American Genetic Association

Journal of Heredity, 2019, 219–228
doi:10.1093/jhered/esy063
Original Article
Advance Access publication November 24, 2018

OXFORD

Original Article

Genome Size Evolution within and between the Sexes

Carl E. Hjelmen, Margaret A. Garrett, V. Renee Holmes, Melissa Mynes, Elizabeth Piron, and J. Spencer Johnston

From the Department of Entomology, Texas A&M University, 2475 TAMU, College Station, TX 77843 (Hjelmen, Holmes, and Johnston); and Department of Biochemistry and Biophysics, Texas A&M University, 2475 TAMU, College Station, TX 77843 (Garrett, Mynes, and Piron).

Address correspondence to Carl E. Hjelmen at the address above, or e-mail: cehjelmen09@tamu.edu.

Received April 26, 2018; First decision November 15, 2018; Accepted November 21, 2018.

Corresponding Editor: Therese Markow

Abstract

Genome sizes are known to vary between closely related species, but the patterns behind this variation have yet to be fully understood. Although this variation has been evaluated between species and within sexes, unknown is the extent to which this variation is driven by differentiation in sex chromosomes. To address this longstanding question, we examine the mode and tempo of genome size evolution for a total of 87 species of Drosophilidae, estimating and updating male genome size values for 44 of these species. We compare the evolution of genome size within each sex to the evolution of the differences between the sexes. Utilizing comparative phylogenetic methods, we find that male and female genome size evolution is largely a neutral process, reflective of phylogenetic relatedness between species, which supports the newly proposed accordion model for genome size change. When similarly analyzed, the difference between the sexes due to heteromorphic sex chromosomes is a dynamic process; the male-female genome size difference increases with time with or without known neo-Y events or complete loss of the Y. Observed instances of rapid change match theoretical expectations and known neo-Y and Y loss events in individual species.

Subject area: Molecular systematics and phylogenetics
Keywords: comparative phylogenetic analyses, *Drosophila*, genome architecture, genome size evolution, sex chromosomes, *Sophophora*

Genome size (also termed C-value) varies widely across organisms, with up to 7000-fold variation in animals alone, yet does not correlate with complexity among eukaryotes (Mirsky and Ris 1951; Palazzo and Gregory 2014). This has been commonly termed the C-value paradox (or enigma) (Gregory 2001). The C-value paradox has been historically been explained by variation in amounts of nongenic and repeat regions resulting from transposable elements, satellite DNA, tandem repeats, and even copy number variation rather than differences in the amount of DNA in coding genes and intergenic regions (Gregory and Hebert 1999; Kidwell 2002; Schaeffer et al. 2008; Kelley et al. 2014). Much of the known variation in genome size between closely related species of plants and animals is explained by differential accumulation of transposable elements (Bennetzen and Kellogg 1997; Ågren and Wright 2011). When sister species *Drosophila melanogaster* and *Drosophila simulans* are compared, there was considerably less evidence of transposable elements in *D. simulans*, the species with a smaller genome (Vieira and Biémont 2004). A more recent example found a significant relationship between genome size and global transposable element content among 26 species of *Drosophila* in a phylogenetic context (Sessegolo et al. 2016).

Although the C-value paradox has been explained through the accumulation of noncoding DNA, the mechanisms and patterns for

the resulting variation are still debated. Many hypotheses for the patterns of long-term change have been proposed, ranging from neutral change via imbalances in insertions and deletions to forces selecting for/against genome size correlates. Each of these hypotheses has incomplete support in the literature, suggesting there is no one answer to the question of genome size variation. The mutational equilibrium hypothesis, which proposes that genome size changes slowly over time due to an imbalance between insertions and deletions (Petrov et al. 2000; Petrov 2001, 2002a, 2002b), has been criticized because it is only presented in small datasets. It is believed that the mechanisms proposed are largely theoretical and too slow in real systems to provide the variation that is currently present between species (Gregory 2003; Gregory 2004). The small effective population size, which hypothesizes that larger, more deleterious genomes are more likely to be fixed and less likely to be selected out of populations when the effective population size is small (Lynch and Conery 2003), loses its statistical support when phylogenetic analyses were introduced with the original data (Whitney and Garland 2010). Current datasets do not reflect this hypothesis when analyzed with phylogenetic methods, but this is not sufficient evidence to disregard it. The recently hypothesized accordion model offers the promising explanation that genome size variation (Kapusta et al. 2017) balances large insertions due to transposable elements with deletions. This may provide a pattern similar to the mutational equilibrium model, but evolves at a fast enough rate to fit large datasets.

Although there have been attempts to study the variation in genome size across organisms (Gregory 2015; Hanrahan and Johnston 2011), much of this work is only related to genome size for the species (average of male and female) or one sex (primarily the female). Recent studies have found significant genome size variation within species may lead to divergent phenotypes (Ellis et al. 2014; Huang et al. 2014; Arnqvist et al. 2015). When *D. melanogaster* DGRP lines with large and small genomes were raised in different thermal environments, genome size and genome size by temperature interactions were significantly related to various important developmental phenotypes, including time to pupation and survival to pupation and adulthood (Ellis et al. 2014). Although there is considerable support for the neutral models of genome size evolution, such results suggest that variation in genome size within a species may have a significant impact on the evolutionary ecology of species, suggesting an adaptive model.

Males and females may have very different life-history selection parameters, which in turn can influence the size of the sex chromosomes. Genome size differences between sexes are positively associated with male competitive fertilization success and female lifetime fecundity in the seed beetle *Callosobruchus maculatus* (Arnqvist et al. 2015). Similarly, for some male grasshoppers, genome size is negatively correlated with the song attractiveness, suggesting that sexual selection may be indirectly acting on genome size evolution between sexes (Schielzeth et al. 2014). Given that there is a known relationship between variation in genome size within a species and divergent phenotypes (Ellis et al. 2014; Huang et al. 2014) and that genome size variation has impacts on reproductive fitness (Schielzeth et al. 2014; Arnqvist et al. 2015), the question becomes: Will the answers change if we look at the difference between males and females? What do these differences tell us about the formation and degradation of sex chromosomes?

Heteromorphic sex chromosomes are hypothesized to originate as a pair of homologous chromosomes containing alleles related to male and female determination and organismal success (Goodfellow et al. 1983; Charlesworth et al. 2005). Inversions and other methods of recombination suppression occur over time effectively preventing genetic exchange of sex determination factors between the heteromorphic sex chromosomes. This suppression is followed by the accumulation of transposable elements. This accumulation is transient because transposable elements may increase the instances of chromosome breaks (Charlesworth et al. 2005; Bachtrog 2013) or insert into coding and regulatory regions of the Y chromosome, inactivating genes. Both lead to eventual loss of genes and gene function (Matsunaga 2009). These changes, along with the differential selection on the X and Y, will result in permanent chromosomal heterozygosity (Muller 1918; Charlesworth et al. 2005; Bachtrog 2013). This process of sex chromosome evolution may first result in an increase in size of the Y chromosome before it inevitably loses coding sequences and, in most cases, physical size (Charlesworth and Charlesworth 2000; Blackmon and Brandvain 2017). In X–Y systems, males with a larger genome size than females could indicate the presence of a neo-Y system through fusions or translocations. However, not all instances of neo-Y systems would provide this pattern given the rapid decay of the Y. During this process of gain and loss, the genome size, particularly the relative size of the X and Y chromosome and the amounts of heterochromatin and euchromatin in each, may change significantly. Cytologically, it is still difficult to disentangle chromosome size from genome size differences due to the tight packing behavior of heterochromatin. Whether or not chromosomal responses to sex determination generally extend to species genome size variation in the 2 sexes is generally unknown.

