doi:10.1111/evo.14

Thoracic underreplication in *Drosophila* species estimates a minimum genome size and the dynamics of added DNA

Carl E. Hjelmen,[1,2,3] Valerie Renee Holmes,[2] Crystal G. Burrus,[1] Elizabeth Piron,[4] Melissa Mynes,[4] Margaret A. Garrett,[4] Heath Blackmon,[1] and John Spencer Johnston[2]

[1]*Department of Biology, Texas A&M University, College Station, Texas*

[2]*Department of Entomology, Texas A&M University, College Station, Texas*

[3]*E-mail: cehjelmen@gmail.com*

[4]*Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas*

Received January 28, 2020
Accepted May 17, 2020

Many cells in the thorax of *Drosophila* were found to stall during replication, a phenomenon known as underreplication. Unlike underreplication in nuclei of salivary and follicle cells, this stall occurs with less than one complete round of replication. This stall point allows precise estimations of early-replicating euchromatin and late-replicating heterochromatin regions, providing a powerful tool to investigate the dynamics of structural change across the genome. We measure underreplication in 132 species across the *Drosophila* genus and leverage these data to propose a model for estimating the rate at which additional DNA is accumulated as heterochromatin and euchromatin and also predict the minimum genome size for *Drosophila*. According to comparative phylogenetic approaches, the rates of change of heterochromatin differ strikingly between *Drosophila* subgenera. Although these subgenera differ in karyotype, there were no differences by chromosome number, suggesting other structural changes may influence accumulation of heterochromatin. Measurements were taken for both sexes, allowing the visualization of genome size and heterochromatin changes for the hypothetical path of XY sex chromosome differentiation. Additionally, the model presented here estimates a minimum genome size in *Sophophora* remarkably close to the smallest insect genome measured to date, in a species over 200 million years diverged from *Drosophila*.

KEY WORDS: *Drosophila*, genome size, heterochromatin, replication, sex chromosomes.

The amount of DNA in an organism, while consistent within a species, is highly variable across the tree of life (Gregory 2005, 2001; Palazzo and Gregory 2014). To date, the smallest insect genome sequenced is the Antarctic midge, *Belgica antarctica*, with a miniscule genome of only 98 Mbp (Kelley et al. 2014). This is a far cry from the massive genomes of grasshoppers surpassing 16,000 Mbp (Gregory 2020). We know this variation is largely due to highly repetitive regions and mobile elements, which are notoriously difficult to sequence and assemble. Without sequencing, it can be difficult to determine components of chromatin structure, amounts of heterochromatin, and the patterns of change across large groups of organisms. Further, yet unanswered is how small we can expect a genome to be when it does not have all of this additional, largely repetitive DNA and we do not know the rate at which extra DNA beyond this minimum size contributes to chromatin structure as euchromatin or heterochromatin.

Historically, chromatin has been described in two cytological forms: heterochromatin and euchromatin (Heitz 1928, 1929; Passarge 1979; Redi et al. 2001). Heterochromatin is typically identified by dark staining of chromosomes, indicating that DNA is tightly packed; whereas euchromatin stains less intensely and is considered to be more "open" than heterochromatin. Typically, euchromatin is gene rich and contains genes that are highly expressed, while heterochromatin is gene poor and rarely transcribed. When DNA is replicated, heterochromatin is typically

replicated last (Barigozzi et al. 1966; Lima-de-Faria and Jaworska 1968), while euchromatic DNA is replicated early (Schübeler et al. 2002; McNairn and Gilbert 2003; MacAlpine et al. 2004; Schwaiger and Schübeler 2006). Early replicating portions of the genome have synonymous substitutions and low mutation rates, while instability of replication forks in late-replicating heterochromatic portions contribute to fast-evolving genes and high mutation rates (Wolfe et al. 1989; Schwaiger and Schübeler 2006; Makunin et al. 2014).

Amounts of heterochromatin and euchromatin in the genomes of flies (Diptera) have been investigated through visualizing the giant polytene chromosomes of salivary glands. The chromosomes of Dipteran salivary glands become polyploid and are concurrently polytenized through a process called endoreplication. In some cases, there are portions of the genome that have endoreplicated less than other regions, and are known to be "underreplicated." Underreplication is where there is a stall in the S phase before DNA can fully replicate (Leach et al. 2000). This results in later replicating regions of the genome, such as pericentric and intercalary heterochromatin, to be underreplicated (Belyaeva et al. 1998), likely due to reduced origin firing frequency or stalling of replication forks through heterochromatin in each endocycle S phase (Hua et al. 2014; Hua and Orr-Weaver 2017). Generally this process has been associated with salivary glands (Rudkin 1969; Hammond and Laird 1985) and nurse cells of *Drosophila* and other Diptera (Painter and Reindorp 1939), in which DNA undergoes many rounds of underreplication. A protein, Suppressor of Underreplication (SuUR), has been found in *Drosophila* to be involved in the suppression of replication (Belyaeva et al. 1998). When SuUR is nonfunctional, replication of heterochromatin is nearly complete, and when overexpressed, more heterochromatic regions were underreplicated (Zhimulev et al. 2003). This protein was found to be "fast-evolving" when comparing across *Drosophila* species, and orthologs could be found in mosquito species, but not outside of Diptera (Yurlova et al. 2010). This suggests that this phenomenon within insects is exclusive to flies.

Uniquely, underreplication was found to occur in the thorax of *Drosophila melanogaster* (Johnston et al. 2013, 2020). Unlike salivary glands and nurse cells, thoracic tissue only undergoes one round of underreplication. The mechanism of underreplication appears to be shared across these tissues, as *D. melanogaster* SuUR mutants were found to have almost no instance of thoracic underreplication (Johnston et al. 2013). Until now, thoracic underreplication had only been identified in *D. melanogaster* and *Drosophila virilis*. We estimate thoracic underreplication for 132 species within the genus *Drosophila*. We leverage these amounts of thoracic underreplication to estimate a minimum expected genome size for both males and females, as well as estimate the fate (euchromatic or heterochromatic) of DNA that is added to the genome. Additionally, we investigate the dynamics of heterochromatin change in respect to genome size evolution across the *Drosophila* genus and between its subgenera and sexes.

Materials and Methods

ESTIMATIONS OF GENOME SIZE AND UNDERREPLICATION

All genome size estimates for female and male *Drosophila* species were taken from (Hjelmen et al. 2019). Species were taken from the *Drosophila* Species Stock Center originally located in San Diego, but recently moved to Cornell (https://blogs.cornell.edu/drosophila/). Underreplication was measured for each of these species and was estimated as in Johnston et al. (2013, 2020). Thoraces of both male and female species of *Drosophila* were dissected and placed in a 2-mL Kontes Dounce tube with 1 mL of Galbraith buffer. Each thorax was then carefully ground with a "loose" A pestle approximately 10-15 times to ensure proper release of nuclei from the thoracic tissue. Each sample was then passed through a 50-micron filter before 25 μL of 1 mg/μL propidium iodide was added for staining. Samples were allowed to incubate with the propidium iodide for at least 20 min to ensure adequate staining had occurred.

Samples were run on a Partec CyFlow SL_3 cytometer with a 532-nm green laser and a Beckman Coulter Cytoflex flow cytometer. Peaks for 2C, 4C, and underreplication were gated to approximate a mean to calculate the percent replication that had occurred for each individual. Underreplication was calculated by subtracting the 2C value from the underreplication (UR) value, then divide by the 2C value. This process was repeated for at least five individuals for each sex within each species to produce a standard error that is less than 1% of the estimated mean. Underreplication values were averaged across replicates for each sex in each species (Females, $n = 132$; Males, $n = 123$). The amount of replicated DNA was calculated by multiplying the estimated percent replication by the genome size. The amount of unreplicated DNA, or the amount of DNA that is not replicated to the 4C amount, was then calculated by subtracting the replicated DNA amount from the whole genome size.

LINEAR MODELS AND STATISTICS

Values for percent replication, replicated DNA, and unreplicated DNA amounts were regressed against genome size for each sex in R version 3.6.0 (R Core Team 2016). These were repeated for species in each subgenus (*Sophophora* and *Drosophila*) and for comparisons between species with differing chromosome numbers. Phylogenetic signal for residuals of each regression model was estimated using the phylosig command in the package phytools (Revell 2012). Phylogenetic generalized least squares models were performed using the gls function in the package

nlme (Pinheiro et al. 2012) with the phylogeny as a correlation structure using the corBrownian function from the package ape (Paradis et al. 2004). The linear model for unreplicated DNA was used to estimate minimum genome size. To do this, we solved for genome size when unreplicated amount was set to zero (Eq. 1):

$$U = \beta G + \varepsilon, \quad (1)$$

where U represents unreplicated DNA, G is genome size, ε is y-intercept, and β is the coefficient estimated for the predictor variable. The slope of the line in the linear model for unreplicated DNA was used as an estimate for the proportion of additional DNA that is heterochromatic, whereas the slope of the line for the replicated DNA linear model was used as an estimate for the proportion of additional DNA that is euchromatic.

