-genomics.com). Individual libraries ($n = 48$) were pooled in equimolar ratios as estimated by Qubit and sequenced on7 Illumina lanes at Rapid Genomics (paired-end 2×100 3 lanes with HiSeq 3000 and paired-end 2×150 2 lanes with HiSeq X and 2 lanes with NovaSeq 6000). External RNA Control Consortium spike-in control was used to evaluate the quality of all RNA-seq sequencing libraries (Jiang, et al. 2011).

For ChIP-seq, ~200 heads from each sample of *D. melanogaster* r301 and *D. simulans* sz11 were collected (2 species $\times$ 6 replicates $\times$ 2 sexes $\times$ 1 genotype = 24 samples). Each sample was used to assay histone marks H3K4me3 (open chromatin), H3K27me2me3 (closed chromatin), and input. (3 antibodies/input $\times$ 24 samples = 72 assays). One r301 female untreated sample contained ~175 heads and one 2 sz11 male ethanol treated sample contained ~120 heads, and one sz11 ethanol treated female sample contained ~50 heads. Despite this low number, ChIP was successful in this sample. A full protocol for the ChIP (supplementary file 3, Supplementary Material online, developed by NM and RR) is available in supplementary file 1, Supplementary Material online. ChIP samples were indexed, pooled, and sequenced on one lane of an Illumina HiSeq2500 (paired-end 2×100) at the University of Florida, ICBR (Gainesville, FL, https://biotech.ufl.edu/).

Genome Annotations

All genome and annotation versions used were from FlyBase release FB2017_04 (http://www.flybase.org) *D. melanogaster* FlyBase r6.17 and *D. simulans* FlyBase r2.02. The FlyBase gene OrthoDB ortholog report (Waterhouse, et al. 2013) (supplementary File 4, Supplementary Material online) was used to identify one-to-one orthologous gene pairs (one gene in *D. melanogaster* associating with one gene in *D. simulans*, and vice versa).

We created BED files for both genic features (exons, exonic features, TSS +/− 150 bp, 5′ UTR, 3′UTR, and introns) and intergenic features (defined as the non-genic features greater than 50 bp in length) for each reference from the relevant GFF annotation file. We note that in areas where there were overlapping exons (where intron/exon boundaries vary by transcript), alternative donor and acceptor sites were defined as exonic and tracked as separate features in downstream analyses (Newman, et al. 2018). Counts of each unique feature type are in supplementary table S4, Supplementary Material online. We note that there are fewer genic features annotated in *D. simulans* compared to *D. melanogaster*.

RNA-seq and ChIP-seq

Results were consistent with high quality data for the technology deployed. Sequencing adapters were removed from both RNA-seq and ChIP-seq reads using Cutadapt version 2.1 (Martin 2011) with a max error rate of 0.1 and a minimum overlap of 3 nt. Forward and reverse reads were merged using BBMerge (Bushnell, et al. 2017). Reads less than 14 bp + 50% original read length were not considered further. Identical reads were identified (fastqSplitDups.py) and removed. The resulting processed reads consisted of 1) merged reads ('single-end'), 2) unmerged reads without a proper pair ('single-end'), and 3) unmerged reads with proper pairs ('paired-end').

Processed RNA-seq and ChIP reads were aligned to the corresponding genome reference (*D. melanogaster* reads mapped to *D. melanogaster* FlyBase r6.17 and *D. simulans* reads mapped to *D. simulans* FlyBase r.202) using BWA-MEM v0.7.15 (Li 2013) as single-end or paired-end with default parameters. To determine if there was any systematic reference bias, processed RNA-seq reads from

D. melanogaster samples were mapped to the *D. simulans* FlyBase r.202 genome, and *D. simulans* samples were mapped to the *D. melanogaster* FlyBase r6.17 genome. A small bias was observed toward mapping to the *D. simulans* genome. In both species female samples tended to have, on average, slightly higher mapping rates in the ChIP experiment. Sensitivity to mapping bias was examined and results are described in detail in (supplementary section 5.3, Supplementary Material online).