This cyclical process of bloat, decay, loss, and regeneration of the Y chromosome may be quicker in higher Diptera (Brachycera), and other organisms with achiasmatic meiosis, relative to the process in other organisms, including most mammals that have chiasmatic meiosis. For example, the entire neo-Y chromosome in *Drosophila* lacks recombination as soon as it is formed (Blackmon and Demuth 2015a). In contrast, organisms with chiasmatic meiosis must fix a series of mutations (often inversions) that suppress recombination in the sex chromosomes before the Y can begin to differentiate from the X, a process that can take millions of years (Lahn and Page 1999). This rapid turnover in Y chromosomes in insects, specifically *Drosophila*, has made them the model system for studying sex chromosome differentiation (Blackmon and Demuth 2015b).

Here, we report the genome sizes for both males and females of 87 species of Drosophilidae, with a focus on the subgenus *Sophophora*, estimating and updating male genome size values for 44 of these species. Species outside of *Sophophora* are included as outgroups for appropriate phylogenetic comparison. Male and female genome size values are analyzed in a phylogenetic context utilizing modern phylogenetic comparative methods. Female *Sophophora* have been analyzed in a phylogenetic context (Hjelmen and Johnston 2017); it is unknown whether male whole genomes exhibit the same or different evolutionary patterns. The difference in male and female genome size is also of interest because it can serve as a relative proxy for sex chromosome differentiation, assuming the change in the X chromosome from the neo-X is small relative to the neo-Y transition. This male–female size difference can be analyzed in a similar fashion to whole-genome size to answer questions related to the evolution of heteromorphic sex chromosomes.

We hypothesize that the X–Y difference will show a spontaneous and sporadic pattern of evolution (explained below), as selection may act rapidly on regions that determine successful production of males and females. The hypothesized pattern would differ from the patterns found for whole-genome size in species of the *Sophophora* subgenus and the whole *Drosophila* genus (Sessegolo et al. 2016;

Hjelmen and Johnston 2017). Any dramatic changes in the X–Y system would produce a markedly different individual, or clade, and a loss of phylogenetic signal or predictability. This would be true in particular when the Y chromosomes are not identical by descent (Carvalho and Clark 2005; Koerich et al. 2008; Carvalho et al. 2009). Hypothesizing that change in a degraded Y is less predictable, we do not expect gradual pattern of change; saltatory changes would be indicated by departures from signals of gradualistic evolution. These nonphylogenetic differences in mode and tempo of change in size would potentially pinpoint major changes in the differentiation of heteromorphic sex chromosomes, the formation of neo-Y chromosomes, and the rapid degradation of Y chromosomes.

Methods

Genome Size Database

Forty-three genome sizes were obtained for male and female *Sophophora* from published datasets (Gregory and Johnston 2008). Forty-four new species values for males and females were estimated concurrently under the same run conditions, with new and updated male values reported here scored along with females from the same strain to ensure that updated run conditions have not changed the estimated female genome values (Supplementary Table S1). Multiple biological replicates were run for each sex to determine the level of variation by sex for each species/strain estimate. Genome sizes were estimated using flow cytometry (Johnston, et al. 2019) for species and strains from the UC San Diego Species Stock Center (http://stockcenter.ucsd.edu). Species from Gregory and Johnston (2008) were obtained from the Tucson Species Stock Center and were not published with stock number information. Newly obtained species were maintained at ambient laboratory temperatures for only 2–3 generations in the laboratory of J. Spencer Johnston before all estimates for each species was complete.

Because genome size is traditionally represented as the 1C amount, or the amount in a haploid cell, genome size for males is an average of the "X" and the "Y" gametes (1 autosomal set + ½[X + Y]), whereas the female 1C is truly half (1 autosomal set + X). Doubling the 1C size recreates the diploid genome size, with the full sex chromosome complement and the diploid autosomes. The difference in genome size due to heteromorphic sex chromosomes for each species was calculated by subtracting double the male genome size estimate from double the female genome size. This can be represented below, where A represents the autosomal portion of the genome:

$$\begin{aligned} ♀ - ♂ &= 2^*(1A + X) - 2^*\left(1A + \left(\frac{1}{2}\right)^*(X + Y)\right) \\ &= (2A + 2X) - (2A + X + Y) = X - Y \end{aligned}$$

It is important to note that this equation shows sex difference due to heteromorphic sex chromosomes as a simple X–Y. This difference may be due to any combination of sex chromosomes, such as XX/XO or a neo-Y, where a neo-Y and Y may be present. In the case of a neo-Y system, such as *Drosophila pseudoobscura* and *Drosophila miranda*, this difference involves multiple Muller elements. In *D. pseudoobscura*, the ancestral Y (Muller A element) found in related species has fused with the Muller F element to make a larger dot chromosome, while one Muller D element has fused with the X to create a novel submetacentric X. The other homologous Muller D element has become a neo-Y (YD). More recently, in *D. miranda*, this neo-Y (YD) has fused with one of the Muller C elements and the other, homologous C element has formed a neo-X (Mahajan et al. 2018).

Genome Size Analysis

A GS means comparison was run between sexes for each species using PROC GLM with a pdiff function in SAS 9.4 (SAS S, Version S 2003) and Benjamini–Hochberg control of false discovery rate set to 0.01 (Benjamini and Hochberg 1995). This was used to test for significant differences between estimates of genome size for each sex using the model $Y_{ij} = u + species_i + sex\ (species)_i + e_{ij}$. A Kolmogorov–Smirnov test was also performed using R 3.2.3 to test for a significant difference in the distribution of female and male genome sizes (R Core Team 2016).

Phylogeny Reconstruction

The *Sophophora* phylogeny used in this study is the same used in an earlier study of genome size evolution in female *Sophophora* (Hjelmen and Johnston 2017). This phylogeny utilized 16 genes (4 mitochondrial and 12 protein-coding genes) to reconstruct a molecular phylogeny using a supermatrix method (using *COI*, *COII*, *COIII*, *Cytb*, *Amy*, *AmyRel*, *Ddc*, *boss*, *SNF*, *Marf*, *Sod*, *per*, *Wee*, *HB*, *ADH*, and *fkh*). Genes were obtained from NCBI Genbank, aligned using MAFFT v.7 online (http://mafft.cbrc.jp/), and corrected by hand after inspection in Mesquite 2.75 (see Hjelmen and Johnston 2017 for accession numbers, phylogeny in Supplementary Figure S1). These genes were selected, as they were found common to species of interest and have been found to be reliable in past studies of *Drosophila* phylogenetic relationships (van der Linde et al. 2010; Kellermann, Loeschcke, et al. 2012; Kellermann, Overgaard, et al. 2012).

