Parallel Faster-X Evolution of Gene Expression and Protein Sequences in *Drosophila*: Beyond Differences in Expression Properties and Protein Interactions

Ana Llopart[1,2]*

1 Department of Biology, The University of Iowa, Iowa City, Iowa, United States of America,
2 Interdisciplinary Graduate Program in Genetics, The University of Iowa, Iowa City, Iowa, United States of America

* ana-llopart@uiowa.edu

Abstract

Population genetics models predict that the *X* (or *Z*) chromosome will evolve at faster rates than the autosomes in *XY* (or *ZW*) systems. Studies of molecular evolution using large data-sets in multiple species have provided evidence supporting this faster-X effect in protein-coding sequences and, more recently, in transcriptomes. However, X-linked and autosomal genes differ significantly in important properties besides hemizygosity in males, including gene expression levels, tissue specificity in gene expression, and the number of interactions in which they are involved (i.e., protein-protein or DNA-protein interactions). Most important, these properties are known to correlate with rates of evolution, which raises the question of whether differences between the *X* chromosome and autosomes in gene properties, rather than hemizygosity, are sufficient to explain faster-X evolution. Here I investigate this possibility using whole genome sequences and transcriptomes of *Drosophila yakuba* and *D. santomea* and show that this is not the case. Additional factors are needed to account for faster-X evolution of both gene expression and protein-coding sequences beyond differences in gene properties, likely a higher incidence of positive selection in combination with the accumulation of weakly deleterious mutations.

Introduction

Population genetics theory predicts that under certain conditions the *X* chromosome will evolve at faster rates than the autosomes in *XY* (or *ZW*) systems, an effect known as faster-X [1,2]. In a population of diploids, newly arisen autosomal mutations are mostly found in heterozygotes and, if recessive for fitness, their effects are masked by the ancestral variants. In contrast, if these mutations arise on the *X* chromosome and are beneficial, natural selection can drive them to fixation more efficiently because their fitness effects will be fully exposed in hemizygous males. As a result, the *X* chromosome is expected to accumulate adaptive

mutations at a faster rate than the autosomes. This difference is expected to be more extreme for mutations that affect only males [3].

Yet in *XY* (or *ZW*) systems faster-X evolution can also result from new mutations that are weakly deleterious [1,2,4,5]. If a population contains equal numbers of breeding females and males, the effective population size of the *X* chromosome (N_{eX}) is expected to be ¾ the effective population size of the autosomes (N_{eA}) (all else being equal). This reduced N_{eX} results in increased genetic drift and, as a consequence, weakly deleterious mutations that are partially dominant attain higher frequencies in the population and reach fixation more often when X-linked than when autosomal, thus producing faster-X. Note that strongly deleterious mutations are eliminated from the population by natural selection before they can reach fixation regardless of whether they are X-linked or autosomal.

The likelihood of faster-X being driven by either beneficial or weakly deleterious mutations is strongly influenced by the N_{eX} to N_{eA} ratio [4,5]. Taxa with small N_{eX}/N_{eA}, such as several species of birds, are expected to be more susceptible to faster-X associated with the fixation of weakly deleterious mutations [6], although positive selection has also been invoked [7,8]. Cases where faster-X is due to weakly deleterious mutations should be also accompanied by a reduction of neutral polymorphism on the *X* chromosome due to the smaller N_{eX}. On the other hand, the observation of faster-X in species with N_{eX}/N_{eA} close to, or slightly above, 1 is more easily explained by a higher incidence of adaptive evolution on the *X* chromosome [4,5]. The N_{eX} to N_{eA} ratio is influenced by multiple factors besides sex ratios, including mating system, differences between the two sexes in the strength of sexual selection, or disparities in recombination rates and linkage effects between the *X* chromosome and autosomes [4,9–13].

Regardless of the evolutionary forces underlying faster-X (selection on beneficial or weakly deleterious mutations), the evidence for this effect was initially controversial. In *Drosophila* early attempts to evaluate the faster-X evolution of protein-coding sequences produced inconsistent results [14–19] but more recent studies did find clear statistical support [20–31]. Importantly the strongest evidence for faster-X is observed in male-biased genes (i.e., genes expressed at higher levels in males than in females, MBGs) [20,26,29]. To date, accelerated rates of evolution for X-linked proteins have been documented in a wide range of taxa, including primates [32–36], mice [37,38], birds [6–8], snakes [39], worms [40], silkmoths [41] and aphids [42], which are an X0 system. Faster-X evolution of gene expression has been shown in primates [35] and more recently in *Drosophila* [43–45].

One limitation of all studies of faster-X is that X-linked and autosomal genes may not be directly comparable [46,47]. MBGs have been proposed to be underrepresented on the *Drosophila X* chromosome [48–50]. Recent studies show that the lack of dosage compensation in the *Drosophila* male germline is an important factor explaining the apparent deficit of X-linked MBGs [51,52]. Levels of gene expression and tissue specificity are significantly different in X-linked than in autosomal genes of *D. melanogaster* [47,52–56]. Transcriptional networks also show unique characteristics for genes on the *X* chromosome [57,58]. Most important, variation in these properties correlates significantly with rates of evolution [46,59–65].

In a comprehensive study of the 12 *Drosophila* genomes, Larracuente *et al.* (2008) showed that gene expression levels were negatively correlated with evolutionary rate, and that tissue specificity in gene expression was independently positively correlated with the rate of evolution [59]. Protein interactions and protein divergence also showed a negative association in *Drosophila* and yeast [59,64,65]. These observations raise the question of whether differences between X-linked and autosomal genes in expression properties (i.e., levels and tissue specificity) and protein interactions, rather than hemizygosity, could entirely explain faster-X. To evaluate this possibility, I examined genomewide gene expression divergence between *Drosophila yakuba* and *D. santomea* and whole-genome patterns of protein-coding sequence divergence

based on a newly obtained *D. santomea* genome sequence. *D. santomea* started diverging allopatrically from its common ancestor with *D. yakuba* ~400,000 years ago [66,67] on the African island of São Tomé where today the two species form a hybrid zone [68]. I show that gene expression properties are good predictors of both gene expression and protein-coding sequence divergence between species, but they are not sufficient to explain faster-X for MBGs. Additional evolutionary factors are thus necessary to account for the rapid evolution of MBGs on the *X* chromosome in this system.

Materials and Methods

Generating a *D. santomea* genome sequence

To generate a high quality, high coverage sequence of the *D. santomea* genome, I obtained Illumina reads from 5 different strains (STO.4, STO.18, QUIJA 650.1, COST1235.1 and CAR1600.3; details on the strains are reported in [67]) according to the manufacturer guidelines ("Preparing Samples for Sequencing Genomic DNA", Part # 1003806 Rev. B, Illumina, San Diego, CA). High molecular weight DNA was isolated from 30 flies following protocol 47 in Ashburner (1989) with minor modifications [69]. Cluster generation and sequencing (75 to 101-bp single-end reads) were carried out in Illumina GAII/HiSeq2000 instruments at the Iowa State DNA Facility (Iowa State University, Ames, IA). Trimming and filtering were done using SAMTools v1.4 [70] and custom scripts.