RNA-seq Feature Detection

A feature was considered detected by RNA-seq if at least one read was present in more than 50% of the replicates for a species-sex combination (e.g., present in at least 7 of the 12 female or male replicates for a given species). The number of detected features for each species-sex combination is summarized in supplementary table S5, Supplementary Material online. There are fewer features in *D. simulans* and, despite the slightly higher mapping rates found in *D. simulans* samples, there are slightly fewer features detected in *D. simulans* samples compared to *D. melanogaster* across all feature types except for 3′ UTR. The 3′UTR features have a higher proportion of detection in *D. simulans* compared to *D. melanogaster*, suggesting there may be a systematic bias in the 3′UTR regions of the two species of either an over-annotation of these regions in *D. melanogaster* or an under-annotation in *D. simulans*. If this is due to under-annotation in *D. simulans*, then we expect more expression in intergenic features in *D. simulans* compared to *D. melanogaster*. However, there is a lower proportion of detected intronic and intergenic features in *D. simulans* samples compared to D. *melanogaster* samples. Exonic feature detection was similar between the species, with slightly higher detection rates in D. *melanogaster* males. A feature was considered sex-limited if the feature was detected in only one of the 2 sexes. Approximately 4% of exonic features were sex-limited in *D. melanogaster* samples (2,530 in males, 1,195 in females) and *D. simulans* (1,801 in males, 1,506 in females).

For the gene expression analysis, exonic regions were separated into non-overlapping exonic features where alternative donor/acceptor sites were quantified separately from shared exonic regions in order to capture the potential sex-specific structures in the gene (Newman et al. 2018). Genes were defined as detected if at least one exonic feature was detected for either sex. There are a similar number but proportionally more genes detected in *D. simulans* (11,543 of 15,385, ∼75%) compared to *D. melanogaster* (11,716 of 17,737, ∼66%). This indicates that there are no large quality differences in the *D. simulans* genome compared to the *D. melanogaster* genome despite the differences in annotation.

To compare genes across *D. melanogaster* and *D. simulans*, we focus on annotated orthologs from the OrthoDB ortholog report (Waterhouse, et al. 2013) to identify one-to-one orthologous gene pairs (one gene in *D. melanogaster* associating with one gene in *D. simulans*, and vice versa) (supplementary file 4, Supplementary Material online). There are 14,601 orthologous gene pairs between the species, 12,386 of which are one-to-one orthologs. Genes on chromosome 4, the Y chromosome, and scaffolds of either species (138 orthologs) were excluded from further analysis. There were 8 genes on the X chromosome of *D. melanogaster* with orthologs on autosomes of *D. simulans*, and 1 gene on the X of *D. simulans* with an ortholog on an autosome of *D. melanogaster*. These 9 genes were also excluded. The remaining 12,239 one-to-one orthologous genes on the X ($n = 1{,}877$) and autosomes ($n = 10{,}362$) of both species were carried forward. Of these, 11,937 (1,840 on the X and 10,097 on the autosomes) had evidence of expression and/or chromatin accessibility.

RNA-seq Differential Expression

For each species, exonic features were quantified as $C_{is} = \left(\sum (d_{ijs})/N_i\right) \times (Q/U_s)$, where d is the depth of reads at nucleotide j of feature i, N is the length of the feature, U_s is the upper quartile of $\left(\sum (d_{ijs})/N_i\right)$ values in sample s, and Q is the median of all U_s values within the given species (Bullard, et al. 2010; Dillies, et al. 2013) (supplementary file 5, Supplementary Material online). Distributions of upper quartile values across exonic features were evaluated for each sample mapped to the genome of the sample species (supplementary fig. S8, Supplementary Material online). Median upper quartile values and associated distributions were strikingly similar across all samples in both species except for one *D. simulans* sz12 male replicate, which was removed from further analysis.