To investigate the impact of sex chromosome differentiation on heterochromatin, the sex difference in genome size for each species was calculated by subtracting the diploid male genome size from the diploid female genome size. The sex difference in heterochromatin was calculated by subtracting the diploid unreplicated male DNA from the diploid unreplicated female DNA. The sex difference in unreplicated DNA was regressed against the sex difference in genome size.

COMPARATIVE PHYLOGENETIC METHODS

The *Drosophila* phylogeny was previously estimated by Hjelmen et al. (2019). The drop.tip function in ape was used to remove any taxa for which underreplication values were not available (Paradis et al. 2004). The drop.tip function was also used to separate the *Drosophila* subgenus phylogeny from the *Sophophora* subgenus phylogeny for further comparative analyses. Amounts of unreplicated and replicated DNA for both sexes, as well as replicated and unreplicated sex differences, were analyzed using the fitContinous function in the package geiger in R version 3.6.0 (Harmon et al. 2009). These analyses allowed for comparison of Brownian motion, Ornstein-Uhlenbeck, and white noise models using AICc (Akaike Information Criterion) values. In addition to these models, fitContinous was used to estimate Pagel's λ for each trait (Pagel 1999). All traits were visualized on the phylogeny using the contMap function in the package phytools (Revell 2012). To ensure results were robust to phylogenetic uncertainty, 100 trees were randomly selected from the Bayesian distribution of trees using the SampleTrees function in the package evobiR for further analyses with fitContinous (Blackmon and Adams 2015).

Patterns of trait evolution in unreplicated DNA and replicated DNA for males and females of each species were investigated using BAMM (Bayesian Analysis of Macroevolutionary Mixtures) in which our traits evolved by Brownian motion allowing for the possibility of rate shifts in our tree. We used the setBAMMpriors function in the BAMMtools package in R version 3.6.0 to set priors for the number of shifts for each trait of interest (Rabosky et al. 2014). This function was used as it constructs a relatively flat prior for the trait evolution parameter. Each BAMM analysis ran for 10,000,000 generations with a 10% burn-in to ensure sufficient effective sizes. The R package coda was used to ensure that all runs reached convergence (Plummer et al. 2006). For each trait, credible rate shift sets for the most likely number of shifts were estimated. Rate of change in each trait was estimated throughout evolutionary time across the entire genus and each subgenus.

Results

ESTIMATES OF THORACIC REPLICATION

Percent thoracic replication was estimated for females of 132 species and males of 123 species (Table S1). The proportion of replication in the different species varied from 41.2% to 99.2% (Table S1). Because underreplication was not found to occur in the outgroup species *Chymomyza* sp., *Scaptodrosophila pattersoni*, *Scaptodrosophila lebanonensis*, and *Hirtodrosophila pictiventris*, these were trimmed from the phylogeny and excluded from analyses. Thoracic replication was found to be negatively correlated with genome size in both sexes (Female: $P < 0.0001$, $R^2 = 0.409$; Male: $P \leq 0.0001$, $R^2 = 0.349$), with percent of the genome replicated decreasing as genome size increases (Figs. S1 and S2). Species with larger genomes had more heterochromatin and stalled earlier in S phase replication.

PROPORTION OF ADDITIONAL DNA ADDED AS HETEROCHROMATIN

The amount of DNA replicated above 2C and, consequently, the amount of DNA not replicated to 4C (referred to as "unreplicated") in megabase pairs (Mbp) were calculated for each sex and can be found in Table S1. We found significant phylogenetic signal in the residuals of the regression models. When comparing the results of the Phylogenetic Generalized Least Squares (PGLS) to linear regression models, we found quantitative, but not qualitative differences in the results. In no model were there differences in the direction of the regression coefficient or the relationships between the subgenera or early and late heterochromatin. To discuss the results in the most simplistic manner, we present the data from the noncorrected model and have the phylogenetically corrected results in Table S2. In linear regression models, both replicated and unreplicated DNA were found to be significantly positively correlated with genome size in each sex (Tables 1 and S3-S6; Figs. S3-S6). Slope and intercept for each of these linear models can be found in Table 1. In the case of unreplicated DNA, slope can be interpreted as the proportion of additional DNA that becomes late-replicating heterochromatin,

Table 1. Linear models and minimum genome size.

Group	Sex	Portion	*P*-value	R^2	Slope	*y*-intercept	Minimum Genome Size (Mbp)
Drosophila genus	Female	Unrep. DNA	<0.0001	0.618	0.617	–88.004	142.6
		Rep. DNA	<0.0001	0.383	0.383	87.992	
	Male	Unrep. DNA	<0.0001	0.625	0.664	−83.59	125.9
		Rep. DNA	<0.0001	0.295	0.336	83.59	
Sophophora subgenus	Female	Unrep. DNA	<0.0001	0.375	0.332	−30.57	92.1
		Rep. DNA	<0.0001	0.713	0.668	30.57	
	Male	Unrep. DNA	<0.0001	0.387	0.387	−33.544	86.7
		Rep. DNA	<0.0001	0.616	0.613	33.539	
Drosophila subgenus	Female	Unrep. DNA	<0.0001	0.801	0.794	−120.254	151.5
		Rep. DNA	<0.0001	0.204	0.206	120.234	
	Male	Unrep. DNA	<0.0001	0.817	0.801	−103.82	129.6
		Rep. DNA	<0.0001	0.206	0.199	103.817	

whereas in replicated DNA, slope can be interpreted as the proportion of additional DNA that becomes early-replicating, mostly euchromatic DNA. In the *Drosophila* genus, it was found that approximately 61.6% of additional DNA in females and 66.4% of additional DNA in males becomes late replicating heterochromatin. Inversely, 38.4% of additional DNA in females and 33.6% of additional DNA in males becomes early-replicating (Table 1). When separated by subgenus, it was found that 79.4% of additional DNA in the *Drosophila* subgenus females (Fig. 1) and 80.1% of additional DNA in *Drosophila* subgenus males (Fig. S7) is late-replicating, whereas 33.2% of additional DNA becomes late-replicating in the *Sophophora* females (Fig. 1) and 38.7% in *Sophophora* subgenus males (Fig. S7) (Table 1).

MINIMUM GENOME SIZE ESTIMATES

Minimum genome size was estimated using the parameters from the linear model for unreplicated DNA as a function of genome size (Eq. 1). Reasoning that the minimal genome would require all DNA to be euchromatic and early replicating, we set the unreplicated DNA to zero and then solved for genome size. Using data for all available *Drosophila* species, a minimum genome size of 142.7 Mbp was estimated for females and 125.8 Mbp for males (Table 1). When separated by subgenus, a minimum genome size of 151.4 Mbp was estimated for the *Drosophila* subgenus females and 129.57 Mbp for *Drosophila* males. A minimum genome size of 92.2 Mbp was estimated for *Sophophora* females and a minimum genome size of 86.6 Mbp was estimated for *Sophophora* males (Table 1). The average of estimated male minimum genome sizes in all cases was found to be lower than those of females, likely due to the presence of two X chromosomes in females, and one X and a reduced Y chromosome in males. Although significant differences were found between estimates for subgenera, the difference was not due to differences in chromosome number. Significant differences in slope were not found when species were separated by chromosome number.

LATE-REPLICATING HETEROCHROMATIN AND DIFFERENTIATION OF SEX CHROMOSOMES

As a proxy for sex chromosome differentiation in the X-Y system of *Drosophila*, the difference in whole genome sizes between sexes was calculated, as well as the difference in the amount of unreplicated DNA between the sexes. Positive values indicate that the female value is larger than the male value (X chromosome large relative to Y chromosome), whereas negative values indicate males have larger values than females (Y chromosome large relative to X chromosome). In most cases, females have larger genome sizes than males, indicating that the X chromosome is larger than the Y chromosome. Sex difference in unreplicated DNA was plotted against sex difference in whole genome size (Fig. 2). The figure shows the change in X and Y chromatin structure. In terms of unreplicated DNA, most cases indicate that the Y chromosome has more heterochromatin than the X chromosome. Most cases of species with neo-sex systems were found to have larger male genomes than female genomes and higher late-replicating heterochromatin content on the Y than the X.