To obtain a *D. santomea* syntenic sequence, 208.1 million filtered reads were aligned to the *D. yakuba* genome project sequence (dyak_r1.3_FB2011_08; http://flybase.org/) [71] using BWA [72]. SAMTools was used to generate pileup and fasta files for each chromosome arm from BAM files. Only nucleotides with coverage greater than 7× and phred-scaled base consensus (or variant) quality greater than $Q = 40$ were included in the final *D. santomea* synthetic and syntenic sequence. The average read depth of this high quality sequence for *D. santomea* is 145.9× with an average phred-scaled quality of $Q = 242$, covering 91.8% of the *D. yakuba* reference genome [71]. To extract the *D. santomea* coding sequences of all genes, I used Galaxy [73–75] and the exon-intron annotation of *D. yakuba* (dyak_r1.3_FB2011_08; http://flybase.org/). All *D. santomea* reads have been deposited in the NCBI Short Read Archive (SRA) Bio-Project SRP049565.

Evolutionary Analysis

To estimate gene expression levels, I used data on transcript abundance from Llopart (2012) [43]. Briefly, gene expression was measured in 1-day old *D. yakuba* and *D. santomea* adult males using whole-genome, custom-designed *D. yakuba* arrays (11,530 genes analyzed; 60-mer probes; 10 probes/gene). For any given gene, expression level was calculated as the average signal (in Log$_2$) of the two species. Gene expression divergence between *D. yakuba* and *D. santomea* and gene expression polymorphism were obtained from [43]. To gauge tissue specificity in gene expression, I calculated the tissue specificity index (τ), which takes into account the number of tissues in which a gene is expressed and the relative expression levels in each tissue [76]. Genes expressed in many different tissues tend to show low τ values, and genes expressed in a narrow number of tissues, high τ values. Information on gene expression for different tissues was inferred from the analysis of all 25 tissues included in *D. melanogaster* FlyAtlas (http://www.flyatlas.org/) [77]. These tissues are brain, head, eye, thoracicoabdominal ganglion, salivary gland, crop, midgut, tubule, hindgut, heart, fat body, ovary, testis, male accessory glands, virgin spermatheca, mated spermatheca, adult carcass, larval CNS, larval salivary gland, larval midgut, larval tubule, larval hindgut, larval fat body, larval trachea and larval carcass. To determine the number of physical interactions, I used information derived from the strict gene

interaction network of *D. melanogaster* proposed by Hansen and Kulathinal (2013) [58]. This is an empirically driven network of physical interactions based on data derived from six experiments, including yeast-two-hybrid and transcription-factor CHIP-seq analyses from the mod-ENCODE project; the network contains 12,749 genes and 486,138 interactions [58].

To estimate protein-coding sequence divergence, I calculated the number of nonsynonymous (d_N) and synonymous (d_S) substitutions per site between *D. yakuba* and *D. santomea* using the codeml program implemented in PAML v4.5 [78,79]. The equilibrium codon frequencies were calculated from average nucleotide frequencies at third codon positions (codon substitution model F3×4). A single d_N/d_S ratio (ω) was assumed for all lineages and across sites (Model 0). There were 16,082 annotated genes in the *D. yakuba* reference sequence (dyak_r1.3_FB2011_08; http://flybase.org/) [71] and, after excluding genes located on either heterochromatic regions or in unknown chromosomal locations, I applied codeml to 14,682 genes. In 316 genes I found stop codons in the *D. santomea* sequence, which are likely due to either inaccuracy in the current annotation of *D. yakuba*, sequencing errors, or differences between species in protein length or exon-intron structure. There were three *D. yakuba* genes with no orthologous sequences recovered in *D. santomea*. Finally I also excluded from the analysis 350 genes with 0 synonymous changes between *D. yakuba* and *D. santomea* and 354 genes with $\omega \geq 9$, as they tend to generate unreliable estimates of divergence. I obtained d_N, d_S and ω in two sets of genes. The first set contains all 13,659 genes sequenced in both species, without stop codons in *D. yakuba* and *D. santomea*, and with reliable estimates of divergence. The second set of genes analyzed contains the 9,203 out of the 13,659 genes for which there was information on gene expression level [43], tissue specificity in gene expression, and the number of protein interactions.

Results and Discussion

Faster-X evolution of male-biased gene expression

A previous study reported faster-X evolution of male-biased gene expression in *D. yakuba* and *D. santomea* [43]. The same significant trend is detected in the current dataset of 9,697 genes with known gene expression levels, tissue specificity, and the number of protein interactions (1,537 X-linked and 8,160 autosomal) (Mann-Whitney *U* test, $P = 2.5\times10^{-8}$; see S1 Fig). To determine whether these properties are significantly associated with gene expression divergence in *D. yakuba* and *D. santomea*, I examined their distributions across the genome. As expected, expression levels and the number of protein interactions are negatively correlated with gene expression divergence (Spearman's rank correlation ρ = -0.25 for levels and ρ = -0.19 for protein interactions, both $P < 1\times10^{-6}$) while tissue specificity shows a positive correlation (Spearman's rank correlation ρ = 0.23, $P < 1\times10^{-6}$). Similar associations have been reported in other *Drosophila* species [59] and references therein,[64,65]. Most important, X-linked and autosomal genes differ in these properties.

X-linked genes are expressed at significantly lower levels in adult males, show less tissue specificity (i.e., expressed in more tissues), and are involved in more protein interactions than autosomal genes (Fig. 1). Based on these differences and genomewide correlations, tissue specificity and protein interactions may be in fact masking faster-X in this dataset. In contrast, lower expression in X-linked than in autosomal genes could potentially explain faster-X evolution. Altogether these results underline the need to control for differences between the *X* chromosome and autosomes in expression properties (i.e., levels and tissue specificity) and protein interactions when assessing faster-X.

To evaluate whether the rapid evolution of X-linked male-biased gene expression is detectable after controlling for differences between the *X* chromosome and autosomes in levels of
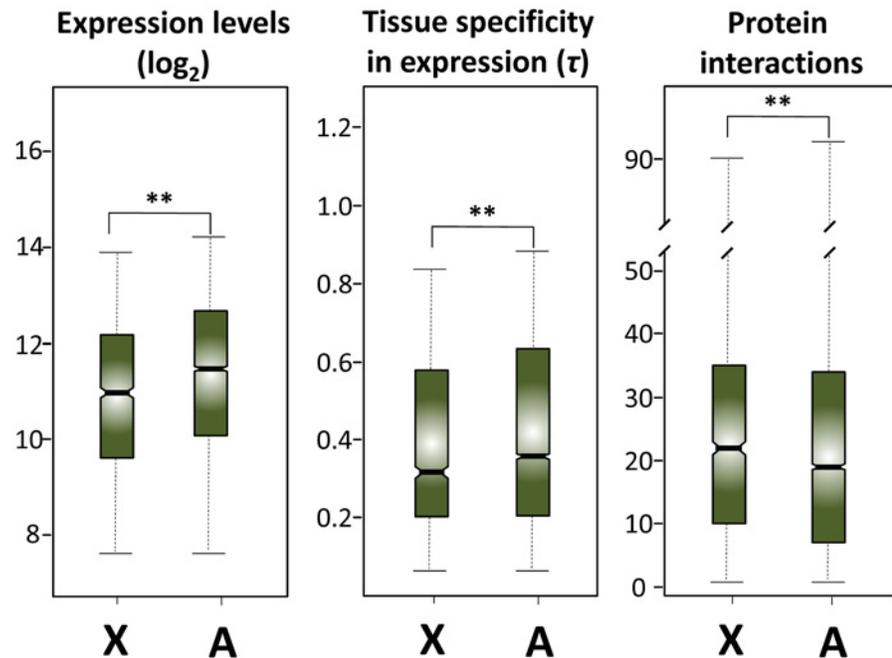
Fig 1. Comparisons of gene expression levels, tissue specificity in gene expression, and number of protein interactions between X-linked (*X*) and autosomal (*A*) genes. The horizontal line inside each box indicates the median. The length of the box and the whiskers represent 50% and 90% confidence intervals, respectively. Asterisks indicate statistically significant differences (Mann-Whitney *U* test; $P < 1 \times 10^{-12}$ for expression levels, $P = 1.6 \times 10^{-4}$ for tissue specificity, and $P = 3 \times 10^{-4}$ for protein interactions).