For each species separately, differential expression between males and females was evaluated for exonic features detected in both sexes. We used the linear fixed effect model $Y_{xp} = \mu + g_x + \varepsilon_{xp}$, where Y is the log-transformed UQ normalized C_{is} values for the x th sex ($x = male,\ female$), p th replicate ($p = 1, 2, \ldots, 12$). We accounted for potential heteroscedasticity of variance between the sexes (Graze, et al. 2012) and used the Kenward–Roger adjustment for the degrees of freedom (Kenward and Roger 1997). Normality of residuals was tested using the Shapiro-Wilk test (Shapiro and Wilk 1965). Fold-change ratios were calculated for each exonic feature i, $r_i = \left(\sum (f_{ip})/k\right)/\left(\sum (m_{il})/n\right)$, where f_{ij} is the UQ normalized C_{is} for exonic region i in female replicate $p = 1 \ldots k$ total female replicates, and m_{il} is the UQ normalized C_{is} for exonic region i in male replicate $l = 1 \ldots n$ total male replicates. Exonic features were classified as male-biased (or female-biased) if the nominal P-value was less than or equal to 0.05 and the fold-change less than (or greater than) 1.

ChIP-seq Feature Detection

While peak calling is a common method of ChIP-seq analysis, it is highly dependent on the algorithm used and the parameters selected (Yang, et al. 2014), especially for ChIP marks that are predicted to show broad peaks such as certain histone modifications (Park 2009; Pepke, et al. 2009; Dahl, et al. 2016). To have a consistent method for

evaluation and comparison of ChIP results across different marks and between males and females, and to compare ChIP results directly to the RNA-seq results in *cis*, we use ChIP-seq reads to quantify features based on the annotations of the reference genomes (Katz, et al. 2010; Anders, et al. 2012; Zhang, et al. 2012; Yang, et al. 2014; Newman, et al. 2018). By focusing on features rather than MACS2 peaks, many more detections above input control are identified at the feature-level and at the gene-level (supplementary section 7.1, Supplementary Material online for detailed results from MACS2).

For both species, features were quantified as $C_{cis} = \left(\sum (d_{cijs})/N_i\right) \times (M_c/R_{cs})$, where d is the depth of reads at nucleotide j of feature i, N is the length of the feature, M is the median read count, and R is the total read count for ChIP c (H3K4me3, H3K27me2me3, Input control) in sample s (Dillies, et al. 2013). A feature was considered detected above the input control (DAI) for H3K4me3 or H3K27me3me4 if $C_{K4,\,is} > C_{Input,\,is}$, in more than 50% of the replicates for that species-sex combination. A gene was considered as having a mark if at least one exonic feature in the gene was DAI. A gene was considered male-limited (or female-limited) if only male-limited (or female-limited) exonic features were identified in the gene. The agreement between histone marks for males and females, and between H3K4me3 and H3K27me2me3 marks within each sex, was estimated using Cohen's kappa (Cohen 1960; Fleiss 1981) in order to account for marginal frequencies and provide a more accurate assessment of the relationship between sexes and the marks (supplementary fig. S5, Supplementary Material online).

Chromatin and Expression

Histone modifications change the availability of chromatin for transcription (Santos-Rosa, et al. 2002; Schneider, et al. 2004; Wang, et al. 2008; Juan, et al. 2016); therefore, we examined the impact of chromatin marks on expression. When sex-biased expression is observed, this may be due to open marks in the sex with the higher expression, or to closed marks in the other sex. Specifically, if there is male-biased expression, we expect open (H3K4me3) marks in males or closed (H3K27me2me3) marks in females for that gene, and if there is female-biased expression, we expect open (H3K4me3) marks in females or closed (H3K27me2me3) marks in males (fig. 2; supplementary fig. S3, Supplementary Material online). As chromatin marks in males do not influence expression in females, or vice versa, the appropriate statistical comparison is not a test of general association between expression and chromatin marks between the sexes.