COMPARATIVE PHYLOGENETIC ANALYSES

To test the robustness of our results, we compared results from the 100 randomly selected trees to the results from the consensus tree. We found quantitative, but not qualitative, differences in the results in comparison to the results from the consensus tree. For this reason, all the results reported are those from the consensus tree. According to analyses with fitContinous, there was evidence for phylogenetic signal in each of the traits investigated, and Ornstein-Uhlenbeck was found to be the best-fit model of change in chromatin structure (Table S3). Values for α and σ^2

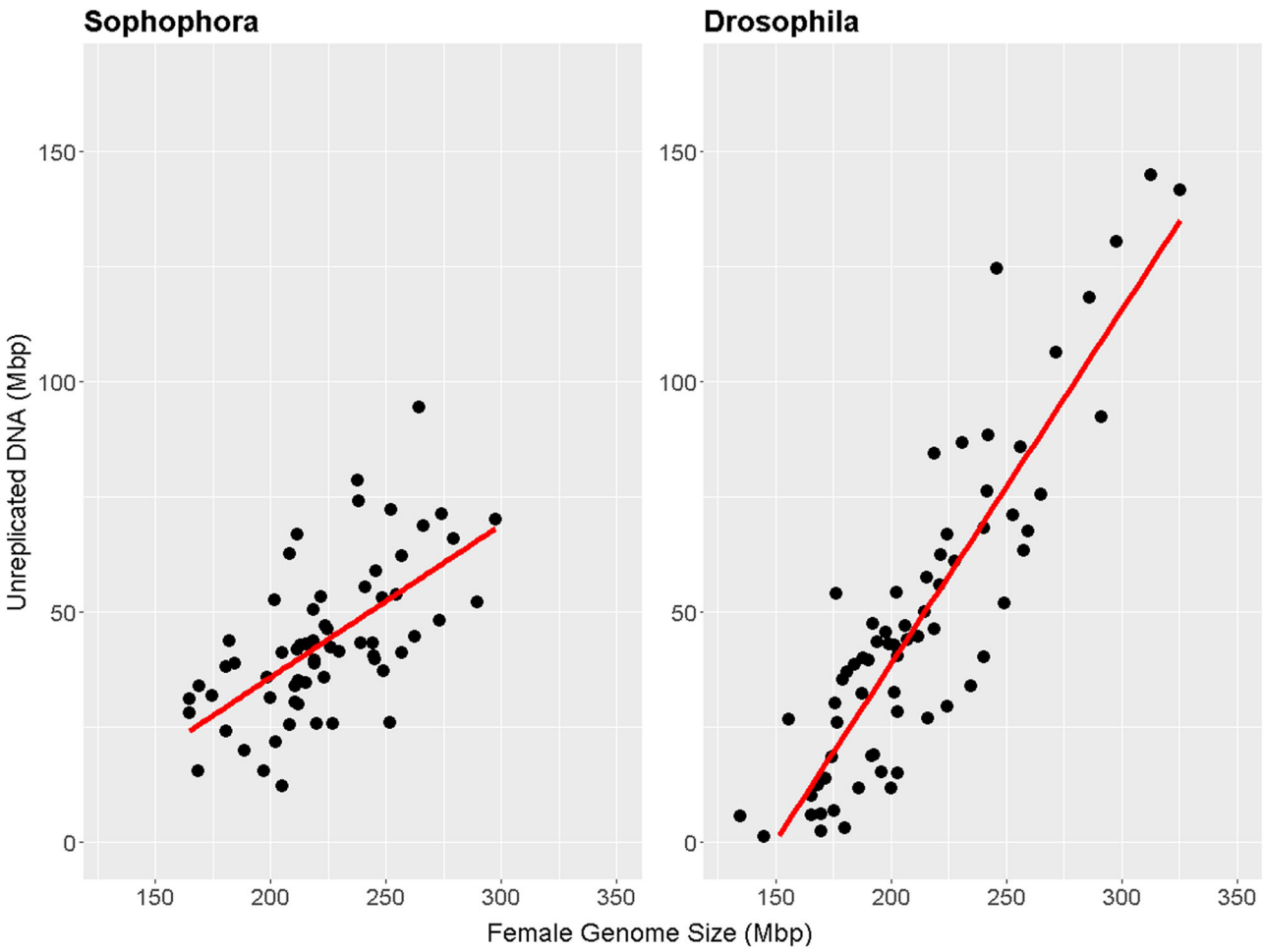

Figure 1. **Unreplicated DNA model for *Sophophora* and *Drosophila* females. Unreplicated DNA (Mbp) (vertical axis) plotted against genome size (horizontal axis) for females of *Sophophora* (left) and *Drosophila* (Right) subgenera. Slope of trend estimates percentage of additional DNA, which is late-replicating heterochromatic.**

parameter were measured for each trait to estimate the strength of selection (α) and the rate of change or drift (σ^2) (Table 2). Across the *Drosophila* genus, higher rates of change (σ^2) were seen in unreplicated DNA than in replicated DNA, suggesting more change in heterochromatic DNA content. This pattern was mirrored in the *Drosophila* subgenus, but with a much higher difference between rate of change in unreplicated versus replicated DNA, suggesting high rates of variation in late-replicating DNA in the *Drosophila* subgenus. There were no remarkable differences in the rate of change of replicated and unreplicated DNA in the *Sophophora* subgenus, yet there was higher strength of selection found in unreplicated DNA than there was in replicated DNA (Table 2). Regarding sex differences, a higher strength of selection was found in the replicated DNA for the *Drosophila* genus and the *Drosophila* subgenus, with similar σ^2 values between unreplicated and replicated sex differences (Table 2). In *Sophophora*, α and σ^2 values were higher in unreplicated sex differences than in replicated sex differences (Table 2). *Sophophora* σ^2 values were higher than either the *Drosophila* subgenus or *Drosophila* genus as a whole, suggesting higher rates of change in overall sex differences in the *Sophophora* subgenus. When comparing the differences results of fitContinuous analysis between the Ornstein-Uhlenbeck and Brownian motion models of trait evolution, there were no qualitative differences in the interpretation of the results, aside the lack of the α parameter in the BM model, suggesting these results are robust across evolutionary models.

Z_0 values estimate a trait value that evolution appears to select for according to the Ornstein-Uhlenbeck model. These Z_0 values were estimated for chromatin structure for each sex across the phylogeny, as well as within each subgenus (Table 2). Z_0 total was calculated by totaling Z_0 for unreplicated and replicated DNA for the sex of each group. Female Z_0 total values were found to be consistently higher than males, supporting females with XX having higher genome sizes than males with XY. No remarkable differences were found in Z_0 total values across any group within

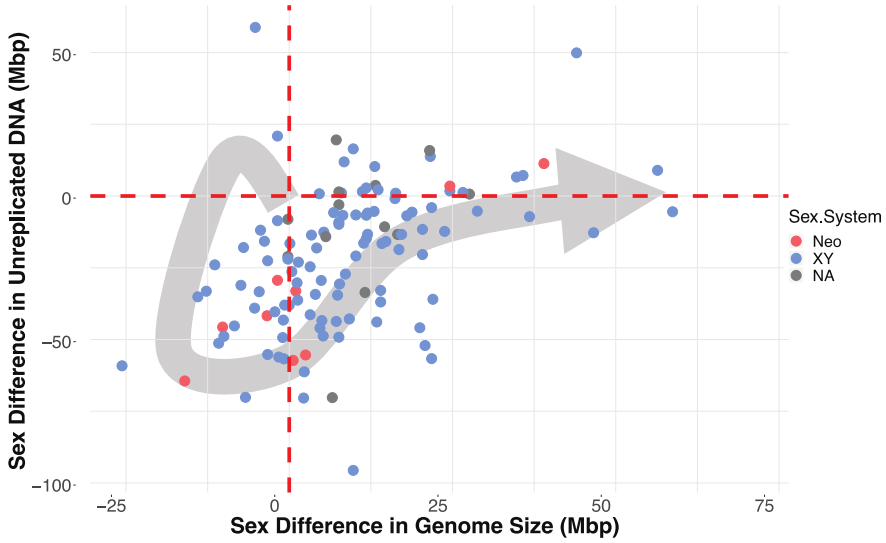

Figure 2. Sex difference in unreplicated DNA against sex difference in genome size suggests pattern for sex chromosome differentiation. Values for sex difference in unreplicated DNA (vertical axis) were plotted against sex difference in genome size (horizontal axis). Each point represents a species and is colored according to documented sex system (XY, Neo-Sex, or NA-indicating not documented). The top left quadrant represents species in which the Y chromosome is larger than the X chromosome and has less late-replicating heterochromatin than the X chromosome. The bottom left quadrant includes species in which the Y chromosome is larger than the X chromosome and has more late-replicating heterochromatin than the X. The bottom right quadrant includes species that have a Y smaller than the X chromosome but which has more late-replicating heterochromatin than the X. The top right quadrant includes species that have larger X chromosomes than Y chromosomes and have more late-replicating heterochromatin on the X chromosome than the Y chromosome. The gray line with arrows indicated the hypothesized path from undifferentiated autosomes to fully differentiated, old sex chromosomes.

Table 2. Ornstein-Uhlenbeck values for each trait.