doi:10.1371/journal.pone.0116829.g001

gene expression, tissue specificity and the number of protein interactions, I performed multiple-regression analysis (Table 1). Residuals of gene expression divergence are significantly greater for X-linked than for autosomal MBGs; a pattern not detected in nonsex-biased genes (NBGs) suggesting a faster-X effect beyond differences in gene properties (Mann-Whitney *U* test, $P = 0.002$ for MBGs and Mann-Whitney *U* test, $P = 0.08$ for NBGs). In addition, I carried out an analysis that is free of the assumptions associated with standard linear regression models and enables comparisons of gene expression divergence between subsets of genes with similar expression properties and number of protein interactions. Expression levels, tissue specificity, and the number of protein interactions were each subdivided into 10 equal-size categories and all MBGs were assigned to their corresponding combined (3-dimensions) category. For each category, I then obtained mean expression divergence for X-linked and autosomal MBGs separately. Finally, I applied a nonparametric test to determine whether there was an

Table 1. Parameters of multiple-regression models and partial correlations for gene expression divergence.

	R (*P*)	*b*Lev* (*P*)	*b*Spe* (*P*)	*b*Int* (*P*)
MBGs	0.33 ($<1 \times 10^{-12}$)	-0.30 ($<1 \times 10^{-12}$)	0.07 (0.003)	-0.04 (0.043)
NBGs	0.35 ($<1 \times 10^{-12}$)	-0.25 ($<1 \times 10^{-12}$)	0.17 ($<1 \times 10^{-12}$)	-0.03 (0.032)
All	0.39 ($<1 \times 10^{-12}$)	-0.28 ($<1 \times 10^{-12}$)	0.19 ($<1 \times 10^{-12}$)	-0.027 (0.004)

Note: *R*, correlation coefficient; *P*, Probability; Lev, average gene expression levels; Spe, tissue specificity in gene expression estimated according to Yanai *et al*. (2005) [76]; Int, protein interactions identified from Hansen and Kulathinal (2012) [58]; MBGs, male-biased genes; NBGs, nonsex-biased genes; All, all genes.

doi:10.1371/journal.pone.0116829.t001

overrepresentation of categories in which X-linked MBGs show higher gene expression divergence than autosomal MBGs. (Only categories that contained both X-linked and autosomal MBGs were considered.) This analysis shows that X-linked MBGs have consistently higher gene expression divergence than autosomal MBGs (Wilcoxon matched pairs test, $T = 4772.0$, $P = 0.0032$) when sharing expression properties and number of protein interactions. Results from both types of analyses imply that the observed excess of gene expression divergence in MBGs on the X chromosome remains significant after taking into account differences among genes in properties. Note that while estimates of expression levels are based directly on transcript abundance in *D. yakuba* and *D. santomea*, tissue specificity and protein interactions are inferred from the close relative *D. melanogaster*. However potential errors in this inference could not explain our results, as they are not expected to impact differentially X-linked and autosomal genes.

The excess of gene expression divergence for X-linked MBGs could be the result of either positive selection or increased genetic drift associated with $N_{eX} < N_{eA}$ and a higher rate of accumulation of weakly deleterious mutations. The analysis of silent nucleotide polymorphism in 26 nuclear regions sequenced in *D. yakuba* and *D. santomea* indicated that the N_{eX} to N_{eA} ratio is ~0.63 [67], which opens the possibility that faster-X evolution of gene expression could be the result of increased genetic drift on the X chromosome. To test this possibility, I investigated gene expression polymorphism in both MBGs and NBGs on the X chromosome [43]. X-linked NBGs are more polymorphic than X-linked MBGs (median gene expression polymorphism is 0.21 for NBGs and 0.16 for MBGs; Mann-Whitney U test, $P = 0.0002$), even though the latter show increased tissue specificity and are involved in fewer protein interactions ([Fig. 2](#)). Indeed this difference remains significant after correcting for variation in expression properties and protein interactions on the X chromosome using multiple-regression analysis (Mann-Whitney U test, $P = 7.3 \times 10^{-6}$ in analysis of residuals; [Table 2](#)). Although the precise relationship between nucleotide changes and variation in gene expression is not yet fully understood, the contrasting patterns of gene expression divergence and polymorphism in MBGs and NBGs suggest that a
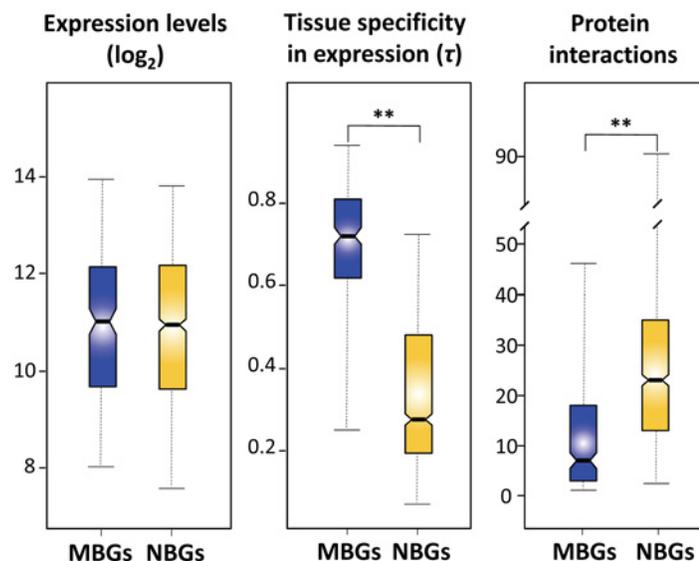
Fig 2. Comparisons of gene expression levels, tissue specificity in gene expression, and number of protein interactions between X-linked MBGs (blue) and X-linked NBGs (yellow). Asterisks indicate statistically significant differences (Mann-Whitney U test; $P < 1 \times 10^{-32}$ for tissue specificity and $P = 7.3 \times 10^{-29}$ for protein interactions). (See [Fig. 1](#) legend for explanation of box plots.)

doi:10.1371/journal.pone.0116829.g002

Table 2. Parameters of multiple-regression models and partial correlations for gene expression polymorphism.