For males, the presence/absence of the chromatin marks, H3K4me3 and H3K27me2me3, was compared to presence/absence of gene expression in males and evaluated for agreement using Cohen's kappa coefficients (Cohen 1960; Fleiss 1981) (supplementary table S3, Supplementary Material online). Females were examined separately in the same manner. For genes with detected expression in both sexes, the presence/absence of sex bias in males was compared to the presence/absence of male H3K4me3 marks using Fisher exact test (Fisher 1934). A one-sided test was used with the alternative hypothesis that male open chromatin marks would be more likely in male-biased expression compared to non-male-biased expression. For genes with sex-biased expression in males, the presence/absence of H3K27me2me3 marks in females was tested using Fisher exact test (Fisher 1934). A one-sided test was used with the alternative hypothesis that female closed chromatin marks would be more likely in genes with male-biased expression compared to nonmale-biased expression. Analogous tests were performed for the presence/absence of sex bias in females compared to the presence/absence of female H3K4me3 and presence/absence of male H3K27me2me3 using Fisher exact test (Fisher 1934).

List Enrichment

Genes with sex-biased gene expression conserved between *D. melanogaster* and *D. simulans* in this study were compared to genes identified in previous studies of sex-biased expression in *D. melanogaster* head tissue (Chang, et al. 2011) using Pearson's chi-square (χ^2) test (Pearson 1900). Additionally, conserved male-biased (or female-biased) genes were compared to genes previously identified as male-biased (or female-biased) in *D. melanogaster* head tissue and in *fru-P1*-expressing neurons (Newell, et al. 2016) using Pearson's chi-square (χ^2) test (Pearson 1900). Based on the extensive knowledge of the sex-specifically spliced Drosophila sex determination gene *fru* (Ryner, et al. 1996; Heinrichs, et al. 1998; reviewed in Salvemini, et al. 2010), we expected *fru* to play a role in conserved sex-biased expression. Genes with male-biased and female-biased expression conserved between *D. melanogaster* and *D. simulans* in this study were compared to genes regulated by the Fru^M protein in *D. melanogaster* males (Dalton, et al. 2013) using Pearson's chi-square (χ^2) test (Pearson 1900).

Divergence of the targets of the terminal sex determination genes may contribute to the divergence of sex-biased expression between the species. To evaluate this, species-specific sex-biased genes identified in this study were compared to genes in a study of *dsx* regulation in *dsx* null females and *dsx* pseudomales of *D. melanogaster* (Arbeitman, et al. 2016) and to genes observed to be regulated downstream of *fru* in *D. melanogaster* males (Dalton, et al. 2013) using Pearson's Chi-square (χ^2) test (Pearson 1900). To validate the patterns of open and closed chromatin in males and females, gene-level presence of open (H3K4me3) and closed (H3K27me2me3) chromatin marks in *D. melanogaster* males and females found in this study were compared to previous observations of H3K4me3 and H3K27me3 marks in *D. melanogaster* male and female (*elav*-expressing) neurons (Palmateer, et al. 2021) using Pearson's chi-square (χ^2) test (Pearson 1900). Tests of agreement between these datasets were carried out for

males and females separately using Cohen's kappa coefficients (Cohen 1960; Fleiss 1981) (supplementary table S3, Supplementary Material online).

To evaluate if the patterns of the chromatin marks in the head tissue described here are consistent with patterns of chromatin marks in neurons known to direct male and female reproductive behaviors (Demir and Dickson 2005; Manoli, et al. 2005; Stockinger, et al. 2005; Kvitsiani and Dickson 2006), the genes we detected with open (or closed) chromatin marks were compared to genes with H3K4me3 (or H3K27me3) marks in *D. melanogaster* male and female *fru-P1*-expressing neurons (Palmateer, et al. 2021) using Pearson's chi-square (χ^2) test (Pearson 1900). We also compared genes we detected with male-limited and female-limited open (or closed) chromatin to the genes with H3K4me3 (or H3K27me3) marks in *D. melanogaster* male and female *fru-P1*-expressing neurons (Palmateer, et al. 2021) using Pearson's chi-square (χ^2) test (Pearson 1900). Tests of agreement of the comparable marks between head tissue and *fru*-P1-expressing neurons were also evaluated for males and females separately using Cohen's kappa coefficients (Cohen 1960; Fleiss 1981) (supplementary table S3, Supplementary Material online).