Trait	Group	Sex	Portion	α	σ^2	Z_0	Z_0 total
Genome size	*Drosophila* genus	Female	Unrep. DNA	0.05	98.77	47.05	215.09
			Rep. DNA	0.032	32.37	168.04	
		Male	Unrep. DNA	0.036	88.7	60.9	211.14
			Rep. DNA	0.028	23.78	150.24	
	Sophophora subgenus	Female	Unrep. DNA	0.071	41.77	43.95	224.00
			Rep. DNA	0.029	38.76	180.05	
		Male	Unrep. DNA	0.04	35.3	52.5	216.00
			Rep. DNA	0.027	37.57	163.5	
	Drosophila subgenus	Female	Unrep. DNA	2.69	7225.2	49.5	212.60
			Rep. DNA	0.035	25.05	163.1	
		Male	Unrep. DNA	0.061	204.28	65.53	209.34
			Rep. DNA	7.66E-218	7.125	143.81	
Sex difference	*Drosophila* genus	–	Unrep. DNA	0.066	86.35	−24.35	8.99
			Rep. DNA	0.117	100.77	33.34	
	Sophophora subgenus	–	Unrep. DNA	0.2	161.53	−15.3	13.48
			Rep. DNA	0.156	129.75	28.78	
	Drosophila subgenus	–	Unrep. DNA	0.066	100.3	−27.44	8.57
			Rep. DNA	0.119	97.28	36.01	

each sex; however, females and males within the *Sophophora* subgenus had higher proportions of replicated DNA than their counterparts did in the *Drosophila* subgenus (Table 2). When investigating sex difference, the *Sophophora* subgenus had a higher Z_0 total value than did the *Drosophila* subgenus (Table 2). This result suggests that Y-chromosomes in *Sophophora* species are proportionally smaller than X-chromosomes and smaller than would be expected for the XY relationship in *Drosophila* species.

Table 3. Number of most likely rate shifts according to BAMM.

	Female rep.			Female unrep.			Male rep.			Male unrep.		
Rank	Prob.	Total	Shifts	Prob.	Total	Shifts	Prob.	Total	Shifts	Prob.	Total	Shifts
1	52.7%	52.7%	0	0.2%	0.2%	3	54.8%	54.8%	1	2.6%	2.6%	4
2	10.4%	63.1%	1	0.1%	0.3%	4	12.0%	66.8%	1	2.5%	5.2%	3
3	3.5%	66.6%	2	0.1%	0.4%	5	11.9%	78.6%	1	1.3%	6.5%	4
4	1.9%	68.5%	1	0.1%	0.5%	4	6.4%	85.0%	1	1.2%	7.7%	3
5	1.8%	70.3%	1	0.1%	0.7%	2	3.1%	88.1%	1	1.0%	8.8%	1

BAMM analyses found that the most likely number of rate shifts for replicated DNA in the *Drosophila* phylogeny is zero in females (52.7% support) and one in males (54.8% support) (Figs. S8 and S9; Table 3). The estimated rate shift in male replicated DNA most often occurs at the split of the *Sophophora* and *Drosophila* subgenus (Fig. S9). When investigated by subgenus, there is a higher rate of change in replicated DNA in the *Sophophora* subgenus than in the *Drosophila* subgenus for both female and male data (Figs. 3A and S9). In the case of unreplicated DNA, the most likely number of shifts is 3 for females and 4 for males; however, predicted numbers of shifts in unreplicated DNA have low amounts of support suggesting high variability in the trait (Table 3; Figs. S11 and S12). When investigated by subgenus, the rates of change of unreplicated DNA in *Drosophila* are much higher than those in *Sophophora* (Figs. 3B and S12). When replicated and unreplicated DNA are plotted in color on the phylogeny, more variation in replicated DNA is seen in *Sophophora*, whereas more variation in unreplicated DNA is seen in the *Drosophila* (Figs. 4 and S13-S17).

Discussion

Underreplication, or stalling of replication before completion of late-replicating heterochromatic regions, has been reported in salivary glands and follicle cells of *Drosophila* as well as in other dipteran species (Painter and Reindorp 1939; Rudkin 1969; Renkawitz-Pohl and Kunz 1975; Hammond and Laird 1985; Belyaeva et al. 2006; Andreyeva et al. 2008). In these tissues, the genome undergoes multiple rounds of incomplete replication. Uniquely, however, one round of underreplication had been found to occur in the thoracic tissue of *D. melanogaster* and *D. virilis* (Johnston et al. 2013, 2020). To explore this phenomenon, thoracic replication values have been reported for females of 132 species and males of 123 species within the *Drosophila* genus (Table S1). Although underreplication is known to occur across Diptera (Yurlova et al. 2010), it is unknown to what extent thoracic underreplication occurs across the insect tree of life. Thoracic underreplication was not found in tested *Chymomyza* species or in *Scaptodrosophila lebanonensis*, *Scaptodrosophila pattersoni*, and *Scaptodrosophila stonei* (Table S1). These results suggest that this phenomenon may be largely restricted to the *Drosophila* genus of flies.

Underreplication in polytene chromosomes has been suggested to be representative of the proportion of heterochromatin across the entire genome (Bosco et al. 2007). Following this suggestion, we represent the percent underreplication in the thorax as the proportion of heterochromatin in the genome and further compare this percentage against genome size. Although Bosco et al. report a relationship between genome size and underreplication for 38 species, all estimations are based on assumptions regarding ploidy level and genome size of *D. virilis*, as well as the assumption of a constant fluorescence ratio of 4′,6-diamidino-2-phenylindole (DAPI) and Propidium Iodide (PI) stained nuclei. Estimates of underreplication in the thoracic tissue provide more precision than do estimates in the polytene cells because thoracic underreplication measures provide a direct comparison of peaks of 2C and underreplicated DNA. Underreplication based on polytene cells requires estimates of GC content and comparisons against *D. melanogaster* for each species and strain. The amount of thoracic replication was found to be significantly correlated with genome size, suggesting as genome sizes get larger, a lower proportion of the genome is replicated early ($P < .0001$). The fit of the relationship between replication percent and genome size is higher in females than in males of the species (Female $R^2 = 0.409$; Male $R^2 = 0.349$), which is not surprising given the highly variable size and heterochromatin content of Y-chromosomes found in males. Although this relationship of underreplication and genome size has been reported (Bosco et al. 2007), the implications of this relationship had not been thoroughly investigated. With this wealth of information on estimated heterochromatin content across a range of *Drosophila* species, it has become possible to estimate the expected heterochromatin content for each genome size, the minimum expected genome size, and the rate at which heterochromatin is added to the genome.

SUGGESTED PATHWAY FOR SEX CHROMOSOME DIFFERENTIATION

Our linear model for unreplicated DNA content found that expected minimum genome sizes were smaller for males than