	R (P)	b*Lev (P)	b*Spe (P)	b*Int (P)
X	0.38 ($<1\times10^{-12}$)	-0.37 ($<1\times10^{-12}$)	0.042 (0.094)	-0.001 (0.97)
A	0.40 ($<1\times10^{-12}$)	-0.32 ($<1\times10^{-12}$)	0.16 ($<1\times10^{-12}$)	-0.02 (0.1)

Note: R, correlation coefficient; P, Probability; Lev, average gene expression levels; Spe, tissue specificity in gene expression estimated following Yanai et al. (2005) [76]; Int, protein interactions identified from Hansen and Kulathinal (2012) [58]; X, X-linked genes; A, autosomal genes.

doi:10.1371/journal.pone.0116829.t002

fraction of beneficial mutations contributes to faster-X evolution of male-biased gene expression in *D. yakuba* and *D. santomea* (Fig. 3). However, weakly deleterious mutations may still play an additional role in faster-X, as a higher incidence of positive selection on the *X* chromosome than on autosomes is expected to reduce N_{eX} due to linkage effects [80–82] and result in an associated increased fixation of weakly deleterious mutations.

Faster-X evolution of protein-coding sequences in male-biased genes

I obtained a high-coverage, high-quality sequence of the *D. santomea* genome that covers >97% of the coding sites annotated in *D. yakuba* (see Materials and Methods for details). To determine whether there is faster-X evolution of protein-coding sequences in *D. yakuba* and *D. santomea*, I examined d_N and d_S in 9,203 genes with information on expression properties and protein interactions (1,480 X-linked and 7,723 autosomal) [43,77]. When all genes are analyzed together, there is no evidence for faster-X in comparisons of d_N (median d_N = 0.003 for X-linked genes and 0.003 for autosomal genes; Mann-Whitney U test, P = 0.35; Fig. 4) or d_N/d_S (ω) (median ω = 0.074 for X-linked genes and 0.070 for autosomal genes; Mann-Whitney U
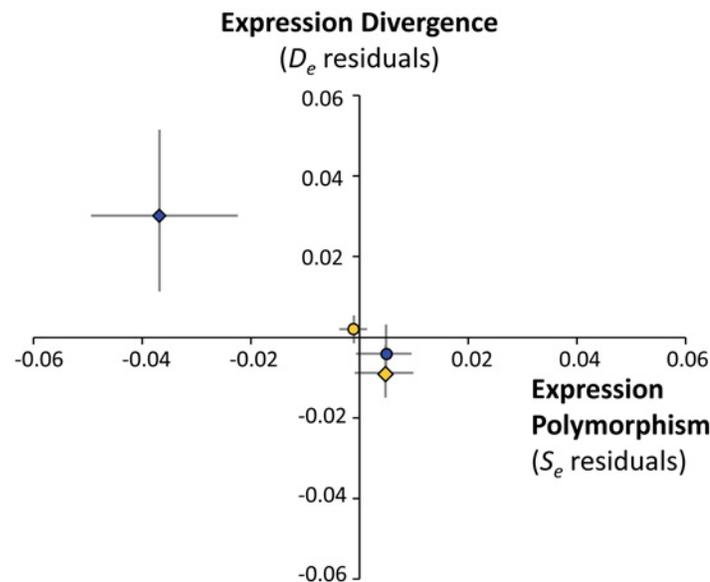
Fig 3. Multiple-regression residuals of gene expression divergence and polymorphism (mean) for X-linked male-biased genes (MBGs) (blue diamonds), autosomal MBGs (blue circles), X-linked nonsex-biased genes (NBGs) (yellow diamonds), and autosomal NBGs (yellow circles). Grey lines represent 90% confidence limits of the residual means.

doi:10.1371/journal.pone.0116829.g003
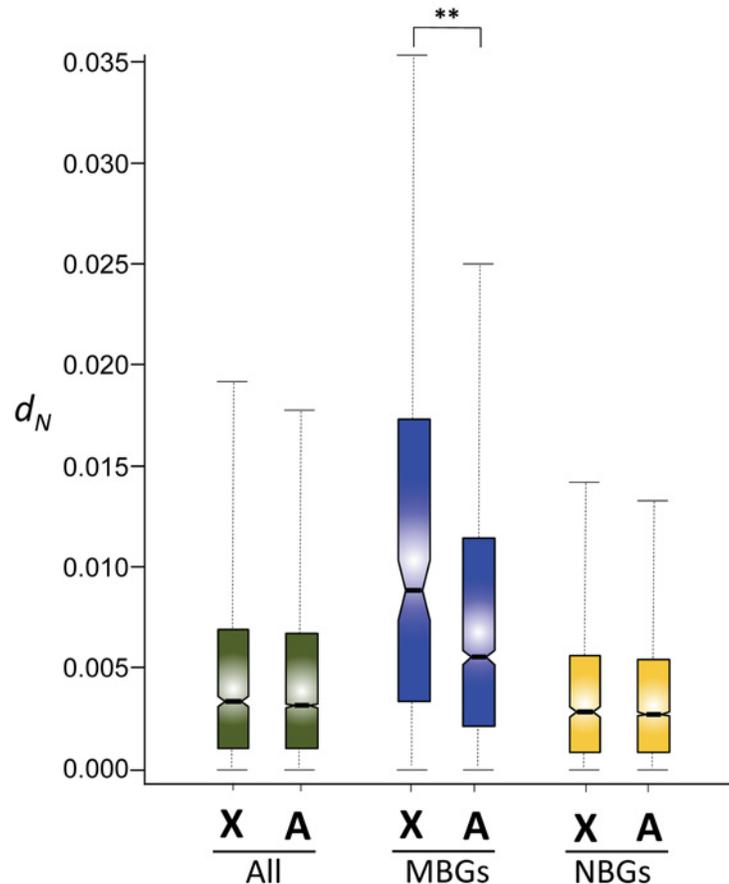
Fig 4. Number of nonsynonymous changes per site (d_N) between *D. yakuba* and *D. santomea* in 9,203 genes (green), male-biased genes (blue, MBGs) and nonsex-biased genes (yellow, NBGs). *X*, X-linked genes; *A*, autosomal genes. Asterisks indicate statistically significant differences in d_N (Mann-Whitney *U* test, $P = 1 \times 10^{-6}$). (See Fig. 1 legend for explanation of box plots.).

doi:10.1371/journal.pone.0116829.g004

test, $P = 0.63$). To confirm that these trends are representative of whole genome patterns, I also analyzed a larger set of 13,659 genes sequenced in both species (2,335 X-linked and 11,424 autosomal; see Materials and Methods for details), which produced the same non-significant results for d_N (median $d_N = 0.0039$ for X-linked genes and 0.0038 for autosomal genes; Mann-Whitney *U* test, $P = 0.76$) and marginally significant results for ω (median $\omega = 0.092$ for X-linked genes and 0.086 for autosomal genes; Mann-Whitney *U* test, $P = 0.026$). The only consistent difference between X-linked and autosomal genes is detected in d_S, which is significantly smaller on the *X* chromosome in both gene sets (9,203 genes, median $d_S = 0.043$ for X-linked genes and 0.045 for autosomal genes; Mann-Whitney *U* test, $P = 1.3 \times 10^{-4}$; 13,659 genes, median $d_S = 0.041$ for X-linked genes and 0.043 for autosomal genes; Mann-Whitney *U* test, $P < 2.8 \times 10^{-8}$). This observation is consistent with previous findings of stronger codon bias on the *X* chromosome in other *Drosophila* species, reflecting higher intensity of natural selection for preferred synonymous variants [13,25,28,83–86].