Supplementary material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

We would like to acknowledge the two reviewers for this paper whose insightful suggestions and patience have made a substantial contribution to our understanding, Michelle Arbeitman for helpful discussion and sharing of pre-publication results, HiPerGator high-performance supercomputer at the University of Florida, and The University of Florida Genetics Institute. Funding provided by National Institute of General Medical Sciences R01GM128193 (L.M.M., S.V.N.) and National Institute of Dental and Craniofacial Research R01DE026707 (R.R.).

Data Availability

Raw short-read data from the RNA-seq and ChIP-seq experiments are available under SRA BioProject accession PRJNA737411. RNA-seq and ChIP-seq mapped read count summary (supplementary table S6, Supplementary Material online) and RNA-seq UQ normalization factors (supplementary file 5, Supplementary Material online) are provided in the supplement. Analyzed data are provided as supplementary files for 1) *D. melanogaster* gene-level chromatin and expression variables (supplementary file 1, Supplementary Material online), 2) *D. melanogaster* feature-level level chromatin and expression variables (supplementary file 6, Supplementary Material online), 3) *D. simulans* gene-level chromatin and expression variables (supplementary file 2, Supplementary Material online), 4) *D. simulans* feature-level chromatin and expression variables (supplementary File 7, Supplementary Material online), and 5) *D. melanogaster-D. simulans* orthologous gene chromatin and expression variables (supplementary file 8, Supplementary Material online). Further method details can be found in supplementary materials, Supplementary Material online. Documentation of all analyses and comparisons as well as scripts are on github (https://github.com/McIntyre-Lab/papers/tree/master/nanni_{PI}{PI}chip_rna_2023).

References

Anders S, Reyes A, Huber W. 2012. Detecting differential usage of exons from RNA-Seq data. *Genome Res.* **22**:2008–2017.

Anholt RRH, O'Grady P, Wolfner MF, Harbison ST. 2020. Evolution of reproductive behavior. *Genetics*. **214**:49–73.

Arbeitman MN, New FN, Fear JM, Howard TS, Dalton JE, Graze RM. 2016. Sex differences in Drosophila somatic gene expression: variation and regulation by doublesex. *G3-Genes Genomes Genetics*. **6**:1799–1808.

Bonduriansky R, Chenoweth SF. 2009. Intralocus sexual conflict. *Trends Ecol Evol (Amst)*. **24**:280–288.

Boros I. 2012. Histone modification in Drosophila. *Brief Funct Genomics*. **11**:319–331.

Brown EJ, Bachtrog D. 2014. The chromatin landscape of *Drosophila*: comparisons between species, sexes, and chromosomes. *Genome Res.* **24**:1125–1137.

Bullard JH, Purdom E, Hansen KD, Dudoit S. 2010. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics*. **11**:94.

Bushnell B, Rood J, Singer E. 2017. BBMerge—accurate paired shotgun read merging via overlap. *Plos One*. **12**:e0185056.

Chang P, Dunham J, Nuzhdin S, Arbeitman M. 2011. Somatic sex-specific transcriptome differences in Drosophila revealed by whole transcriptome sequencing. *Bmc Genomics*. **12**:364.

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

Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman TC, Kellis M, Gelbart W, Iyer VN, *et al.* 2007. Evolution of genes and genomes on the Drosophila phylogeny. *Nature*. **450**:203–218.

Cohen J. 1960. A coefficient of agreement of nominal scales. *Educ Psychol Meas.* **20**:37–46.

Cox RM, Calsbeek R. 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. *American Naturalist*. **173**:176–187.