Phylogenetic Analyses

Male genome size, female genome size, and the difference in genome size between the male and female of each sex were analyzed on the reconstructed phylogeny using the fitContinuous function in the package "geiger" in R 3.2.3 (Harmon et al. 2008; R Core Team 2016). This function allows for likelihood and Akaike information criterion (AICc) comparisons of 5 different models of trait evolution (Ornstein–Uhlenbeck [OU], Brownian motion [BM], trend diffusion, drift, and the white-noise) to determine the best model for trait evolution. The BM model assumes that the amount of similarity in a trait is proportional to the shared ancestry of the species in question (Felsenstein 1973). The OU model assumes a random walk along a branch in a directional fashion, with strength proportional to the estimated alpha value (Butler and King 2004). The trend diffusion model estimates a linear trend of change in the trait, estimating a slope value. The drift model of trait evolution estimates directional change in trait values across the phylogeny, estimating negative values of drift for downward change and positive values for upward change. Like the BM model, the White-noise model assumes randomness in the model, yet the White-noise model is nonphylogenetic, assuming the trait values are randomly assigned from a normal distribution and do not have a covariance structure (signal).

Male and female genome sizes were analyzed in relation to the phylogeny using Pagel's parameters of evolution (Pagel 1999) for comparison to the results found in Hjelmen and Johnston (2017) for female genome size. Pagel's parameters of evolution have been found robust to type I error and provide comparable and more reliable results than other phylogenetic comparative methods (Münkemüller et al. 2012; Hjelmen and Johnston 2017). The Pagel's analyses were run assuming a BM model, given the results from the fitContinuous analysis. The raw values for the phylogenetic parameters were compared between male and female genome sizes, and the significance

from the bounds was compared between this study and the previously published study on female genome size evolution (Hjelmen and Johnston 2017).

In addition to whole-genome size comparison, the difference in genome size of the female and male of each species was calculated as above and analyzed using the same comparative methods as the whole-genome size. All estimates of Pagel's parameters were completed in R utilizing the function PGLS from package "caper" (Orme et al. 2018). Genome size (for males and females) and sex differences, which are primarily, if not entirely due to the heteromorphic sex chromosomes, were then mapped onto the phylogeny using the ContMap function from the phytools package from R 3.3.0 (Revell 2012). Phylogenetic signal is measured by the parameter λ, which ranges from no signal to complete phylogenetic signal (0–1). The parameters kappa and delta measure where change occurs, on individual branches or in regard to the entire tree, respectively. Values of "1" indicate gradualistic change. The hypothesized pattern of change due to sex differences would be supported by a reduced amount of phylogenetic signal ($\lambda < 1$), and departure from gradual, predictable change throughout evolutionary time ($\delta \neq 1$, $\kappa \neq 1$).

Results

Genome Size Information

Genome size and the difference in genome size between sexes due to the heteromorphic sex chromosomes for each species can be found in Supplementary Table S1. The distribution of genome sizes for females and males is shown in Figure 1. There was no statistical difference between male and female genome size distributions across the *Sophophora* (Figure 1) when tested with Kolmogorov–Smirnov ($D = 0.12644$, $P = 0.49$). When analyzed within each species, 43 of the 87 species were found to have statistically different female and male sizes according to the PDIFF analysis ($P < 0.01$) (Table 1, Supplementary Table S1). The male genome is larger than that of the female in 3 of the 43 statistically different values (Table 1); only one of these significantly larger males is found within the *Sophophora* clade, *D. paulistorum*. The distribution of sex differences is plotted in Figure 2, with color indicating XY, XO, and neo-Y systems. As expected, instances of large Y chromosomes are rare; there are very few instances of species where the difference due to sex is less than zero (Supplementary Table S1), resulting in a very small tail on the left side of the distribution (Figure 2), with the majority of sex difference values positive.

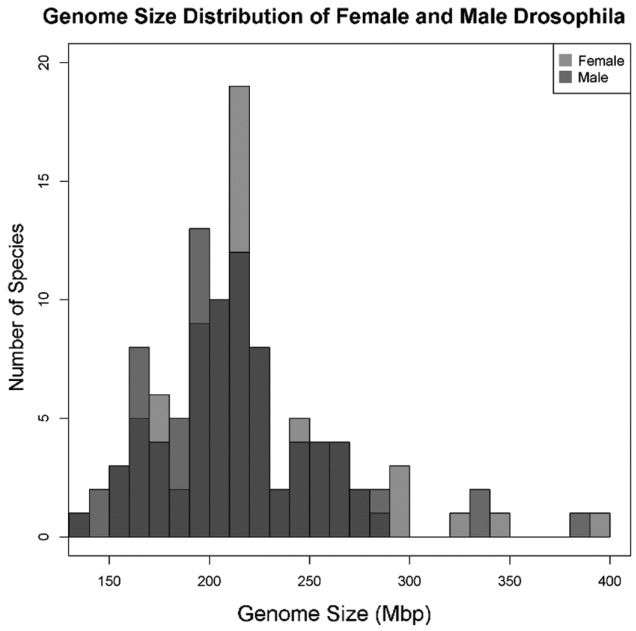

Figure 1. Distribution of female and male genome size in *Sophophora* and outgroups. Genome size is plotted in histogram form to visualize the differences between females and males of *Sophophora* species. The distributions are not statistically different (Kolmogorov–Smirnov, $D = 0.12644$, $P = 0.49$). See online version for full colors.

Comparative Phylogenetic Results

When comparing models of trait evolution (OU, BM, Trend, Drift, and White) using whole male and female genome sizes, all models performed better than the White-noise model of trait evolution. The lowest support in the White-noise model suggests that genome size is not assigned randomly among taxa, but is rather correlated with species relatedness (Table 2). BM was found to perform better than the OU, with OU providing an alpha value of 0.0 in both males and females (Table 2), which indicates there is not a strong directional change in the evolution of genome size in either sex. This in conjunction with the lower AICc value suggests BM is an acceptable model for Pagel's parameters of evolution. The slope estimated from the Trend model (female = −1.43, male = −1.50, Table 2) suggests that whole-genome sizes in these *Drosophila* species are decreasing throughout evolutionary time. The latter result is supported by the negative Drift values obtained from the Drift model (Table 2).

When comparing the tested models of trait evolution using sex difference information for *Drosophila* species, there is uniformity across likelihood scores, yet there is higher performance in the White-noise model (Table 2). The increased likelihood and AICc score of the White-noise model suggests that sex differences may not be as related to phylogeny as whole-genome size. There is, however, phylogenetic signal, suggesting that a phylogenetic model should be used. When comparing BM and OU, OU was found to perform slightly better, indicating a directional change in sex differences. The alpha value from the OU model suggests that there is a slight attraction toward positive change (alpha = 2.72, Table 2) or an increasing difference between a relatively large female genome size and a smaller male genome size. The relative decrease in the size of the Y is also supported with the positive slope from the Trend model (slope = 16.02, Table 2) and the positive Drift value (Drift = 28.71, Table 2).