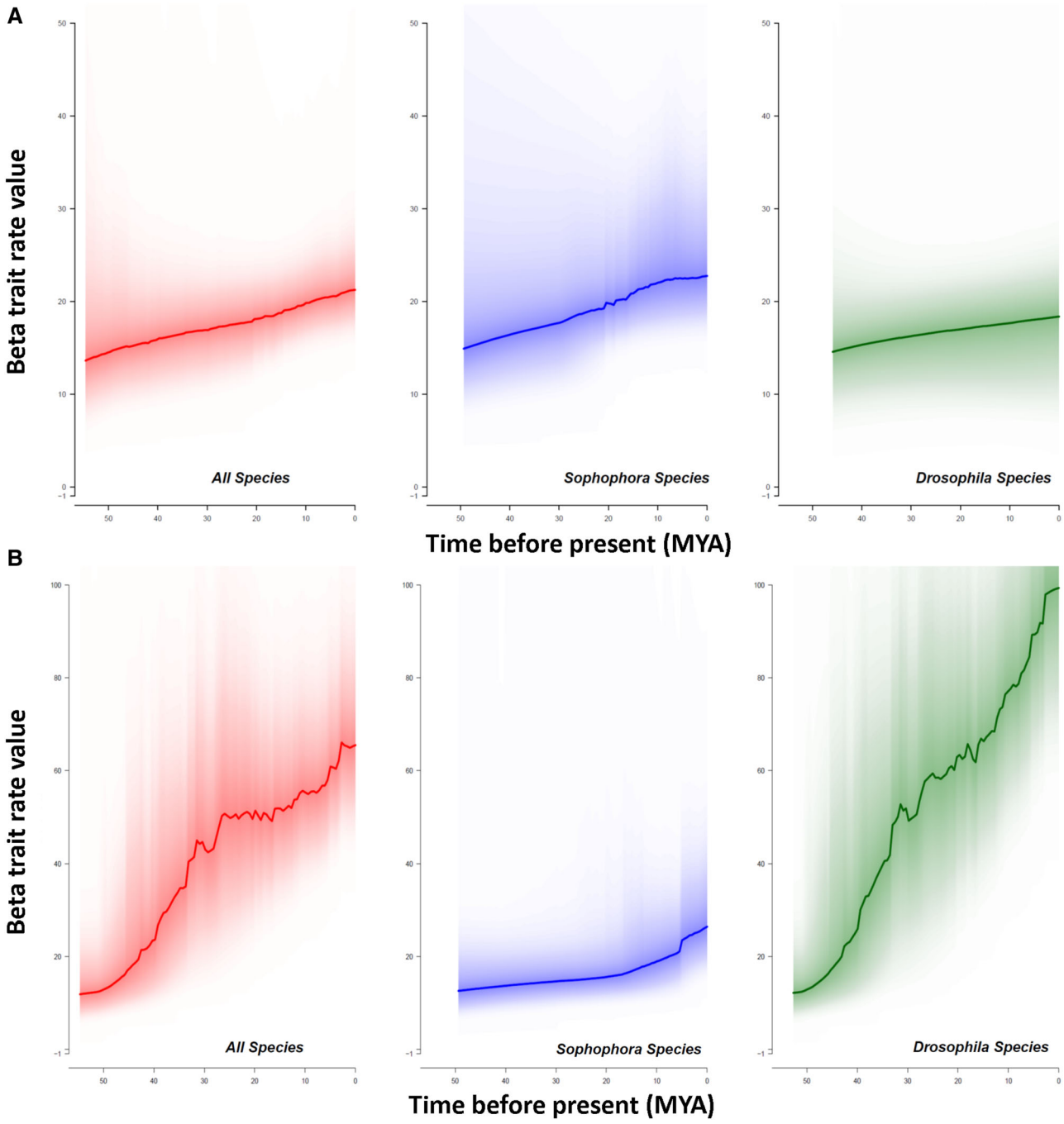

Figure 3. Rate of change throughout time for replicated and unreplicated DNA. Estimates of rate of change throughout time for (A) replicated DNA and (B) unreplicated DNA. Horizontal axis is the time before present in millions of years. The vertical axis represents the Brownian rate parameter, or the beta trait value for trait evolution. Light coloration around the line represents the 95% credible interval on the distribution of rates through time. Left panel (red) shows rate of change across entire *Drosophila* genus, middle panel (blue) shows rate of change across the *Sophophora* subgenus, and right panel (green) shows rate of change across *Drosophila* subgenus.

females across the entire *Drosophila* genus, as well as in both the *Drosophila* and *Sophophora* subgenera (Table 1), supporting that Y-chromosomes lose DNA content as they differentiate from the X-chromosome (Bachtrog 2008, 2013; Hjelmen et al. 2018, 2019). With estimates of the difference in genome size and amount of heterochromatin (unreplicated DNA) due to heteromorphic sex chromosomes, we can investigate the hypothetical path of X-Y chromosome differentiation and the loss/gain of DNA and heterochromatin on the Y chromosome (Fig. 2).

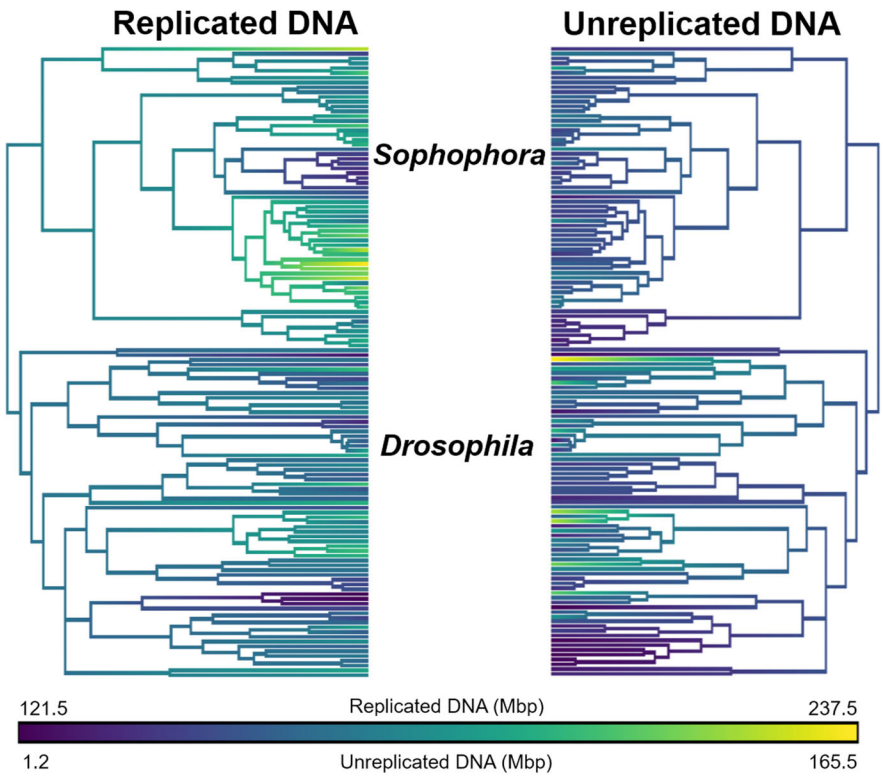

Figure 4. **Replicated and unreplicated DNA plotted across the *Drosophila* phylogeny. Replicated (left) and unreplicated (right) DNA plotted by color across the *Drosophila* genus phylogeny. *Sophophora* subgenus species are on the upper portion of the phylogeny, whereas *Drosophila* subgenus species are on the lower portion. Darker colors represent less DNA, whereas lighter colors represent more DNA.**

The gray arrow throughout the plot represents one common, hypothetical path of sex chromosome differentiation. Species with young sex chromosomes are expected to have similar female and male genome sizes, as the sex chromosomes are less differentiated. Early in differentiation of these sex chromosomes, the Y chromosome is expected to increase in size and become more heterochromatized due to increased mobile element activity and loss of genic content (Fig. 2) (Charlesworth et al. 2005; Matsunaga 2009; Bachtrog 2013). It is then expected that the Y chromosome will continue to become heterochromatized as it subsequently degrades due to transposable element suppression and deletion bias (Fig. 2). This process will further degrade the Y, decreasing its physical size, and eventually result in heterochromatin content equal to or less than the X chromosome (Fig. 2). Eventually, there may be sex chromosome loss or turnover, in which the process of decay will partially reset (Charlesworth and Charlesworth 2000; Bachtrog et al. 2008; Vicoso and Bachtrog 2015; Vicoso 2019). Although these events can be rare, there are representative species with sex chromosome turnover in *Drosophila*. We find a large proportion of these species with known neo-sex chromosomes to be located in the bottom left quadrant, supporting this hypothesis of change (species in red, Fig. 2).

It is important to note that we have presented one interpretation of our data here based on the canonical linear/cycle hypothesis of sex chromosome differentiation. This is not the only possible interpretation of our data. Y chromosomes are often portrayed as an "evolutionary trap," destined to degrade until they disappear (Aitken and Graves 2002; Graves 2004), yet the rate of gene loss may actually decrease over time until there is no further loss (Bachtrog 2008; Bachtrog et al. 2014), or sex chromosomes may even remain homomorphic (Bachtrog et al. 2014). However, sex chromosomes can be remarkably stable, and may even show signs of gene increase (Koerich et al. 2008). Sex chromosome stability, decreased gene loss, and instances of gene increase make it difficult to predict when a sex chromosome turnover may occur, and certainly makes it difficult to predict an ideal level of sex chromosome differentiation. Therefore, it is possible that each species has reached some level of sex chromosome stability and may not degrade or heterochromatize any further. It can be confidently stated, however, that in the majority of species here, the Y chromosome is smaller than the X (positively located on the horizontal axis) and the Y chromosome has more heterochromatin than the X chromosome (negatively located on vertical axis).

MINIMUM GENOME SIZES DIFFER BETWEEN THE SUBGENERA

When we look at all of the analyzed species in the *Drosophila* genus, the estimate of minimum genome size reflects the smallest genome size we see within the group of estimated species. Although these genome sizes are small relative to many insects, they are not the smallest seen within Diptera. It is important to recognize the diversity in genome configurations and divergence; however, this model relies on estimations of heterochromatin by measuring underreplication, which is only documented in *Drosophila* species. To investigate the impacts of these differences in genome configurations, we take advantage of the two subgenera within the *Drosophila* genus: *Sophophora* and *Drosophila*. These subgenera separated by an estimated 40-65 million years and differ notably in their chromosome numbers (Russo et al. 1995, 2013; Tamura et al. mar. 2004; Obbard et al. 2012). The majority of *Drosophila* subgenus species maintain the ancestral karyotype of six chromosomes, whereas *Sophophora* species have reduced chromosome numbers (three to five chromosomes) due to fusions (Schulze et al. 2006). The karyotypic differences between the subgenera may therefore result in different patterns of evolution for early- and late-replicating DNA. There was a remarkable difference in minimum genome size estimates when separating the *Sophophora* subgenus from the *Drosophila* subgenus (Table 1), with much smaller genomes predicted in *Sophophora* than *Drosophila*, even though there is no significant difference in whole genome size between the subgenera (Hjelmen et al. 2019). Interestingly, there were no significant differences in the linear models for unreplicated DNA when separated by chromosome number, suggesting that something else about the configuration of their genomes allows smaller sizes in the *Sophophora* subgenus than the *Drosophila* subgenus.