Dahl JA, Jung I, Aanes H, Greggains GD, Manaf A, Lerdrup M, Li GQ, Kuan S, Li B, Lee AY, *et al.* 2016. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature*. **537**:548–552.

Dalton JE, Fear JM, Knott S, Baker BS, McIntyre LM, Arbeitman MN. 2013. Male-specific fruitless isoforms have different regulatory roles conferred by distinct zinc finger DNA binding domains. *Bmc Genomics*. **14**:659.

Demir E, Dickson BJ. 2005. Fruitless splicing specifies male courtship behavior in Drosophila. *Cell*. **121**:785–794.

Dillies MA, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, Keime C, Marot G, Castel D, Estelle J, *et al.* 2013. A comprehensive evaluation of normalization methods for illumina high-throughput RNA sequencing data analysis. *Brief Bioinform*. **14**:671–683.

Ding Y, Lilivis JL, Cande J, Berman GJ, Arthur BJ, Long X, Xu M, Dickson BJ, Stern DL. 2019. Neural evolution of context-dependent fly song. *Curr Biol.* **29**:1089–1099.e7.

Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nature Reviews Genetics*. **8**: 689–698.
Fisher RA. 1934. *Statistical methods for research workers*. Edinburgh: Oliver and Boyd.
Fleiss JL. 1981. *Statistical methods for rates and proportions*. 2nd ed. New York: Wiley, John and Sons, Inc.
Goldman TD, Arbeitman MN. 2007. Genomic and functional studies of Drosophila sex hierarchy regulated gene expression in adult head and nervous system tissues. *PLoS Genet.* **3**:e216.
Graze RM, Novelo LL, Amin V, Fear JM, Casella G, Nuzhdin SV, McIntyre LM. 2012. Allelic imbalance in Drosophila hybrid heads: exons, isoforms, and evolution. *Mol Biol Evol*. **29**: 1521–1532.
Harrison PW, Wright AE, Zimmer F, Dean R, Montgomery SH, Pointer MA, Mank JE. 2015. Sexual selection drives evolution and rapid turnover of male gene expression. *Proc Natl Acad Sci USA*. **112**:4393–4398.
Heinrichs V, Ryner LC, Baker BS. 1998. Regulation of sex-specific selection of *fruitless* 5 'splice sites by *transformer* and *transformer-2*. *Mol Cell Biol.* **18**:450–458.
Innocenti P, Morrow E. 2010. The sexually antagonistic genes of Drosophila melanogaster. *PLoS Biol.* **8**:e1000335.
Ito H, Fujitani K, Usui K, ShimizuNishikawa K, Tanaka S, Yamamoto D. 1996. Sexual orientation in Drosophila is altered by the satori mutation in the sex-determination gene fruitless that encodes a zinc finger protein with a BTB domain. *Proc Natl Acad Sci USA*. **93**:9687–9692.
Ito H, Sato K, Koganezawa M, Ote M, Matsumoto K, Hama C, Yamamoto D. 2012. Fruitless recruits two antagonistic chromatin factors to establish single-neuron sexual dimorphism. *Cell*. **149**:1327–1338.
Jiang L, Schlesinger F, Davis CA, Zhang Y, Li R, Salit M, Gingeras TR, Oliver B. 2011. Synthetic spike-in standards for RNA-Seq experiments. *Genome Res*. **21**:1543–1551.
Juan AH, Wang S, Ko KD, Zare H, Tsai PF, Feng XS, Vivanco KO, Ascoli AM, Gutierrez-Cruz G, Krebs J, *et al.* 2016. Roles of H3K27me2 and H3K27me3 examined during fate specification of embryonic stem cells. *Cell Rep.* **17**:1369–1382.
Katz Y, Wang ET, Airoldi EM, Burge CB. 2010. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods*. **7**:1009–1015.
Kenward MG, Roger JH. 1997. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*. **53**:983–997.
Kido A, Ito K. 2002. Mushroom bodies are not required for courtship behavior by normal and sexually mosaic Drosophila. *J Neurobiol*. **52**:302–311.
Kimura H. 2013. Histone modifications for human epigenome analysis. *J Hum Genet.* **58**:439–445.
Kimura K, Hachiya T, Koganezawa M, Tazawa T, Yamamoto D. 2008. Fruitless and doublesex coordinate to generate male-specific neurons that can initiate courtship. *Neuron*. **59**:759–769.
Kopp A, Barmina O, Hamilton AM, Higgins L, McIntyre LM, Jones CD. 2008. Evolution of gene expression in the Drosophila olfactory system. *Mol Biol Evol.* **25**:1081–1092.
Kvitsiani D, Dickson BJ. 2006. Shared neural circuitry for female and male sexual behaviours in Drosophila. *Curr Biol.* **16**:R355–R356.
Landry CR, Castillo-Davis CI, Ogura A, Liu JS, Hartl DL. 2007. Systems-level analysis and evolution of the phototransduction network in *Drosophila*. *Proc Natl Acad Sci USA*. **104**:3283–3288.
Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
Liang HQ, Katoh T, Sato K, Yamamoto D, Wen SY. 2021. Evolution of a neuromuscular sexual dimorphism in the Drosophila montium species group. *Sci Rep.* **11**(1):1–16.
Lorbeck MT, Singh N, Zervos A, Dhatta M, Lapchenko M, Yang C, Elefant F. 2010. The histone demethylase dmel\Kdm4A controls genes required for life span and male-specific sex determination in Drosophila. *Gene*. **450**:8–17.
Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, Casillas S, Han Y, Magwire MM, Cridland JM, *et al.* 2012. The Drosophila melanogaster genetic reference panel. *Nature*. **482**: 173–178.
Manoli DS, Foss M, Villella A, Taylor BJ, Hall JC, Baker BS. 2005. Male-specific fruitless specifies the neural substrates of Drosophila courtship behaviour. *Nature*. **436**:395–400.
Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal*. **17**(1): 10–12.
Newell NR, New FN, Dalton JE, McIntyre LM, Arbeitman MN. 2016. Neurons that underlie *Drosophila melanogaster* reproductive behaviors: detection of a large male-bias in gene expression in *fruitless*-expressing neurons. *G3 Genes|Genomes|Genetics*. **6**: 2455–2465.
Newman J, Concannon P, Tardaguila M, Conesa A, McIntyre L. 2018. Event analysis: using transcript events to improve estimates of abundance in RNA-Seq data. *G3-Genes Genomes Genetics*. **8**: 2923–2940.
Palmateer CM, Moseley SC, Ray S, Brovero SG, Arbeitman MN. 2021. Analysis of cell-type-specific chromatin modifications and gene expression in Drosophila neurons that direct reproductive behavior. *PLoS Genet.* **17**:e100924.
Park PJ. 2009. ChIP-seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics*. **10**:669–680.
Pearson K. 1900. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine*. **50**:157–175.
Pepke S, Wold B, Mortazavi A. 2009. Computation for ChIP-seq and RNA-Seq studies. *Nat Methods*. **6**:S22–S32.
Rice WR, Holland B. 1997. The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific red queen. *Behav Ecol Sociobiol (Print).* **41**:1–10.
Rideout EJ, Billeter JC, Goodwin SF. 2007. The sex-determination genes fruitless and doublesex specify a neural substrate required for courtship song. *Curr Biol.* **17**:1473–1478.
Rideout EJ, Dornan AJ, Neville MC, Eadie S, Goodwin SF. 2010. Control of sexual differentiation and behavior by the doublesex gene in *Drosophila melanogaster*. *Nat Neurosci.* **13**:458–466.
Ryner LC, Goodwin SF, Castrillon DH, Anand A, Villella A, Baker BS, Hall JC, Taylor BJ, Wasserman SA. 1996. Control of male sexual behavior and sexual orientation in Drosophila by the fruitless gene. *Cell*. **87**:1079–1089.
Salvemini M, Polito C, Saccone G. 2010. Fruitless alternative splicing and sex behaviour in insects: an ancient and unforgettable love story? *J Genet.* **89**:287–299.
Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NCT, Schreiber SL, Mellor J, Kouzarides T. 2002. Active genes are tri-methylated at K4 of histone H3. *Nature*. **419**: 407–411.
Sato K, Tanaka R, Ishikawa Y, Yamamoto D. 2020. Behavioral evolution of Drosophila: unraveling the circuit basis. *Genes (Basel).* **11**:157.
Schenkel MA, Pen I, Beukeboom LW, Billeter JC. 2018. Making sense of intralocus and interlocus sexual conflict. *Ecol Evol.* **8**: 13035–13050.
Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. 2004. Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol.* **6**:73–77.
Shapiro SS, Wilk MB. 1965. An analysis of variance test for normality (complete samples). *Biometrika*. **52**:591–611.
Shiao MS, Chang JM, Fan WL, Lu MYJ, Notredame C, Fang S, Kondo R, Li WH. 2015. Expression divergence of chemosensory genes between *Drosophila sechellia* and its sibling Species and its implications for host shift. *Genome Biol Evol.* **7**:2843–2858.
Shukla JN, Nagaraju J. 2010. Doublesex: a conserved downstream gene controlled by diverse upstream regulators. *J Genet.* **89**:341–356.
Signor S. 2017. Population genomics of Wolbachia and mtDNA in Drosophila simulans from California. *Sci Rep*. **7**:13369.