Pagel's Parameters of Evolution

Male whole-genome size was found to have significant phylogenetic signal ($\lambda = 1$) with high levels of support (Table 3). The male genome size was also found to change gradually along branch lengths ($\kappa = 1.244$) with high support for change early in the phylogeny ($\delta = 0.670$) (Table 3). The values for male genome size are visualized on the phylogeny (Supplementary Figure S2). Here we can see that closely related species have similar sizes, represented by similar colors, while there is a large change early in the phylogeny. Overall, there is a gradual change in size downwards, visually supporting the values obtained by Pagel's parameters.

In contrast, the heteromorphic sex chromosome difference was found to have only partial phylogenetic signal ($\lambda = 0.829$) with significant departure from full phylogenetic signal (H_0: $\lambda = 1$, Table 4). The sex difference was found to have rapid early change in branches ($\kappa = 0.399$) with longer paths in the tree contributing more to change

Table 1. Significant sex differences

Species	Sex difference (Mbp)	Significance	Species	Sex difference (Mbp)	Significance
C. amoena[a]	20.8	0.0016	***D. paulistorum***	**−10.8**	0.0005
C. procnemis[a]	74.4	<0.0001	*D. pectinifera*	34.1	<0.0001
D. affinis	46.6	<0.0001	*D. persimilis*	58.7	<0.0001
D. algonquin	22	<0.0001	*D. phaeopleura*	31.3	<0.0001
D. ambigua	23.8	<0.0001	*D. pseudoananassae*	19.6	0.0049
D. auraria	12	<0.0001	*D. pseudoobscura*	39	<0.0001
D. azteca	14.7	<0.0001	*D. punjabiensis*	10.2	0.0011
D. baimaii	11.2	0.0002	*D. rufa*	21.6	<0.0001
D. biauraria	51.8	<0.0001	*D. sechellia*	8.7	0.0029
D. bicornuta	36.8	<0.0001	*D. simulans*	24.8	0.0004
D. bifasciata	28.2	<0.0001	*D. suzukii*	18.6	<0.0001
D. capricorni	12.9	<0.0001	*D. takahashii*	27.1	<0.0001
D. emarginata	21.2	<0.0001	*D. tani*	21.5	<0.0001
D. greeni	14.8	<0.0001	*D. tolteca*	20.5	0.0031
D. hydei[a]	15.8	<0.0001	*D. triauraria*	8.6	0.0039
D. jambulina	17.2	<0.0001	*D. tsacasi*	26.6	<0.0001
D. kikkawai	10.5	0.0008	*D. varians*	21.5	0.0022
D. lacteicornis	35.8	<0.0001	***D. virilis***[a]	**−25.6**	0.0003
D. lucipennis	77.3	<0.0001	*D. vulcana*	9.8	0.0004
D. mayri	12	<0.0001	*H. pictiventris*[a]	41.1	<0.0001
D. nebulosa	16.8	<0.0001	***S. pattersoni***[a]	**−22.5**	0.0012
D. paralutea	24.5	0.0005			

Species in which there was a significant difference in female and male genome size utilizing a PDiff analysis in SAS 9.4 with *P* values less than the Benjamini–Hochberg critical value for 0.01 false discovery rate. In 3 of the 43 significant differences, males have larger genomes than the female, as indicated by negative sex difference values (in bold).

[a]Significantly different species outside of *Sophophora*.

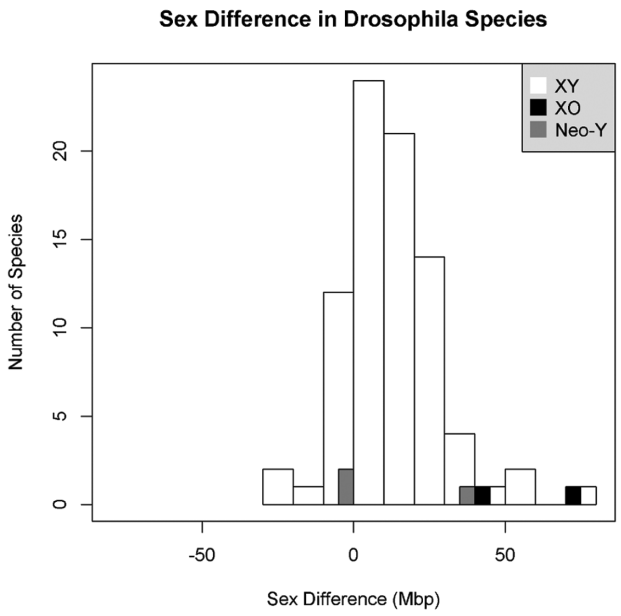

Figure 2. Distribution in difference in genome size due to heteromorphic sex chromosomes in *Sophophora*. The difference in genome size between females and males for 82 XY, 2 XO, and 3 neo-Y species was calculated by subtracting double the male genome size from double the female genome size. These values were then visualized in histogram form for all species. These sex differences were largely positive, indicating that females have larger genome sizes than males. All species differences are in Supplementary Table S1. Any species with significantly different sex differences are given in Table 1. Sex systems known to be XY, XO, and neo-Y are shown in white, black, and gray, respectively.

(δ = 1.691). These tests of mode and rate of change depart significantly from gradual change (H_0: κ and δ = 1, Table 2). These patterns can be visualized on a color phylogeny (Figure 3). Most of the phylogeny shares a similar color, which probably gives the phylogenetic signal. When there is change, it occurs late in the tree, in individual species, supporting the δ value.

Through the use of Pagel's parameters of evolution, the statistical difference from bounds and the low type I error rate (Münkemüller et al. 2012) allow for direct comparison of results between trait values. When the phylogenetic values for male genome sizes are compared with those of female genome values, there is no substantial difference in the results (including likelihood scores; Table 5). The parameter values found for heteromorphic sex chromosome size differences were not the same as those found for the whole-genome sizes of males and females, indicating different patterns in trait evolution (Table 5).

Discussion

We report the genome sizes for males and females and calculate the differences due to heteromorphic sex chromosomes for 87 species of Drosophilidae, focusing on *Sophophora* (Supplementary Table S1). In most cases (43 of 87), the female genome is significantly larger than that of the male, as indicated by the positive sex difference values in Supplementary Table S1. These 43 are consistent with a reduction in the Y chromosome. There are 3 instances (two of which are in the outgroup species) where males have a larger whole-genome size than females of the same species, suggesting the presence of an inflated neo-Y chromosome.

Male genome size was found to have complete phylogenetic signal (λ = 1), with change early in the phylogeny (δ = 0.670) and a gradual

Table 2. Comparison of AICc and likelihood values from different models of trait evolution

	Female genome size			Male genome size			Sex difference		
Model	Value	AICc	Log-likelihood	Value	AICc	Log-likelihood	Value	AICc	Log-likelihood
BM	—	860.57	−428.21	—	862.69	−429.27	—	806.67	−401.26
OU (alpha)	0.00	862.72	−428.21	0.00	864.83	−429.27	2.72	800.31	−397.01
Trend (Slope)	−1.43	862.23	−427.97	−1.50	864.30	−429.01	16.02	805.21	−399.46
Drift (Drift)	−100.00	861.79	−427.75	−100.00	863.89	−428.80	28.71	808.79	−401.25
White noise	—	903.77	−449.82	—	902.23	−449.04	—	746.35	−371.10

Five different models of trait evolution were compared for female and male genome size as well as sex differences utilizing the fitContinuous function in the "geiger" package of R3.2.3. For both female and male genome sizes, BM model performed the best, with similar performance for other models, aside from decreased performance in the White-noise model, suggesting trait values are correlated with species relationships. Directional models (OU, Trend, and Drift), suggest there is a slight downward trend in genome size change in both males and females across the phylogeny. For sex differences, the White-noise model performed best; however, it is a nonphylogenetic model. When accounting for signal, the OU model performed best with regard to sex differences. Directional models (OU, Trend, and Drift) suggest there is a trend toward increasing difference between larger female genomes and smaller male genomes.