In this manuscript, estimations of minimum genome size are restricted to the genus *Drosophila*, one group within the diverse order of flies. However, the predicted minimum size for the *Sophophora* subgenus was remarkably close to the smallest insect genome sequenced to date, *Belgica antarctica* (98 Mbp) (Kelley et al. 2014). Although located within Diptera, *Drosophila* and *Belgica* are diverged at least 200 million years (Hedges and Kumar 2009; Wiegmann et al. 2011). Diptera karyotypes are highly conserved in terms of Muller elements—six highly syntenic elements known to comprise all known Diptera chromosomes (Muller 1940; Schaeffer 2018). The conservation of chromosomal elements suggests these predictive minimum genome sizes in *Drosophila* may further extend to the order Diptera. Although current genome size estimations place Diptera among the smallest genome sizes in insects, it is not clear how these minimum genome size estimations extend to the rest of insects. Fluctuations in noncoding regions, such as repeat content and transposable elements, drive genome size change (Gregory and Hebert 1999; Kidwell 2002; Sessegolo et al. 2016; Wright 2017). Although the relationship between genome size and these noncoding regions is clear, the patterns by which these late-replicating heterochromatic regions and early-replicating, mostly euchromatic regions evolve are largely unknown.

RATES AT WHICH ADDITIONAL DNA IS HETEROCHROMATIC OR EUCHROMATIC DIFFER BETWEEN SUBGENERA

Although it seems genome size in both subgenera is selected to be similar (Z_0 values, Table 2), the proportion of euchromatin to heterochromatin suggests different dynamics of genome size change between the subgenera. The slope parameter from our linear models allows us to investigate the rate at which additional DNA becomes either early- or late-replicating DNA (Table 1). In all cases, males had a higher proportion of additional DNA becoming late-replicating heterochromatic, likely driven by the heterochromatic Y chromosome. When separated by subgenus, there is a much higher proportion of additional DNA contributed as late-replicating heterochromatin in females of the *Drosophila* subgenus than of females in *Sophophora* (Table 1). These differences in the rates of change in early- and late-replicating DNA are further supported by the Ornstein-Uhlenbeck and BAMM analyses (Tables 2 and 3; Fig. 3). We propose that the difference in the proportion of additional DNA that becomes heterochromatic may reflect the cost of additional DNA in excess of the minimum possible genome. Based on our models, genome size change within species of the *Sophophora* subgenus is more reliant on changes in euchromatic DNA than heterochromatic DNA. It is expected that changes in euchromatic DNA are more likely to impact coding sequences, and are therefore less tolerable than changes in heterochromatic sequences. However, these data suggest that changes in euchromatic DNA in *Sophophora* may be more tolerated than in the *Drosophila* subgenus. Although *Sophophora* may have mechanisms to allow changes in euchromatin, addition of DNA in *Sophophora* may be disruptive and may account for the small minimum genome size estimates and the lower rates of change seen in this subgenus, whereas genome size change in the *Drosophila* subgenus is driven by large changes in heterochromatin and is therefore less likely to directly impact coding sequence. These changes are more likely tolerated, resulting in more rate shifts in change of unreplicated DNA (Figs. S11 and S12), and consequently higher rates of genome size change, seen in the *Drosophila* subgenus (Hjelmen et al. 2019).

Although there is high support for zero rate shifts of replicated DNA in females across the *Drosophila* genus, there is high support (88.1%) for one rate shift to occur in replicated DNA of males (Table 3). In fact, eight of nine top models for rate shifts in males suggest a rate shift at the split of the subgenera (Fig. S9), with higher rates of change in the *Sophophora*

subgenus (Fig. S10). This difference in the rates of change in replicated DNA is not seen in females (Table 3; Fig. 3A), suggesting that this shift is due to changes in replicated content of the Y chromosome, supporting higher rates of sex chromosome differentiation in the *Sophophora* subgenus. The *Sophophora* subgenus has known neo-sex chromosomes in species such as *Drosophila miranda* (Bachtrog 2004; Bachtrog et al. 2008) and *Drosophila pseudoobscura* (Carvalho and Clark 2005). These results are supportive of higher rates of sex chromosome differentiation and sex chromosome turnover in *Sophophora* predicted by our previous study (Hjelmen et al. 2019).

IMPLICATIONS OF CHANGE IN EUCHROMATIN AND HETEROCHROMATIN ACROSS *DROSOPHILA*

Large structural changes in the genome, whether through replicated or unreplicated DNA, are expected to decrease the amount of successful hybridization between species, subsequently leading to reproductive isolation and eventually speciation (Brown and O'Neill 2010). In the case of meiosis, heterochromatic content may facilitate speciation by allowing chromosomal rearrangements, forming genetic barriers against hybridization (Yunis and Yasmineh 1971). Within the *Drosophila* genus, it is suggested that speciation is impacted by qualitative variation in heterochromatin types (Gatti et al. 1976). In the case of *D. melanogaster* and *Drosophila simulans*, rapid evolution of heterochromatic DNA may be driving evolution of genes related to hybrid incompatibility and potential speciation (Brideau et al. 2006). Visualizing replicated and unreplicated DNA on the *Drosophila* genus phylogeny highlights the remarkable variation in replicated DNA across the phylogeny (Fig. 4). Within smaller clades, species closely related to one other tend to have similar amounts of replicated DNA. Some clades have increases in replicated DNA, such as the montium group within the *Sophophora* subgenus and the *virilis* group within the *Drosophila* subgenus, whereas other clades have reductions in replicated DNA, such as the *nannoptera* group within the *Drosophila* and the *melanogaster* group in the *Sophophora*, likely corresponding to the overall decrease in genome size for this clade. These notable differences between clades suggest that reconfiguration of early and late replicating portions of the genome may associate with bursts of speciation; however, further work must be done to verify this statement.

The most striking of the aforementioned differences in replicated DNA is the *nannoptera* group of *Drosophila*. These species have been noted to have a high proportion of heterochromatic DNA content, including some almost entirely heterochromatic chromosomes (Ward and Heed 1970). Although the amount of unreplicated DNA they have is not remarkable compared to sister species, the proportion of their genome which is unreplicated is striking. Although having female genome sizes ranging from 155.8 to 245.9 Mbp, each species only replicates just over 120 Mbp of DNA in the female and only about 100 Mbp in the male (Table S1). Comparatively speaking, on average, *Drosophila* species replicate 171.7 Mbp of DNA in females and 155.6 Mbp in males (S1 Table). Population sizes of *Drosophila pachea*, one member of this group, have been noted to be small (Breitmeyer and Markow 1998), suggesting that drift might have contributed substantially to the dramatic reconfiguration of heterochromatin in these genomes. However, one interesting explanation for the uniqueness of these species is the pressure experienced in their desert dwelling, cactophilic lifestyle. The unique characteristics of species in the *nannoptera* group, such as specialization on columnar cacti and asymmetric genitalia, have been studied extensively in relation to their ecological speciation, phylogenetic relationships, and specialization (Lang et al. 2012, 2014; Acurio et al. 2019). The exposure to extreme heat and potential for long dispersal distances may be detrimental to maintaining genomic integrity. The increase in heterochromatic proportion of the genome may therefore play a protective role. It is thought that heterochromatin may play a protective role in some areas of the genome; it is present around vital areas in chromosomes, thereby providing strength and potentially absorbing effects of mutagenic agents (Yunis and Yasmineh 1971). This group requires further investigation with modern sequencing methods to identify repeat regions and chromatin makeup to investigate the potential adaptive advantage of such a high proportion of heterochromatic content.

The advantage of underreplication in specific tissues is unknown. Underreplication (and forms of endoreduplication) is assumed to have an impact on gene expression and cell size. Although it seems plausible that an increase in gene copy number through replication may increase the expression of those genes (Wu et al. 2010), evidence that this replication impacts expression has yet to be confirmed (reviewed in Neiman et al. 2017). In bacteria and archaea, gene expression does, in fact, increase with replication of DNA; however, these patterns of increased expression are not seen when investigated in yeast (Voichek et al. 2016). Like expression, there is evidence for a positive relationship between size and amount of DNA replicated (Nagl 1976, 1978), but conflicting evidence also exist (Neiman et al. 2017). Because underreplication is known to be found in thoracic tissue, specific tissues within the thorax of these flies have been further dissected to quantify underreplication levels in specific flight muscle tissues (Johnston et al. 2020). It is possible that underreplication has evolved to increase the muscle cell size or transcriptional output of these highly energetic cells to better flight performance of these flies. This supposition, however, is conjecture and must be further investigated with transcriptomic analyses and flight performance assays.