Faster-X evolution is expected to be strongest in MBGs. Tests to the faster-X hypothesis are thus optimally performed when analyzing MBGs and NBGs separately. Two trends became apparent from this analysis. First, MBGs show strong evidence for faster-X evolution at nonsynonymous sites (1,799 genes; median $d_N = 0.009$ for X-linked genes and 0.006 for autosomal

genes; Mann-Whitney U test, $P = 1×10^{-6}$) but NBGs do not (5,684 genes; median $d_N = 0.003$ for X-linked genes and 0.003 for autosomal genes; Mann-Whitney U test, $P = 0.5$; Fig. 4). Similar patterns are detected for ω ($\omega = 0.22$ *vs.* $\omega = 0.11$; Mann-Whitney U test, $P = 2.5×10^{-8}$ for MBGs and $\omega = 0.07$ *vs.* $\omega = 0.06$; Mann-Whitney U test, $P = 0.053$ for NBGs). Second, MBGs appear to evolve significantly faster than NBGs across the genome (median $d_N = 0.006$ for MBGs and 0.003 for NBGs; Mann-Whitney U test, $P < 1×10^{-32}$), on the X chromosome (median $d_N = 0.009$ for MBGs and 0.003 for NBGs; Mann-Whitney U test, $P = 2.6×10^{-28}$) and on autosomes (median $d_N = 0.006$ for MBGs and 0.003 for NBGs; Mann-Whitney U test, $P < 1×10^{-32}$). All together these results support the idea that MBGs and NBGs evolve under different selective regimes, possibly connected with differences in function [46,87–92], pleiotropic effects, and/or differences in expression properties [46,47] and protein interactions [65].

Genomewide correlations of gene expression levels, tissue specificity, and the number of protein interactions with d_N, d_S and ω, open the possibility that faster-X could be the by-product of differences in these properties between X-linked and autosomal genes (Table 3). To take into account this possibility, I applied a multiple-regression analysis (Table 4). Residuals of d_N are significantly greater for X-linked than for autosomal MBGs (Mann-Whitney U test, $P = 6×10^{-5}$) but this difference is not observed in NBGs (Mann-Whitney U test, $P = 0.48$ for NBGs). In addition, the comparison of divergence between genes with similar expression properties and number of protein interactions using a nonparametric test indicates that MBGs on the X chromosome show consistently higher d_N than autosomal MBGs (Wilcoxon matched pairs test, $T = 4615.5$, $P = 0.014$; see above). Both results imply that faster-X protein evolution in MBGs of *D. yakuba* and *D. santomea* remains detectable even after correcting for differences between the X chromosome and autosomes in expression levels, tissue specificity, and number of protein interactions.

Although expression levels, tissue specificity, and the number of protein interactions are important to our understanding of X-linked divergence of both gene expression and protein sequences in the *D. yakuba—D. santomea* system, additional factors are necessary to explain faster-X evolution. The hemizygosity of the X chromosome in males has been often used as the default explanation [1,2]. Faster-male evolution driven by positive selection [93,94] could

Table 3. Spearman's rank correlations (ρ) of d_N, d_S and ω with gene expression levels, tissue specificity, and protein interactions.

	ρ (P)		
	d_N	d_S	ω
Levels			
MBGs	-0.28 (<0.001)	-0.094 (<0.001)	-0.23 (<0.001)
NBGs	-0.23 (<0.001)	-0.13 (<0.001)	-0.17 (<0.001)
All	-0.24 (<0.001)	-0.12 (<0.001)	-0.19 (<0.001)
Specificity			
MBGs	0.30 (<0.001)	0.10 (<0.001)	0.24 (<0.001)
NBGs	0.22 (<0.001)	0.02 (0.13)	0.21 (<0.001)
All	0.33 (<0.001)	0.057 (<0.001)	0.30 (<0.001)
Interactions			
MBGs-	-0.24 (<0.001)	-0.045 (0.06)	-0.21 (<0.001)
NBGs	-0.23 (<0.001)	-0.07 (<0.001)	-0.19 (<0.001)
All	-0.31 (<0.001)	-0.07 (<0.001)	-0.27 (<0.001)

Note: P, Probability; MBGs, 1,799 male-biased genes; NBGs, 5,684 nonsex-biased genes; All, 9,203 genes.

doi:10.1371/journal.pone.0116829.t003

Table 4. Parameters of multiple-regression models and partial correlations for rates of protein-coding sequence evolution.

	R (P)	b*Lev (P)	b*Spe (P)	b*Int (P)
ω				
MBG	0.21 ($<1\times10^{-12}$)	-0.11 ($<5\times10^{-6}$)	0.14 ($<7\times10^{-7}$)	-0.03 (ns)
NBG	0.17 ($<1\times10^{-12}$)	-0.07 ($<1\times10^{-12}$)	0.13 ($<1\times10^{-12}$)	-0.002 (ns)
All	0.26 ($<1\times10^{-12}$)	-0.06 ($<3.7\times10^{-9}$)	0.23 ($<1\times10^{-12}$)	-0.009 (ns)
d_N				
MBG	0.30 ($<1\times10^{-12}$)	-0.16 ($<1\times10^{-12}$)	0.21 ($<1\times10^{-12}$)	-0.03 (ns)
NBG	0.26 ($<1\times10^{-12}$)	-0.13 ($<1\times10^{-12}$)	0.18 ($<1\times10^{-12}$)	-0.03 (0.012)
All	0.36 ($<1\times10^{-12}$)	-0.10 ($<1\times10^{-12}$)	0.31 ($<1\times10^{-12}$)	-0.029 (0.0027)
d_S				
MBG	0.13 ($<1\times10^{-12}$)	-0.05 (0.03)	0.10 ($<9\times10^{-5}$)	0.001 (ns)
NBG	0.12 ($<1\times10^{-12}$)	-0.12 ($<1\times10^{-12}$)	-0.04 (0.032)	-0.03 (0.011)
All	0.11 (0.026)	-0.10 ($<1\times10^{-12}$)	0.02 (ns)	-0.03 (0.008)

Note: R, correlation coefficient; P, Probability; Lev, average gene expression levels; Spe, tissue specificity in gene expression estimated following Yanai *et al*. (2005) [76]; Int, protein interactions identified from Hansen and Kulathinal (2012) [58]; MBG, male-biased genes; NBGs, nonsex-biased genes; All, all genes; ns, $P > 0.05$.

doi:10.1371/journal.pone.0116829.t004

additionally facilitate faster-X for MBGs in systems where males are the heterogametic sex [95]. Other differences between the *X* chromosome and autosomes such as recombination rates could also be contributing factors to faster-X [19]. Because *Drosophila* males lack crossing over, the *X* chromosome is more exposed to recombination than the autosomes, as it spends 2/3 of its time in females [9,11,19]. The difference is particularly exaggerated in *D. melanogaster*, in which the *X* chromosome also shows an increased rate of recombination per female meiosis relative to that of autosomes [96] and there are abundant autosomal polymorphic inversions [31]. As a result of higher recombination, the efficacy of selection is expected to be higher on the *X* chromosome than on autosomes, which could potentially lead to a higher incidence of adaptive evolution on the *X* chromosome [13,21].