Table 3. Comparative phylogenetic values for male *Sophophora* genome size

Test	Value	Significance
λ	1	<2.22e−16 (from 0), 1 (from 1)
δ	0.670	2.27e−10 (from 0), 0.56124 (from 0.5), <2.22e−16 (from 3)
κ	1.244	3.44e−15 (from 0), 0.067835 (from 1), <2.22e−16 (from 3)

Male genome size was found to have complete phylogenetic signal. Although κ was found to be greater than 1, it was not significantly different from 1, indicating gradual change in individual branches. The δ value less than 1 indicates that change probably happened early in the phylogeny.

Table 4. Comparative phylogenetic values for the sex differences in *Sophophora*

Test	Value	Significance
λ	0.829	2.689e−06 (from 0), 0.0175 (from 0.5), 7.77e−17 (from 1)
δ	1.691	4.62e−15 (from 0), 0.0921 (from 1), 2.00e−04 (from 3)
κ	0.399	0.00053 (from 0), 0.30757 (from 0.5), <2.22e−16 (from 3)

Pagel's parameters of evolution for the difference in size due to heteromorphic sex chromosomes/neo-sex chromosomes suggest partial phylogenetic signal (λ < 1.0) and early change in branch lengths (κ < 1.0). Overall change was found to occur late in the phylogeny (δ > 1.0).

change along branches (κ = 1.244). Substantial change early in the phylogeny followed by gradual subsequent change suggests imbalance of insertion and deletion of DNA (Petrov 2002a). Although there are instances of genome size increase (i.e., *Drosophila suzukii*), there is a general, gradual trend downwards in size, as noted by the decrease in size since the split with *Chymomyza* and most notably the reductions in the *melanogaster* and *obscura* groups. This downward trend is suggested through the trait model tests, which found decreasing Trend and Drift values supporting decreases in genome size along the phylogeny (Trend Slope = −1.50, Drift = −100.00; Table 2).

The values and likelihood values for each phylogenetic parameter and the significance from the parameter bounds (Table 5) for males are not different from those found in the females of *Sophophora* (Hjelmen and Johnston 2017). Gradual change and phylogenetic signal in genome size are as proposed for the mutational equilibrium model (Petrov 2002b). Because the rate by which this change would occur with small insertions and deletions is probably too slow to give us results such as these (Gregory 2003, 2004), it is more likely that these results are supportive of the accordion model of genome size evolution, where large deletions are sufficient to balance out new insertions of transposable elements (Kapusta et al. 2017). However, much of this is conjecture, and formal tests of these changes in the genome are necessary before making concrete conclusions on these hypotheses.

Compared with the whole male and female genomes, the difference in genome size due to the relative sizes of the X and Y chromosome behaved very differently over evolutionary time (Table 5); the sex difference was found to have an incomplete phylogenetic signal (λ = 0.829). This reduction in signal was supported through the increased likelihood value from the White-noise model of trail evolution (Table 2). There is an indication of more change happening on long paths (from root to tip) in the tree suggesting that more change occurs later in the phylogeny (δ = 1.691). There is also evidence for early change on individual branches (κ = 0.399), suggesting that when change occurs later in the phylogeny, it happens rapidly (Table 4). A visual phylogenetic representation of this is shown in Figure 3. The seemingly random changes (support for White-Noise model, Table 2) in sex differences are responsible for the λ value of less than 1.0 and an incomplete phylogenetic signal. Although these results are dramatically different from those found when looking at the evolution of whole-genome size (Table 3), they are not unexpected given the hypothesized models of sex chromosome differentiation and Y chromosome evolution and degradation (Charlesworth et al. 2005; Bachtrog 2013).

Trend and Drift models found evidence for increases in the difference between the relatively larger female genomes and smaller male genomes throughout evolutionary time (Trend Slope = 16.02, Drift = 28.71, Table 2). Although this could indicate that female genomes (X chromosomes) are getting larger, the proposed model of Y-chromosome degradation (reviewed by Charlesworth and Charlesworth 2000; Charlesworth et al. 2005; Bachtrog 2013) predicts a trend toward a degenerated Y chromosome. Along with degeneration of the genic qualities of the Y chromosome, a decrease in the physical amount of DNA as represented by the consistently smaller Y chromosomes is expected and found across species in Diptera (Supplementary Table S1) (Gregory and Johnston 2008; Picard et al. 2012).