Conclusions

In conclusion, the strong relationship between thoracic underreplication and genome size in *Drosophila* allows us to predict a minimum genome size within flies, as well as estimate the rate at which additional DNA becomes heterochromatic. Our estimates for minimum size of the *Sophophora* subgenus are remarkably close to the smallest insect genome sequenced to date, *Belgica antarctica*, which is over 200 million years diverged from *Drosophila*. We find that 61.7% of additional DNA is partitioned to late-replicating heterochromatic DNA, with a much lower proportion of late-replicating DNA in the *Sophophora* subgenus. These data support the idea that the rate of genome size change is much higher in the *Drosophila* subgenus than in the *Sophophora* subgenus, and that this is driven by higher rates of change in heterochromatic content. These estimates of underreplication and heterochromatin content provide a powerful resource for efficiently and cost-effectively identifying species with unique heterochromatin profiles for further investigation via long read sequencing, chromatin profiling, and transcriptomic studies. The advantage of thoracic underreplication remains unknown and must be further investigated with phenotypic and physiological performance studies.

AUTHOR CONTRIBUTIONS

CEH and JSJ conceptualized the study. CEH, JSJ, and HB were associated with the methodology of the study. CEH, VRH, CGB, EP, MM, MAG, and JSJ collected the data. CEH, JSJ, and HB performed formal analysis. CEH was associated with data curation. CEH prepared the original draft. CEH, VRH, CGB, EP, MM, MAG, HB, and JSJ reviewed and edited the manuscript. All authors have read and approved of the manuscript.

ACKNOWLEDGMENTS

We would like to thank the Editor and Associate editor, as well at the two anonymous reviewers, for their helpful comments throughout the manuscript submission and review process.

DATA ARCHIVING

Data are archived in supplementary files. Tree files and data csv file are archived on Dryad (https://doi.org/10.5061/dryad.sqv9s4n1m).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

LITERATURE CITED

Acurio, A. E., F. T. Rhebergen, S. Paulus, V. Courtier-Orgogozo, and M. Lang. 2019. Repeated evolution of asymmetric genitalia and right-sided mating behavior in the *Drosophila nannoptera* species group. BMC Evol. Biol. 19:109.

Aitken, R. J., and J. A. M. Graves. 2002. Human spermatozoa: the future of sex. Nature 415:963–963.

Andreyeva, E. N., T. D. Kolesnikova, E. S. Belyaeva, R. L. Glaser, and I. F. Zhimulev. 2008. Local DNA underreplication correlates with accumulation of phosphorylated H2Av in the *Drosophila melanogaster* polytene chromosomes. Chromosome Res. 16:851–862.

Bachtrog, D. 2004. Evidence that positive selection drives Y-chromosome degeneration in *Drosophila miranda*. Nat. Genet. 36:518–522.

———. 2008. The temporal dynamics of processes underlying Y chromosome degeneration. Genetics 179:1513–1525.

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

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

Bachtrog, D., J. E. Mank, C. L. Peichel, M. Kirkpatrick, S. P. Otto, T.-L. Ashman, M. W. Hahn, J. Kitano, I. Mayrose, and R. Ming. 2014. Sex determination: why so many ways of doing it? PLoS Biol. 12:e1001899.

Barigozzi, C., S. Dolfini, M. Fraccaro, G. R. Raimondi, and L. Tiepolo. 1966. In vitro study of the DNA replication patterns of somatic chromosomes of *Drosophila melanogaster*. Exp. Cell Res. 43:231–234.

Belyaeva, E. S., I. F. Zhimulev, E. I. Volkova, A. A. Alekseyenko, Y. M. Moshkin, and D. E. Koryakov. 1998. Su(UR)ES: a gene suppressing DNA underreplication in intercalary and pericentric heterochromatin of *Drosophila melanogaster* polytene chromosomes. Proc. Natl. Acad. Sci. USA 95:7532–7537.

Belyaeva, E. S., S. A. Demakov, G. V. Pokholkova, A. A. Alekseyenko, T. D. Kolesnikova, and I. F. Zhimulev. 2006. DNA underreplication in intercalary heterochromatin regions in polytene chromosomes of *Drosophila melanogaster* correlates with the formation of partial chromosomal aberrations and ectopic pairing. Chromosoma 115:355–366.

Blackmon, H., and R. Adams. 2015. R package 'evobiR'v. 1.1. Cran R.

Bosco, G., P. Campbell, J. T. Leiva-Neto, and T. A. Markow. 2007. Analysis of *Drosophila* species genome size and satellite DNA content reveals significant differences among strains as well as between species. Genetics 177:1277–1290.

Breitmeyer, C. M., and T. A. Markow. 1998. Resource availability and population size in cactophilic *Drosophila*. Funct. Ecol. 12:14–21.

Brideau, N. J., H. A. Flores, J. Wang, S. Maheshwari, X. Wang, and D. A. Barbash. 2006. Two Dobzhansky-Muller genes interact to cause hybrid lethality in *Drosophila*. Science 314:1292–1295.

Brown, J. D., and R. J. O'Neill. 2010. Chromosomes, conflict, and epigenetics: chromosomal speciation revisited. Annu. Rev. Genomics Hum. Genet. 11:291–316.

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

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

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

Gatti, M., S. Pimpinelli, and G. Santini. 1976. Characterization of *Drosophila* heterochromatin. Chromosoma 57:351–375.

Graves, J. A. M. 2004. The degenerate Y chromosome–can conversion save it? Reprod. Fertil. Dev. 16:527–534.

Gregory, T. R. 2001. Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. Biol. Rev. 76:65–101.

———. 2005. The evolution of the genome. Elsevier, Amsterdam, the Netherlands.

———. 2020. Animal genome size database. Available via http://www.genomesize.com.

Gregory, T. R., and P. D. Hebert. 1999. The modulation of DNA content: proximate causes and ultimate consequences. Genome Res. 9:317–324.

Hammond, M. P., and C. D. Laird. 1985. Control of DNA replication and spatial distribution of defined DNA sequences in salivary gland cells of *Drosophila melanogaster*. Chromosoma 91:279–286.

Harmon, L., J. Weir, C. Brock, R. Glor, W. Challenger, G. Hunt, R. FitzJohn, M. Pennell, G. Slater, and J. Brown. 2009. geiger: analysis of evolutionary diversification. R package version 1.

Hedges, S. B., and S. Kumar. 2009. The timetree of life. Oxford Univ. Press, Oxford, U.K.

Heitz, E. 1928. Das heterochromatin der moose. Jahrb. Wiss. Bot. 69:762–818.

———. 1929. Heterochromatin, chromocentren, chromomeren. Ber. Botan. Ges. 47:274–284.

Hjelmen, C. E., M. A. Garrett, V. R. Holmes, M. Mynes, E. Piron, and J. S. Johnston. 2018. Genome size evolution within and between the sexes. J. Hered. 110:219–228.

Hjelmen, C. E., H. Blackmon, V. R. Holmes, C. G. Burrus, and J. S. Johnston. 2019. Genome size evolution differs between *Drosophila* subgenera with striking differences in male and female genome size in *Sophophora*. G3 9:3167–3179.

Hua, B. L., and T. L. Orr-Weaver. 2017. DNA replication control during Drosophila development: insights into the onset of S phase, replication initiation, and fork progression. Genetics 207:29–47.

Hua, B. L., S. Li, and T. L. Orr-Weaver. 2014. The role of transcription in the activation of a *Drosophila* amplification origin. G3 4:2403–2408.

Johnston, J. S., M. Schoener, and D. P. McMahon. 2013. DNA underreplication in the majority of nuclei in the *Drosophila melanogaster* thorax: evidence from Suur and flow cytometry. J. Mol. Biol. Res. 3:p47.

Johnston, J. S., M. E. Zapalac, and C. E. Hjelmen. 2020. Flying high—muscle-specific underreplication in *Drosophila*. Genes 11:246.

Kelley, J. L., J. T. Peyton, A. S. Fiston-Lavier, N. M. Teets, M. C. Yee, J. S. Johnston, C. D. Bustamante, R. E. Lee, and D. L. Denlinger. 2014. Compact genome of the Antarctic midge is likely an adaptation to an extreme environment. Nat. Commun. 5:4611.

Kidwell, M. G. 2002. Transposable elements and the evolution of genome size in eukaryotes. Genetica 115:49–63.

Koerich, L. B., X. Wang, A. G. Clark, and A. B. Carvalho. 2008. Low conservation of gene content in the *Drosophila* Y chromosome. Nature 456:949–951.

Lang, M., S. Murat, A. G. Clark, G. Gouppil, C. Blais, L. M. Matzkin, É. Guittard, T. Yoshiyama-Yanagawa, H. Kataoka, R. Niwa, et al. 2012. Mutations in the *neverland* gene turned *Drosophila pachea* into an obligate specialist species. Science 337:1658–1661.