Conclusions

I report here parallel trends of faster-X evolution of gene expression and protein-coding sequences in *D. yakuba* and *D. santomea*. The increased divergence is specific to X-linked MBGs and not detected in NBGs. Multiple-regression and nonparametric analyses indicate that faster-X is not the by-product of differences between X-linked and autosomal genes in properties known to correlate with rates of evolution, such as gene expression levels, tissue specificity in expression, and the number of protein interactions. Because MBGs are significantly less polymorphic for gene expression than NBGs on the *X* chromosome, the observed excess of gene expression divergence is consistent with a higher incidence of positive selection, possibly associated with the hemizygosity of the *X* chromosome in males. Besides male heterogamety, other differences between X-linked and autosomal genes, such as recombination rates, could also contribute to faster-X. Note that an additional contribution of weakly deleterious mutations to faster-X cannot be ruled out, as recurrent adaptive evolution on the *X* chromosome will inevitably reduce N_{eX} due to linkage effects, which in time will result in an increased fixation of weakly deleterious mutations. It is thus possible that the excess of X-linked divergence is the compound effect of both beneficial and weakly deleterious mutations.

Supporting Information

S1 Fig. Gene expression divergence between *D. yakuba* and *D. santomea*. Gene expression divergence (D_e) is shown for all genes analyzed (green, All), male-biased genes (blue, MBGs), and nonsex-biased genes (yellow, NBGs). X, X-linked genes; A, autosomal genes. The heavy horizontal line in each box indicates the median. The length of the box and the whiskers represent 50% and 90% confidence intervals, respectively. Asterisks indicate significant differences (Mann-Whitney U test; ∗∗, $P < 0.001$).
(PDF)

Acknowledgments

Author Contributions

Conceived and designed the experiments: AL. Performed the experiments: AL. Analyzed the data: AL. Contributed reagents/materials/analysis tools: AL. Wrote the paper: AL.

References

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

2. Vicoso B, Charlesworth B (2006) Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet 7: 645–653. PMID: 16847464

3. Herrig DK, Llopart A (2014) Evolution of X-linked male-baised genes in *Drosophila*. In: Encyclopedia of Life Sciences: John Wiley & Sons, Ltd: Chichester. doi: 10.1002/9780470015902.a0025537.

4. Vicoso B, Charlesworth B (2009) Effective population size and the faster-X effect: an extended model. Evolution 63: 2413–2426. doi: 10.1111/j.1558-5646.2009.00719.x PMID: 19473388

5. Mank JE, Vicoso B, Berlin S, Charlesworth B (2010) Effective population size and the Faster-X effect: empirical results and their interpretation. Evolution 64: 663–674. doi: 10.1111/j.1558-5646.2009.00853.x PMID: 19796145

6. Mank JE, Nam K, Ellegren H (2010) Faster-Z evolution is predominantly due to genetic drift. Mol Biol Evol 27: 661–670. doi: 10.1093/molbev/msp282 PMID: 19926635

7. Mank JE, Axelsson E, Ellegren H (2007) Fast-X on the Z: rapid evolution of sex-linked genes in birds. Genome Res 17: 618–624. PMID: 17416747

8. Ellegren H (2009) Genomic evidence for a large-Z effect. Proc Biol Sci 276: 361–366. doi: 10.1098/rspb.2008.1135 PMID: 18826931

9. Langley CH, Montgomery E, Hudson R, Kaplan N, Charlesworth B (1988) On the role of unequal exchange in the containment of transposable element copy number. Genet Res 52: 223–235. PMID: 2854088

10. Charlesworth B (2012) The role of background selection in shaping patterns of molecular evolution and variation: evidence from variability on the *Drosophila* X chromosome. Genetics 191: 233–246. doi: 10.1534/genetics.111.138073 PMID: 22377629

11. Vicoso B, Charlesworth B (2009) Recombination rates may affect the ratio of X to autosomal noncoding polymorphism in African populations of *Drosophila melanogaster*. Genetics 181: 1699–1701. doi: 10.1534/genetics.108.098004 PMID: 19189953

12. Charlesworth B (2001) The effect of life-history and mode of inheritance on neutral genetic variability. Genet Res 77: 153–166. PMID: 11355571

13. Comeron JM (2014) Background Selection as Baseline for Nucleotide Variation across the *Drosophila* Genome. PLoS Genet 10: e1004434. doi: 10.1371/journal.pgen.1004434 PMID: 24968283

14. Thornton K, Long M (2002) Rapid divergence of gene duplicates on the *Drosophila melanogaster* X chromosome. Mol Biol Evol 19: 918–925. PMID: 12032248

15. Thornton K, Long M (2005) Excess of amino acid substitutions relative to polymorphism between X-linked duplications in *Drosophila melanogaster*. Mol Biol Evol 22: 273–284. PMID: 15483321

16. Betancourt AJ, Presgraves DC, Swanson WJ (2002) A Test for Faster X Evolution in *Drosophila*. Mol Biol Evol 19: 1816–1819. PMID: 12270910

17. Thornton K, Bachtrog D, Andolfatto P (2006) X chromosomes and autosomes evolve at similar rates in *Drosophila*: no evidence for faster-X protein evolution. Genome Res 16: 498–504. PMID: 16520459

18. Counterman BA, Ortiz-Barrientos D, Noor MA (2004) Using comparative genomic data to test for fast-X evolution. Evolution 58: 656–660. PMID: 15119449

19. Connallon T (2007) Adaptive protein evolution of X-linked and autosomal genes in *Drosophila*: implications for faster-X hypotheses. Mol Biol Evol 24: 2566–2572. PMID: 17884828

20. Grath S, Parsch J (2012) Rate of amino acid substitution is influenced by the degree and conservation of male-biased transcription over 50 million years of *Drosophila* evolution. Genome Biol Evol 4: 346–359. doi: 10.1093/gbe/evs012 PMID: 22321769

21. Campos JL, Halligan DL, Haddrill PR, Charlesworth B (2014) The relation between recombination rate and patterns of molecular evolution and variation in *Drosophila melanogaster*. Mol Biol Evol 31: 1010–1028. doi: 10.1093/molbev/msu056 PMID: 24489114

22. Begun DJ, Holloway AK, Stevens K, Hillier LW, Poh YP, et al. (2007) Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. PLoS Biol 5: e310. PMID: 17988176

23. Musters H, Huntley MA, Singh RS (2006) A genomic comparison of faster-sex, faster-X, and faster-male evolution between *Drosophila melanogaster* and *Drosophila pseudoobscura*. J Mol Evol 62: 693–700. PMID: 16612543

24. Hu TT, Eisen MB, Thornton KR, Andolfatto P (2013) A second-generation assembly of the *Drosophila simulans* genome provides new insights into patterns of lineage-specific divergence. Genome Res 23: 89–98. doi: 10.1101/gr.141689.112 PMID: 22936249

25. Vicoso B, Haddrill PR, Charlesworth B (2008) A multispecies approach for comparing sequence evolution of X-linked and autosomal sites in *Drosophila*. Genet Res (Camb) 90: 421–431. doi: 10.1017/S0016672308009804 PMID: 19061532

26. Baines JF, Sawyer SA, Hartl DL, Parsch J (2008) Effects of X-linkage and sex-biased gene expression on the rate of adaptive protein evolution in *Drosophila*. Mol Biol Evol 25: 1639–1650. doi: 10.1093/molbev/msn111 PMID: 18477586

27. Bachtrog D, Jensen JD, Zhang Z (2009) Accelerated adaptive evolution on a newly formed X chromosome. PLoS Biol 7: e82. doi: 10.1371/journal.pbio.1000082 PMID: 19402745