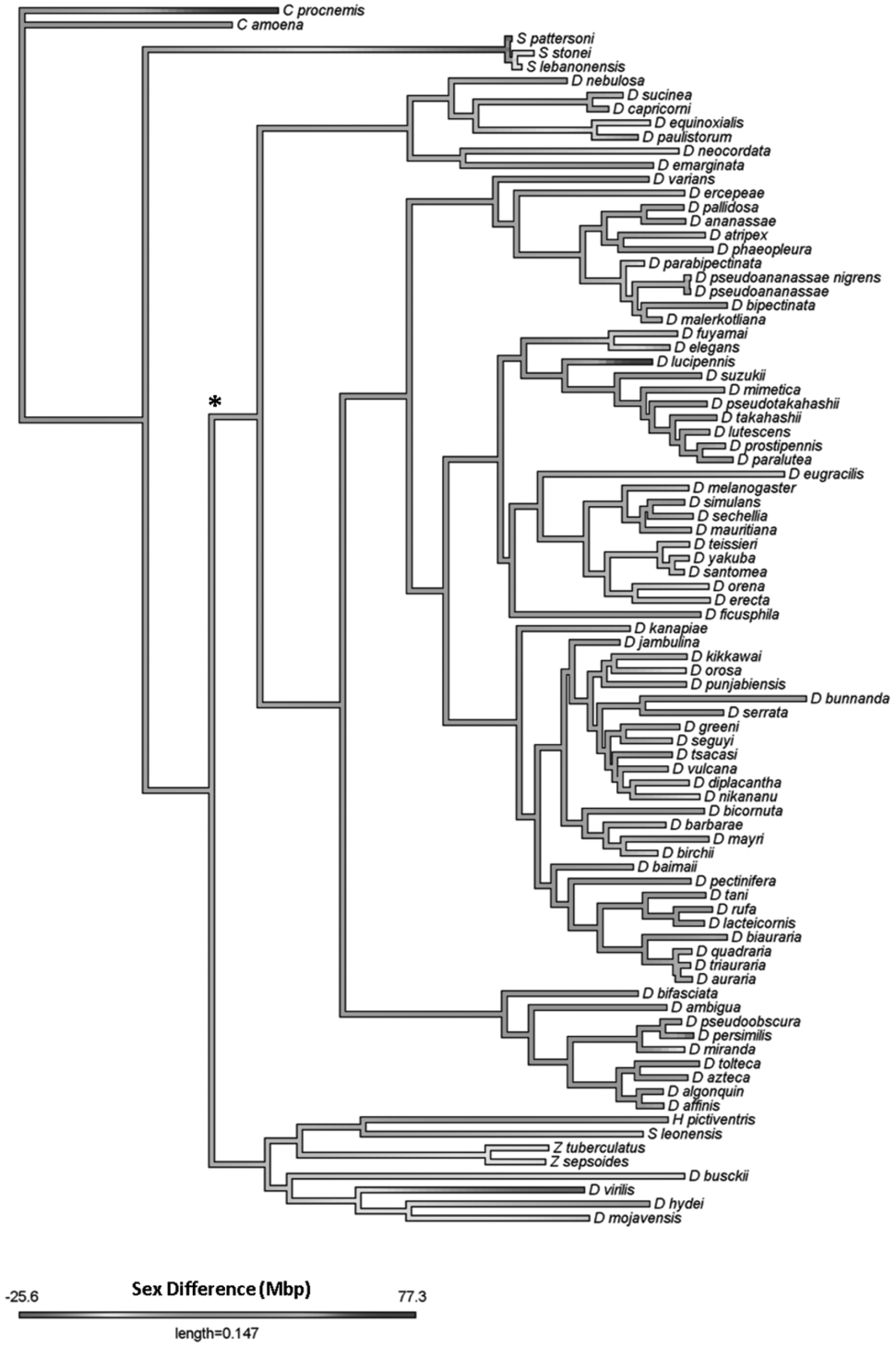

Figure 3. Difference due to heteromorphic sex chromosomes plotted on *Sophophora* phylogeny. The phylogeny of 87 Drosophilidae reconstructed with MrBayes 3.2.3 with a focus on *Sophophora*. The split to *Sophophora* is indicated in the figure with "*". Species with XO configuration are indicated with "O" and species with neo-Y are indicated with "+". The difference in genome size due to the heteromorphic sex chromosomes is visualized in color utilizing ancestral character state reconstruction in the contMap function in the Phytools package in R. Positive values indicate females of the species are larger than males, whereas negative values indicate males have larger genomes than the females. See online version for full colors.

To evaluate whether these conclusions were due to exceptional sex differences associated with species with neo-Y and XO systems, we reran the analysis with the 5 known neo-Y and XO cases removed. Despite the lack of statistical significance, these new values strengthened the conclusions (Supplementary Table S2). There was incomplete phylogenetic signal (λ = 0.68 vs. 0.82), early change in branches (κ = 0.22 vs. 0.36), and late change in the phylogeny (δ = 2.99 vs. 1.69). For sex differences, white noise still performed

Table 5. A comparison of Pagel's parameters between sexes

	Female		Male		Sex difference	
Test	Value	Log-likelihood	Value	Log-likelihood	Value	Log-likelihood
λ	1	−428.21	1	−429.27	0.829	−368.79
δ	0.658	−427.54	0.670	−428.66	1.691	−399.84
κ	1.363	−428.21	1.244	−429.27	0.399	−376.04

The differences due to heteromorphic sex chromosomes were unlike the phylogenetic patterns for female and male whole-genome size. The λ values show that phylogenetic signal was less for the sex difference, κ values provides support for early evolutionary sex difference change along branches, and the δ values show overall sex chromosome differences late in evolution, whereas the overall change of whole-genome size occurred early in the phylogeny.

as the best model, and OU suggested a directional change. The directional change was still supported with a positive trend and drift value, but with a higher value in the reduced dataset (Supplementary Table S2).

The decrease in size of the Y has limits. Over time, the X and Y become almost entirely different in size and genetic content, which can then result in the loss of the older X–Y system and result in a new neo-Y/neo-X system or an XO system. The evidence provided here suggests that analysis of differences in male and female genome size may be able to identify exceptions to the canonical model of sex chromosome evolution and identify those lineages that have had sex chromosome turnovers or transitions. For example, outside of *Sophophora*, *Hirtodrosophila pictiventris* is known to have an XO sex determination system, and we have found it to have one of the largest differences in female and male genome size (41.1 Mbp, $P < 0.0001$, Table 1, Figure 3) (Clayton 1986). Another example outside of *Sophophora* is *Chymomyza procnemis*. Members of this species group are known to exhibit XO sex determination (Watabe 1998; Matsuda 2002) and a significant female and male difference (74.4 Mbp, $P < 0.0001$, Table 1, Figure 3).

When a neo-Y system emerges, the X–Y difference is expected to be small or negative, because degradation is not expected to occur immediately. The neo-Y chromosome that begins at the size of the X may over time result in males with larger genomes than the females of the same species. This increase in size probably occurs with the inevitable initial inflation of the Y with transposable elements, as is the case in the *Drosophila miranda* neo-Y chromosome. In addition to interacting with the common element A found in *Drosophila* sex chromosomes, the neo-Y system in *D. miranda* is hypothesized to have been formed by a Y-autosome fusion with Muller elements C and D, about 1.2 million years ago (Bachtrog et al. 2008; Matsunaga 2009; Mahajan et al. 2018). This chromosome still harbors many functional genes, yet has more than 20-fold greater accumulation of repetitive sequences than the X chromosome (Bachtrog et al. 2008). Although the male–female difference in *D. miranda* is not significant, the difference is negative, reflecting a male genome larger than the female by 2.9 Mbp. The male–female difference in *D. melanogaster* is similarly nonsignificant, although in the latter, the Y is almost entirely heterochromatic and the male smaller than the female by 4.9 Mbp.

Drosophila pseudoobscura has also been reported to have neo-Y chromosomes but is at a different stage of Y chromosome evolution (reviewed in Carvalho and Clark 2005; Koerich et al. 2008; Carvalho et al. 2009; Bachtrog 2013). This neo-Y system is documented to have resulted from the fusion of the Y to the Muller D element and is ancestrally present in the *affinis* and *pseudoobscura* groups (Carvalho and Clark 2005). The process of becoming heterochromatic from an autosome is estimated to have occurred within 17 million years, a relatively short period of time (Bachtrog 2013). However, we find that the male–female genome size is different in *D. pseudoobscura*, producing a significantly smaller male genome (39.0 Mbp, $P < 0.0001$, Table 1).

If not for the sequence data for *D. pseudoobscura*, the difference in genome size between sexes would not have implicated this species as one with a neo-sex system. There was also no significant difference in the well-studied *D. miranda* neo-Y system. For this reason, it is important to note the drawbacks of using genome size as a measure to identify neo-sex systems. With the constantly expanding mapping data for genomes in *Drosophila*, and other *Diptera* (Blackmon and Demuth 2015b), it is becoming increasingly possible to look at the patterns of fusions/translocations of sex chromosomes with other Muller elements and how such changes may shape the speed by which sex differences may change. However, the lack of Muller element data in a large number of *Drosophila* species inhibits a broad phylogenetic analysis to genome size currently (Crosby et al. 2007).