Stanley CE, Kulathinal RJ. 2016. flyDIVaS: a comparative genomics resource for Drosophila divergence and selection. *G3-Genes Genomes Genetics*. **6**:2355–2363.

Stockinger P, Kvitsiani D, Rotkopf S, Tirian L, Dickson BJ. 2005. Neural circuitry that governs Drosophila male courtship behavior. *Cell*. **121**:795–807.

Swanson WJ, Nielsen R, Yang Q. 2003. Pervasive adaptive evolution in mammalian fertilization proteins. *Mol Biol Evol*. **20**:18–20.

Tanaka R, Higuchi T, Kohatsu S, Sato K, Yamamoto D. 2017. Optogenetic activation of the *fruitless*-labeled circuitry in *Drosophila subobscura* males induces mating motor acts. *Journal of Neuroscience*. **37**:11662–11674.

Tregenza T, Wedell N, Chapman T. 2006. Introduction. Sexual conflict: a new paradigm? *Philosophical Transactions of the Royal Society B-Biological Sciences*. **361**:229–234.

von Philipsborn AC, Liu TX, Yu JY, Masser C, Bidaye SS, Dickson BJ. 2011. Neuronal control of Drosophila courtship song. *Neuron*. **69**:509–522.

Wang ZB, Zang CZ, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui KR, Roh TY, Peng WQ, Zhang MQ, *et al.* 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet.* **40**:897–903.

Waterhouse R, Tegenfeldt F, Li J, Zdobnov E, Kriventseva E. 2013. OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic Acids Res.* **41**:D358–D365.

Wong WS, Yang Z, Goldman N, Nielsen R. 2004. Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics*. **168**:1041–1051.

Yamamoto D, Koganezawa M. 2013. Genes and circuits of courtship behaviour in Drosophila males. *Nature Reviews Neuroscience*. **14**:681–692.

Yang Y, Fear J, Hu J, Haecker I, Zhou L, Renne R, Bloom D, McIntyre LM. 2014. Leveraging biological replicates to improve analysis in ChIP-seq experiments. *Comput Struct Biotechnol J.* **9**:e201401002.

Yang ZH. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci.* **13**:555–556.

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

Zhang YJ, Lameijer EW, t'Hoen PAC, Ning ZM, Slagboom PE, Ye K. 2012. PASSion: a pattern growth algorithm-based pipeline for splice junction detection in paired-end RNA-Seq data. *Bioinformatics*. **28**:479–486.