Lang, M., M. Polihronakis Richmond, A. E. Acurio, T. A. Markow, and V. Orgogozo. 2014. Radiation of the *Drosophila nannoptera* species group in Mexico. J. Evol. Biol. 27:575–584.

Leach, T. J., H. L. Chotkowski, M. G. Wotring, R. L. Dilwith, and R. L. Glaser. 2000. Replication of heterochromatin and structure of polytene chromosomes. Mol. Cell. Biol. 20:6308–6316.

Lima-de-Faria, A., and H. Jaworska. 1968. Late DNA synthesis in heterochromatin. Nature 217:138–142.

MacAlpine, D. M., H. K. Rodriguez, and S. P. Bell. 2004. Coordination of replication and transcription along a *Drosophila* chromosome. Genes Dev. 18:3094–3105.

Makunin, I. V., T. D. Kolesnikova, and N. G. Andreyenkova. 2014. Underreplicated regions in *Drosophila melanogaster* are enriched with fast-evolving genes and highly conserved noncoding sequences. Genome Biol. Evol. 6:2050–2060.

Matsunaga, S. 2009. Junk DNA promotes sex chromosome evolution. Heredity 102:525–526.

McNairn, A. J., and D. M. Gilbert. 2003. Epigenomic replication: linking epigenetics to DNA replication. Bioessays 25:647–656.

Muller, H. 1940. Bearings of the "*Drosophila*" work on systematics. Pp. 185–268 *in* J. Huxley, ed. The new systematics. Clarendon Press, Lond.

Nagl, W. 1976. DNA endoreduplication and polyteny understood as evolutionary strategies. Nature 261:614–615.

———. 1978. Endopolyploidy and polyteny in differentiation and evolution. North-Holland Publishing Company, Amsterdam, the Netherlands.

Neiman, M., M. J. Beaton, D. O. Hessen, P. D. Jeyasingh, and L. J. Weider. 2017. Endopolyploidy as a potential driver of animal ecology and evolution. Biol. Rev. 92:234–247.

Obbard, D. J., J. Maclennan, K.-W. Kim, A. Rambaut, P. M. O'Grady, and F. M. Jiggins. 2012. Estimating divergence dates and substitution rates in the *Drosophila* phylogeny. Mol. Biol. Evol. 29:3459–3473.

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

Painter, T. S., and E. C. Reindorp. 1939. Endomitosis in the nurse cells of the ovary of *Drosophila melanogaster*. Z. Zellforsch. Mikrosk. Anat. Abt. B Chromosoma 1:276–283.

Palazzo, A. F., and T. R. Gregory. 2014. The case for junk DNA. PLoS Genet. 10:e1004351.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290.

Passarge, E. 1979. Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. Am. J. Hum. Genet. 31:106–115.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and R. C. Team. 2012. nlme: linear and nonlinear mixed effects models. R package version 3.

Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. CODA: convergence diagnosis and output analysis for MCMC. R News 6:7–11.

R Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available via www.R-project.org.

Rabosky, D. L., M. Grundler, C. Anderson, P. Title, J. J. Shi, J. W. Brown, H. Huang, and J. G. Larson. 2014. BAMM tools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. Methods Ecol. Evol. 5:701–707.

Redi, C., S. Garagna, H. Zacharias, M. Zuccotti, and E. Capanna. 2001. The other chromatin. Chromosoma 110:136–147.

Renkawitz-Pohl, R., and W. Kunz. 1975. Underreplication of satellite DNAs in polyploid ovarian tissue of *Drosophila virilis*. Chromosoma 49:375–382.

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

Rudkin, G. T. 1969. Non replicating DNA in *Drosophila*. Genetics 61:227–238.

Russo, C., N. Takezaki, and M. Nei. 1995. Molecular phylogeny and divergence times of drosophilid species. Mol. Biol. Evol. 12:391–404.

Russo, C. A., B. Mello, A. Frazão, and C. M. Voloch. 2013. Phylogenetic analysis and a time tree for a large drosophilid data set (Diptera: Drosophilidae). Zool. J. Linn. Soc. 169:765–775.

Schaeffer, S. W. 2018. Muller "elements" in *Drosophila*: how the search for the genetic basis for speciation led to the birth of comparative genomics. Genetics 210:3–13.

Schübeler, D., D. Scalzo, C. Kooperberg, B. van Steensel, J. Delrow, and M. Groudine. 2002. Genome-wide DNA replication profile for *Drosophila melanogaster*: a link between transcription and replication timing. Nat. Genet. 32:438–442.

Schulze, S. R., B. F. McAllister, D. A. R. Sinclair, K. A. Fitzpatrick, M. Marchetti, S. Pimpinelli, and B. M. Honda. 2006. Heterochromatic genes in *Drosophila*: a comparative analysis of two genes. Genetics 173:1433–1445.

Schwaiger, M., and D. Schübeler. 2006. A question of timing: emerging links between transcription and replication. Curr. Opin. Genet. Dev. 16:177–183.

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

Tamura, K., S. Subramanian, and S. Ku mar. 2004. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. Mol. Biol. Evol. 21:36–44.

Vicoso, B. 2019. Molecular and evolutionary dynamics of animal sex-chromosome turnover. Nat. Ecol. Evol. 3:1632–1641.

Vicoso, B., and D. Bachtrog. 2015. Numerous transitions of sex chromosomes in Diptera. PLoS Biol. 13:e1002078.

Voichek, Y., R. Bar-Ziv, and N. Barkai. 2016. Expression homeostasis during DNA replication. Science 351:1087–1090.

Ward, B. L., and W. B. Heed. 1970. Chromosome phylogeny of *Drosophila pachea* and related species. J. Hered. 61:248–258.

Wiegmann, B. M., M. D. Trautwein, I. S. Winkler, N. B. Barr, J.-W. Kim, C. Lambkin, M. A. Bertone, B. K. Cassel, K. M. Bayless, A. M. Heimberg, et al. 2011. Episodic radiations in the fly tree of life. Proc. Natl. Acad. Sci. USA 108:5690–5695.

Wolfe, K. H., P. M. Sharp, and W. H. Li. 1989. Mutation rates differ among regions of the mammalian genome. Nature 337:283–285.

Wright, S. I. 2017. Evolution of genome size. John Wiley & Sons, Ltd., Hoboken, NJ.

Wu, C.-Y., P. A. Rolfe, D. K. Gifford, and G. R. Fink. 2010. Control of transcription by cell size. PLoS Biol. 8:e1000523.

Yunis, J. J., and W. G. Yasmineh. 1971. Heterochromatin, satellite DNA, and cell function. Science 174:1200–1209.

Yurlova, A. A., I. V. Makunin, and I. F. Zhimulev. 2010. Phylogenetic analysis of the fast-evolving SuUR gene in insects. Russ. J. Genet. 46:1127–1129.

Zhimulev, I. F., E. S. Belyaeva, V. F. Semeshin, V. V. Shloma, I. V. Makunin, and E. I. Volkova. 2003. Overexpression of the SuUR gene induces reversible modifications at pericentric, telomeric and intercalary heterochromatin of *Drosophila melanogaster* polytene chromosomes. J. Cell Sci. 116:169–176.

Associate Editor: C. Burch
Handling Editor: T. Chapman

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Female Replication Percentage (vertical axis) plotted against female genome size (horizontal axis).
Figure S2. Male Thoracic Replication Percentage (vertical axis) plotted against male genome size (horizontal axis).
Figure S3. Female replicated DNA (vertical axis) plotted against female genome size (horizontal axis).
Figure S4. Female unreplicated DNA (vertical axis) plotted against female genome size (horizontal axis).
Figure S5. Male Replicated DNA (vertical axis) plotted against male genome size (horizontal axis).
Figure S6. Male unreplicated DNA (vertical axis) plotted against male genome size (horizontal axis).
Figure S7. Male Unreplicated DNA linear model by subgenera.
Figure S8. Credible Shift set for female replicated DNA. Circles represent rate shifts.
Figure S9. Credible Shift set for Male replicated DNA. Circles represent rate shifts.
Figure S10. Male replicated DNA rate shifts throughout time in the entire Drosophila genus (left), in the Sophophora subgenus (middle) and the Drosophila subgenus (right).
Figure S11. Credible Shift set for Female unreplicated DNA. Circles represent rate shifts.
Figure S12. Credible Shift sets for Male unreplicated DNA. Circles represent rate shifts.
Figure S13. Rate shifts throughout time for male unreplicated DNA.
Figure S14. Female Replicated DNA Phylogeny.
Figure S15. Male Replicated DNA Phylogeny.
Figure S16. Female Unreplicated DNA Phylogeny.
Figure S17. Male Unreplicated DNA Phylogeny.
Table S1. Underreplication Measurements, Genome Size, Chromosome Numbers.
Table S2. Linear Model and Minimum Genome Size Corrected for Phylogeny.
Table S3. fitContinous Model Testing AICc Results.