Interestingly, *Drosophila persimilis* is a species known to hybridize with *D. pseudoobscura* but is characterized by a very different sex difference than that of *D. pseudoobscura*. Evaluation of how differences allow these species to hybridize requires follow-up with crosses and studies of genome size. It is also important to note that genome size may vary intraspecifically. With this in mind, future studies should inspect the genome sizes of *D. persimilis* and *D. pseudoobscura* in regions and within strains in which hybridization is common.

The high δ value in the sex difference analysis found suggests that more change occurs later in the tree or that more change happens in later branches relative to the earlier branches of the phylogeny. Given the rapid changes in genome size between sexes for those species with neo-Y and XO systems, it is not surprising to find evidence for large changes later in the tree. Because the change is occurring late in the phylogeny where most of the radiation of species has occurred, it may be inferred that the change is related to speciation. This idea is not new and has been highlighted in studies of hybrid incompatibility and speciation (Haldane 1922; Johnson and Lachance 2012; Abbott et al. 2017). It has been suggested that sex chromosome turnover is related to speciation and divergence of major mammal groups (Graves 2016) as well as contributing to reproductive isolation in fish (Kitano et al. 2009). The formation of a neo-Y system has the potential to increase instances of reproductive isolation, leading to the formation of new species, but the speculative nature of this statement requires investigation by future studies.

It can be concluded that using either male or female genome size data leads to equivalent results in terms of whole-genome size studies. In contrast, when differences between male and female genomes are examined phylogenetically, support is given to the hypothesis that Y chromosomes have experienced both physical reduction in size and genic degradation of neo-Y chromosomes followed by stasis. The somewhat unpredictable occurrences of neo-Y chromosomes

and XO systems will dramatically shift the expected X–Y difference, probably reducing the phylogenetic signal. This supports our initial hypothesis that the degrading Y chromosome and rapidly changing sex chromosome difference is consistent but unpredictable.

The X–Y sex chromosome system results in significantly different levels of heterochromatin and presumably increased transposable element content. When dissecting components of genome size evolution, such as heterochromatin, repeat, and transposable element content, the differences in sex may result in significant differences in genome size. Also, noteworthy is the observation that these patterns appear consistent throughout the *Sophophora* subgenus, with a few additional outgroups. It will be informative to see whether these patterns are maintained throughout the analysis of the *Drosophila* genus in its entirety, with the inclusion of an equal number of *Drosophila* subgenus species.

Supplementary material

Supplementary data are available at *Journal of Heredity* online.

Funding

C.E.H. was supported by the Texas A&M University Merit Fellowship.

Acknowledgments

We thank Heath Blackmon, Aaron Tarone, Raul Medina, and Mariana Mateos for their comments on the manuscript throughout preparation.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability

Genome size data available in Supplementary Table S1. Accession numbers for phylogeny are available in Hjelmen and Johnston (2017).

References

Abbott JK, Nordén AK, Hansson B. 2017. Sex chromosome evolution: historical insights and future perspectives. *Proc R Soc Biol Sci*. 284: 20162806.

Ågren JA, Wright SI. 2011. Co-evolution between transposable elements and their hosts: a major factor in genome size evolution? *Chromosome Res*. 19:777–786.

Arnqvist G, Sayadi A, Immonen E, Hotzy C, Rankin D, Tuda M, Hjelmen CE, Johnston JS. 2015. Genome size correlates with reproductive fitness in seed beetles. *Proc R Soc B Biol Sci*. 282:20151421.

Bachtrog D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat Rev Genet*. 14:113–124.

Bachtrog D, Hom E, Wong KM, Maside X, de Jong P. 2008. Genomic degradation of a young Y chromosome in *Drosophila miranda*. *Genome Biol*. 9:R30.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B (Methodol)*. 57:289–300.

Bennetzen JL, Kellogg EA. 1997. Do plants have a one-way ticket to genomic obesity? *Plant Cell*. 9:1509–1514.

Blackmon H, Brandvain Y. 2017. Long-term fragility of Y chromosomes is dominated by short-term resolution of sexual antagonism. *Genetics*. 207:1621–1629.

Blackmon H, Demuth JP. 2015a. The fragile Y hypothesis: Y chromosome aneuploidy as a selective pressure in sex chromosome and meiotic mechanism evolution. *Bioessays*. 37:942–950.

Blackmon H, Demuth JP. 2015b. Genomic origins of insect sex chromosomes. *Curr Opin Insect Sci*. 7:45–50.

Butler MA, King AA. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am Nat*. 164:683–695.

Carvalho AB, Clark AG. 2005. Y chromosome of *D. pseudoobscura* is not homologous to the ancestral *Drosophila* Y. *Science*. 307:108–110.

Carvalho AB, Koerich LB, Clark AG. 2009. Origin and evolution of Y chromosomes: *Drosophila* tales. *Trends Genet*. 25:270–277.

Charlesworth B, Charlesworth D. 2000. The degeneration of Y chromosomes. *Philos Trans R Soc Lond B Biol Sci*. 355:1563–1572.

Charlesworth D, Charlesworth B, Marais G. 2005. Steps in the evolution of heteromorphic sex chromosomes. *Heredity (Edinb)*. 95:118–128.

Clayton F. 1986. Overview of chromosomal evolution in the family Drosophilidae. In: Ashburner M, Carson HL, Thompson JN, editors. *The genetics and biology of Drosophila*. 3rd ed. New York: Academic Press. p. 1–38.

Crosby MA, Goodman JL, Strelets VB, Zhang P, Gelbart WM; FlyBase Consortium. 2007. FlyBase: genomes by the dozen. *Nucleic Acids Res*. 35:D486–D491.

Ellis LL, Huang W, Quinn AM, Ahuja A, Alfrejd B, Gomez FE, Hjelmen CE, Moore KL, Mackay TF, Johnston JS, *et al*. 2014. Intrapopulation genome size variation in *D. melanogaster* reflects life history variation and plasticity. *PLoS Genet*. 10:e1004522.

Felsenstein J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. *Am J Hum Genet*. 25:471–492.

Goodfellow P, Banting G, Sheer D, Ropers HH, Caine A, Ferguson-Smith MA, Povey S, Voss R. 1983. Genetic evidence that a Y-linked gene in man is homologous to a gene on the X chromosome. *Nature*. 302:346–349.

Graves JA. 2016. Did sex chromosome turnover promote divergence of the major mammal groups? *Bioessays*. 38:734–743.

Gregory TR. 2001. Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biol Rev Camb Philos Soc*. 76:65–101.

Gregory TR. 2003. Is small indel bias a determinant of genome size? *Trends Genet*. 19:485–488.

Gregory TR. 2004. Insertion-deletion biases and the evolution of genome size. *Gene*. 324:15–34.

Gregory TR. 2015. *Animal genome size database*. 2015. Available from: http://www.genomesize.com

Gregory TR, Hebert PD. 1999. The modulation of DNA content: proximate causes and ultimate consequences. *Genome Res*. 9:317–324.

Gregory TR, Johnston JS. 2008. Genome size diversity in the family Drosophilidae. *Heredity (Edinb)*. 101:228–238.

Haldane JBS. 1922. Sex ratio and unisexual sterility in hybrid animals. *J Genetics*. 12:101–109.