28. Singh ND, Larracuente AM, Clark AG (2008) Contrasting the efficacy of selection on the X and autosomes in *Drosophila*. Mol Biol Evol 25: 454–467. PMID: 18083702

29. Muller L, Grath S, von Heckel K, Parsch J (2012) Inter- and intraspecific variation in *Drosophila* genes with sex-biased expression. Int J Evol Biol 2012: 963976. doi: 10.1155/2012/963976 PMID: 22315698

30. Garrigan D, Kingan SB, Geneva AJ, Vedanayagam JP, Presgraves DC (2014) Genome Diversity and Divergence in *Drosophila mauritiana*: Multiple Signatures of Faster X Evolution. Genome Biol Evol 6: 2444–2458. doi: 10.1093/gbe/evu198 PMID: 25193308

31. Langley CH, Stevens K, Cardeno C, Lee YC, Schrider DR, et al. (2012) Genomic variation in natural populations of *Drosophila melanogaster*. Genetics 192: 533–598. doi: 10.1534/genetics.112.142018 PMID: 22673804

32. Lu J, Wu CI (2005) Weak selection revealed by the whole-genome comparison of the X chromosome and autosomes of human and chimpanzee. Proc Natl Acad Sci U S A 102: 4063–4067. PMID: 15728731

33. Veeramah KR, Gutenkunst RN, Woerner AE, Watkins JC, Hammer MF (2014) Evidence for increased levels of positive and negative selection on the X chromosome versus autosomes in humans. Mol Biol Evol.

34. Nielsen R, Bustamante C, Clark AG, Glanowski S, Sackton TB, et al. (2005) A scan for positively selected genes in the genomes of humans and chimpanzees. PLoS Biol 3: e170. PMID: 15869325

35. Khaitovich P, Hellmann I, Enard W, Nowick K, Leinweber M, et al. (2005) Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. Science 309: 1850–1854. PMID: 16141373

36. Hvilsom C, Qian Y, Bataillon T, Li Y, Mailund T, et al. (2012) Extensive X-linked adaptive evolution in central chimpanzees. Proc Natl Acad Sci U S A 109: 2054–2059. doi: 10.1073/pnas.1106877109 PMID: 22308321

37. Baines JF, Harr B (2007) Reduced X-linked diversity in derived populations of house mice. Genetics 175: 1911–1921. PMID: 17287527

38. Kousathanas A, Halligan DL, Keightley PD (2014) Faster-X adaptive protein evolution in house mice. Genetics 196: 1131–1143. doi: 10.1534/genetics.113.158246 PMID: 24361937

39. Vicoso B, Emerson JJ, Zektser Y, Mahajan S, Bachtrog D (2013) Comparative sex chromosome genomics in snakes: differentiation, evolutionary strata, and lack of global dosage compensation. PLoS Biol 11: e1001643. doi: 10.1371/journal.pbio.1001643 PMID: 24015111

40. Jovelin R (2013) Pleiotropic constraints, expression level, and the evolution of miRNA sequences. J Mol Evol 77: 206–220. doi: 10.1007/s00239-013-9588-6 PMID: 24100521

41. Sackton TB, Corbett-Detig RB, Nagaraju J, Vaishna L, Arunkumar KP, et al. (2014) Positive Selection Drives Faster-Z Evolution in Silkmoths. Evolution.

42. Jaquiery J, Stoeckel S, Rispe C, Mieuzet L, Legeai F, et al. (2012) Accelerated evolution of sex chromosomes in aphids, an x0 system. Mol Biol Evol 29: 837–847. doi: 10.1093/molbev/msr252 PMID: 21998277

43. Llopart A (2012) The rapid evolution of X-linked male-biased gene expression and the large-X effect in *Drosophila yakuba*, *D. santomea* and their hybrids. Mol Biol Evol 29: 3873–3886. doi: 10.1093/molbev/mss190 PMID: 22844069

44. Kayserili MA, Gerrard DT, Tomancak P, Kalinka AT (2012) An excess of gene expression divergence on the X chromosome in *Drosophila* embryos: implications for the faster-X hypothesis. PLoS Genet 8: e1003200. doi: 10.1371/journal.pgen.1003200 PMID: 23300473

45. Meisel RP, Malone JH, Clark AG (2012) Faster-X evolution of gene expression in *Drosophila*. PLoS Genet 8: e1003013. doi: 10.1371/journal.pgen.1003013 PMID: 23071459

46. Meisel RP (2011) Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein-coding sequence evolution. Mol Biol Evol 28: 1893–1900. doi: 10.1093/molbev/msr010 PMID: 21239389

47. Meisel RP, Malone JH, Clark AG (2012) Disentangling the relationship between sex-biased gene expression and X-linkage. Genome Res 22: 1255–1265. doi: 10.1101/gr.132100.111 PMID: 22499666

48. Parisi M, Nuttall R, Naiman D, Bouffard G, Malley J, et al. (2003) Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. Science 299: 697–700. PMID: 12511656

49. Sturgill D, Zhang Y, Parisi M, Oliver B (2007) Demasculinization of X chromosomes in the *Drosophila* genus. Nature 450: 238–241. PMID: 17994090

50. 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. PMID: 17994089

51. Meiklejohn CD, Presgraves DC (2012) Little evidence for demasculinization of the *Drosophila* X chromosome among genes expressed in the male germline. Genome Biol Evol 4: 1007–1016. doi: 10.1093/gbe/evs077 PMID: 22975718

52. Meiklejohn CD, Landeen EL, Cook JM, Kingan SB, Presgraves DC (2011) Sex chromosome-specific regulation in the *Drosophila* male germline but little evidence for chromosomal dosage compensation or meiotic inactivation. PLoS Biol 9: e1001126. doi: 10.1371/journal.pbio.1001126 PMID: 21857805

53. Mikhaylova LM, Nurminsky DI (2011) Lack of global meiotic sex chromosome inactivation, and paucity of tissue-specific gene expression on the *Drosophila* X chromosome. BMC Biol 9: 29. doi: 10.1186/1741-7007-9-29 PMID: 21542906

54. Vibranovski MD, Lopes HF, Karr TL, Long M (2009) Stage-specific expression profiling of *Drosophila* spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes. PLoS Genet 5: e1000731. doi: 10.1371/journal.pgen.1000731 PMID: 19936020

55. Chang PL, Dunham JP, Nuzhdin SV, Arbeitman MN (2011) Somatic sex-specific transcriptome differences in *Drosophila* revealed by whole transcriptome sequencing. BMC Genomics 12: 364. doi: 10.1186/1471-2164-12-364 PMID: 21756339

56. Catalan A, Hutter S, Parsch J (2012) Population and sex differences in *Drosophila melanogaster* brain gene expression. BMC Genomics 13: 654. doi: 10.1186/1471-2164-13-654 PMID: 23170910

57. Marbach D, Roy S, Ay F, Meyer PE, Candeias R, et al. (2012) Predictive regulatory models in *Drosophila melanogaster* by integrative inference of transcriptional networks. Genome Res 22: 1334–1349. doi: 10.1101/gr.127191.111 PMID: 22456606

58. Hansen ME, Kulathinal RJ (2013) Sex-biased networks and nodes of sexually antagonistic conflict in *Drosophila*. Int J Evol Biol 2013: 545392. doi: 10.1155/2013/545392 PMID: 23431497