Hanrahan SJ, Johnston JS. 2011. New genome size estimates of 134 species of arthropods. *Chromosome Res*. 19:809–823.

Harmon LJ, Jason TW, Chad DB, Richard EG, Wendell C. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics* 24:129–131.

Hjelmen CE, Johnston JS. 2017. The mode and tempo of genome size evolution in the subgenus Sophophora. *PLoS One*. 12:e0173505.

Huang W, Massouras A, Inoue Y, Peiffer J, Ràmia M, Tarone AM, Turlapati L, Zichner T, Zhu D, Lyman RF, *et al*. 2014. Natural variation in genome architecture among 205 *Drosophila melanogaster* genetic reference panel lines. *Genome Res*. 24:1193–1208.

Johnson NA, Lachance J. 2012. The genetics of sex chromosomes: evolution and implications for hybrid incompatibility. *Ann N Y Acad Sci*. 1256:E1–E22.

Johnston JS, Bernardini A, Hjelmen CE. 2019. Genome size estimation and quantitative cytogenetics in insects. In: Brown SJ, Pfrender ME, editors. *Insect Genomics*. New York, NY: Humana Press. p. 15–26.

Kapusta A, Suh A, Feschotte C. 2017. Dynamics of genome size evolution in birds and mammals. *Proc Natl Acad Sci USA*. 114:E1460–E1469.

Kellermann V, Loeschcke V, Hoffmann AA, Kristensen TN, Fløjgaard C, David JR, Svenning JC, Overgaard J. 2012. Phylogenetic constraints in key

functional traits behind species' climate niches: patterns of desiccation and cold resistance across 95 *Drosophila* species. *Evolution*. 66:3377–3389.

Kellermann V, Overgaard J, Hoffmann AA, Fløjgaard C, Svenning JC, Loeschcke V. 2012. Upper thermal limits of *Drosophila* are linked to species distributions and strongly constrained phylogenetically. *Proc Natl Acad Sci USA*. 109:16228–16233.

Kelley JL, Peyton JT, Fiston-Lavier AS, Teets NM, Yee MC, Johnston JS, Bustamante CD, Lee RE, Denlinger DL. 2014. Compact genome of the Antarctic midge is likely an adaptation to an extreme environment. *Nat Commun*. 5:4611.

Kidwell MG. 2002. Transposable elements and the evolution of genome size in eukaryotes. *Genetica*. 115:49–63.

Kitano J, Ross JA, Mori S, Kume M, Jones FC, Chan YF, Absher DM, Grimwood J, Schmutz J, Myers RM, *et al*. 2009. A role for a neo-sex chromosome in stickleback speciation. *Nature*. 461:1079–1083.

Koerich LB, Wang X, Clark AG, Carvalho AB. 2008. Low conservation of gene content in the *Drosophila* Y chromosome. *Nature*. 456:949–951.

Lahn BT, Page DC. 1999. Four evolutionary strata on the human X chromosome. *Science*. 286:964–967.

Lynch M, Conery JS. 2003. The origins of genome complexity. *Science*. 302:1401–1404.

Mahajan S, Wei KH, Nalley MJ, Gibilisco L, Bachtrog D. 2018. De novo assembly of a young *Drosophila* Y chromosome using single-molecule sequencing and chromatin conformation capture. *PLoS Biol*. 16:e2006348.

Matsuda M. 2002. Meiotic chromosome configurations of XY and "XO" type males from four species of the genus *Chymomyza* (Diptera, Drosophilidae) with a description of the karyotype of *C. rufithorax*. *Cytologia*. 67:383–388.

Matsunaga S. 2009. Junk DNA promotes sex chromosome evolution. *Heredity (Edinb)*. 102:525–526.

Mirsky AE, Ris H. 1951. The desoxyribonucleic acid content of animal cells and its evolutionary significance. *J Gen Physiol*. 34:451–462.

Muller HJ. 1918. Genetic variability, twin hybrids and constant hybrids, in a case of balanced lethal factors. *Genetics*. 3:422–499.

Münkemüller T, Lavergne S, Bzeznik B, Dray S, Jombart T, Schiffers K, Thuiller W. 2012. How to measure and test phylogenetic signal. *Methods Ecol Evol*. 3:743–756.

Orme D, Freckleton R, Thomas G, Petzoldt T, Fritz S, Isaac N, Pearse W. 2018. *Caper: comparative analyses of phylogenetics and Evolution in R*. R package version 1.0.1. Available from: https://CRAN.R-project.org/package=caper

Pagel M. 1999. Inferring the historical patterns of biological evolution. *Nature*. 401:877–884.

Palazzo AF, Gregory TR. 2014. The case for junk DNA. *PLoS Genet*. 10:e1004351.

Petrov DA. 2001. Evolution of genome size: new approaches to an old problem. *Trends Genet*. 17:23–28.

Petrov DA. 2002a. DNA loss and evolution of genome size in *Drosophila*. *Genetica*. 115:81–91.

Petrov DA. 2002b. Mutational equilibrium model of genome size evolution. *Theor Popul Biol*. 61:531–544.

Petrov DA, Sangster TA, Johnston JS, Hartl DL, Shaw KL. 2000. Evidence for DNA loss as a determinant of genome size. *Science*. 287:1060–1062.

Picard CJ, Johnston JS, Tarone AM. 2012. Genome sizes of forensically relevant Diptera. *J Med Entomol*. 49:192–197.

R Core Team. 2016. *R: a language and environment for statistical computing*. Vienna (Austria): R Foundation for Statistical Computing. Available from: http.//www R-project org.

Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol*. 3:217–223.

SAS S, Version S. 2003. *9.4 [Computer Program]*. Cary (NC): SAS Institute.

Schaeffer SW, Bhutkar A, McAllister BF, Matsuda M, Matzkin LM, O'Grady PM, Rohde C, Valente VL, Aguadé M, Anderson WW, *et al*. 2008. Polytene chromosomal maps of 11 *Drosophila* species: the order of genomic scaffolds inferred from genetic and physical maps. *Genetics*. 179:1601–1655.

Schielzeth H, Streitner C, Lampe U, Franzke A, Reinhold K. 2014. Genome size variation affects song attractiveness in grasshoppers: evidence for sexual selection against large genomes. *Evolution*. 68:3629–3635.

Sessegolo C, Burlet N, Haudry A. 2016. Strong phylogenetic inertia on genome size and transposable element content among 26 species of flies. *Biol Lett*. 12:20160407.

van der Linde K, Houle D, Spicer GS, Steppan SJ. 2010. A supermatrix-based molecular phylogeny of the family Drosophilidae. *Genet Res (Camb)*. 92:25–38.

Vieira C, Biémont C. 2004. Transposable element dynamics in two sibling species: *Drosophila melanogaster* and *Drosophila simulans*. *Genetica*. 120:115–123.

Watabe H-A. 1998. XO type of sex determination in the genus *Chymomyza* (Diptera: Drosophilidae). *Entomol Sci*. 1:309–311.

Whitney KD, Garland T Jr. 2010. Did genetic drift drive increases in genome complexity? *PLoS Genetics*. 6:e1001080.