59. Larracuente AM, Sackton TB, Greenberg AJ, Wong A, Singh ND, et al. (2008) Evolution of protein-coding genes in *Drosophila*. Trends Genet 24: 114–123. doi: 10.1016/j.tig.2007.12.001 PMID: 18249460

60. Marais G, Domazet-Loso T, Tautz D, Charlesworth B (2004) Correlated evolution of synonymous and nonsynonymous sites in *Drosophila*. J Mol Evol 59: 771–779. PMID: 15599509

61. Nuzhdin SV, Wayne ML, Harmon KL, McIntyre LM (2004) Common pattern of evolution of gene expression level and protein sequence in *Drosophila*. Mol Biol Evol 21: 1308–1317. PMID: 15034135

62. Drummond DA, Bloom JD, Adami C, Wilke CO, Arnold FH (2005) Why highly expressed proteins evolve slowly. Proc Natl Acad Sci U S A 102: 14338–14343. PMID: 16176987

63. Pal C, Papp B, Hurst LD (2001) Highly expressed genes in yeast evolve slowly. Genetics 158: 927–931. PMID: 11430355

64. Lemos B, Bettencourt BR, Meiklejohn CD, Hartl DL (2005) Evolution of proteins and gene expression levels are coupled in *Drosophila* and are independently associated with mRNA abundance, protein length, and number of protein-protein interactions. Mol Biol Evol 22: 1345–1354. PMID: 15746013

65. Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C, Feldman MW (2002) Evolutionary rate in the protein interaction network. Science 296: 750–752. PMID: 11976460

66. Cariou ML, Silvain JF, Daubin V, Da Lage JL, Lachaise D (2001) Divergence between *Drosophila santomea* and allopatric or sympatric populations of *D. yakuba* using paralogous amylase genes and migration scenarios along the Cameroon volcanic line. Molecular Ecology 10: 649–660. PMID: 11298976

67. Llopart A, Lachaise D, Coyne JA (2005) Multilocus analysis of introgression between two sympatric sister species of *Drosophila*: *Drosophila yakuba* and *D. santomea*. Genetics 171: 197–210. PMID: 15965264

68. Llopart A, Lachaise D, Coyne JA (2005) An anomalous hybrid zone in *Drosophila*. Evolution 59: 2602–2607. PMID: 16526507

69. Ashburner M (1989) Drosophila: A Laboratory Manual. Cold Spring Harbor, NY.: Cold Spring Harbor Laboratory Press.

70. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079. doi: 10.1093/bioinformatics/btp352 PMID: 19505943

71. Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, et al. (2007) Evolution of genes and genomes on the *Drosophila* phylogeny. Nature 450: 203–218. PMID: 17994087

72. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760. doi: 10.1093/bioinformatics/btp324 PMID: 19451168

73. Blankenberg D, Kuster GV, Coraor N, Ananda G, Lazarus R, et al. (2001) Galaxy: A Web-Based Genome Analysis Tool for Experimentalists. Current Protocols in Molecular Biology: John Wiley & Sons, Inc. doi: 10.1002/0471142700.nca01bs00 PMID: 18428806

74. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, et al. (2005) Galaxy: a platform for interactive large-scale genome analysis. Genome Res 15: 1451–1455. PMID: 16169926

75. Goecks J, Nekrutenko A, Taylor J (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol 11: R86. doi: 10.1186/gb-2010-11-8-r86 PMID: 20738864

76. Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, et al. (2005) Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. Bioinformatics 21: 650–659. PMID: 15388519

77. Chintapalli VR, Wang J, Dow JA (2007) Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. Nat Genet 39: 715–720. PMID: 17534367

78. Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Computer Applications in Biosciences 13: 555–556. PMID: 9367129

79. Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution 24: 1586–1591. PMID: 17483113

80. Maynard Smith J, Haigh J (1974) The hitch-hiking effect of a favorable gene. Genetical Research 23: 23–35. PMID: 4407212

81. Gillespie JH (2000) Genetic drift in an infinite population. The pseudohitchhiking model. Genetics 155: 909–919. PMID: 10835409

82. Leffler EM, Bullaughey K, Matute DR, Meyer WK, Segurel L, et al. (2012) Revisiting an old riddle: what determines genetic diversity levels within species? PLoS Biol 10: e1001388. doi: 10.1371/journal.pbio.1001388 PMID: 22984349

83. Campos JL, Zeng K, Parker DJ, Charlesworth B, Haddrill PR (2013) Codon usage bias and effective population sizes on the X chromosome versus the autosomes in *Drosophila melanogaster*. Mol Biol Evol 30: 811–823. doi: 10.1093/molbev/mss222 PMID: 23204387

84. Singh ND, Davis JC, Petrov DA (2005) Codon bias and noncoding GC content correlate negatively with recombination rate on the *Drosophila* X chromosome. J Mol Evol 61: 315–324. PMID: 16044248

85. Singh ND, Davis JC, Petrov DA (2005) X-linked genes evolve higher codon bias in *Drosophila* and *Caenorhabditis*. Genetics 171: 145–155. PMID: 15965246

86. Vicario S, Moriyama EN, Powell JR (2007) Codon usage in twelve species of *Drosophila*. BMC Evol Biol 7: 226. PMID: 18005411

87. Begun DJ, Whitley P, Todd BL, Waldrip-Dail HM, Clark AG (2000) Molecular population genetics of male accessory gland proteins in *Drosophila*. Genetics 156: 1879–1888. PMID: 11102381

88. Haerty W, Jagadeeshan S, Kulathinal RJ, Wong A, Ravi Ram K, et al. (2007) Evolution in the fast lane: rapidly evolving sex-related genes in *Drosophila*. Genetics 177: 1321–1335. PMID: 18039869

89. Jiang ZF, Machado CA (2009) Evolution of sex-dependent gene expression in three recently diverged species of *Drosophila*. Genetics 183: 1175–1185. doi: 10.1534/genetics.109.105775 PMID: 19720861

90. Zhang Z, Hambuch TM, Parsch J (2004) Molecular evolution of sex-biased genes in *Drosophila*. Mol Biol Evol 21: 2130–2139. PMID: 15282334

91. Proschel M, Zhang Z, Parsch J (2006) Widespread adaptive evolution of *Drosophila* genes with sex-biased expression. Genetics 174: 893–900. PMID: 16951084

92. Perry JC, Harrison PW, Mank JE (2014) The ontogeny and evolution of sex-biased gene expression in *Drosophila melanogaster*. Mol Biol Evol 31: 1206–1219. doi: 10.1093/molbev/msu072 PMID: 24526011

93. Wu C-I, Davis AW (1993) Evolution of postmating reproductive isolation: the composite nature of Haldane's rule and its genetic bases. Am Nat 142: 187–212. doi: 10.1086/285534 PMID: 19425975

94. Wu CI, Johnson N, Palopoli MF (1996) Haldane's rule and its legacy: why are there so many sterile males. Trends Ecol Evol 11: 281–284. PMID: 21237844

95. Begun DJ, Whitley P (2000) Reduced X-linked nucleotide polymorphism in *Drosophila simulans*. Proc Natl Acad Sci USA 97: 5960–5965. PMID: 10823947

96. Comeron JM, Ratnappan R, Bailin S (2012) The many landscapes of recombination in *Drosophila melanogaster*. PLoS Genetics 8: e1002905. doi: 10.1371/journal.pgen.1002905 PMID: 23071443