

# Support for faster and more adaptive Z chromosome evolution in two divergent lepidopteran lineages\*

Andrew J. Mongue,<sup>1,2,3</sup>  Megan E. Hansen,<sup>1</sup> and James R. Walters<sup>1</sup> 

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas 66045

<sup>2</sup>Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH93FL, United Kingdom

<sup>3</sup>E-mail: ajmongue@gmail.com

Received February 14, 2021

Accepted August 6, 2021

The rate of divergence for Z or X chromosomes is usually observed to be greater than autosomes, but the proposed evolutionary causes for this pattern vary, as do empirical results from diverse taxa. Even among moths and butterflies (Lepidoptera), which generally share a single-origin Z chromosome, the handful of available studies give mixed support for faster or more adaptive evolution of the Z chromosome, depending on the species assayed. Here, we examine the molecular evolution of Z chromosomes in two additional lepidopteran species: the Carolina sphinx moth and the monarch butterfly, the latter of which possesses a recent chromosomal fusion yielding a segment of newly Z-linked DNA. We find evidence for both faster and more adaptive Z chromosome evolution in both species, although this effect is strongest in the neo-Z portion of the monarch sex chromosome. The neo-Z is less male-biased than expected of a Z chromosome, and unbiased and female-biased genes drive the signal for adaptive evolution here. Together these results suggest that male-biased gene accumulation and haploid selection have opposing effects on long-term rates of adaptation and may help explain the discrepancies in previous findings as well as the repeated evolution of neo-sex chromosomes in Lepidoptera.

**KEY WORDS:** Chromosomal evolution, evolutionary genomics, molecular evolution, population genetics, sex chromosomes.

Explaining patterns of genetic variation in natural populations is a foundational goal of population genetics. In basic terms, variation is shaped by either selective or neutral processes. But beneath this simplicity, dynamics quickly become more complicated. For example, the efficiency of selection relative to drift depends on the effective population size of the genes in question (Ohta 1992). Simple census population size is often a poor proxy for the effective population size, as historical population size changes have long-lasting effects (Tajima 1989). Also, different parts of the genome may have different population sizes due to either differences in ploidy or conditional limitations on expression. For organisms with chromosomal sex determination, the sex chromosomes present a particularly complex confluence of the above processes (Wilson Sayres 2018).

Relative to the rest of the genome, sex chromosomes have smaller population sizes, occurring at either one fourth (Y or W) or three fourths (X or Z) the frequency of autosomes. Evolution of the Y and W is thought to be driven mainly by a lack of recombination, leading to the degeneration of all but the essential genes in many cases (Charlesworth and Charlesworth 2000; Bachtrog 2013). X and Z chromosomes, however, maintain a large set of functional genes despite often having a smaller population size than the autosomes. This should decrease the efficiency of selection and increase genetic drift on sex-linked genes (Vicoso and Charlesworth 2009). Conversely, because the X/Z is hemizygous in one sex, assuming differentiation between X-Y or Z-W, new mutations may be more exposed to selection than on autosomes, increasing rates of adaptation (Rice 1984; Charlesworth et al. 1987). Both of these scenarios (increased drift or increased selection) may lead to more rapid rates of molecular evolution on the X/Z relative to autosomes, a phenomenon called “Faster-X”/ “Faster-Z.” As such, although increased divergence of sex

\*This article corresponds to Homa Papoli Yazdi 2022. Digest: Neo-sex chromosome underlying the Faster-Z effect in a species of Lepidoptera. *Evolution*. <https://doi.org/10.1111/evo.14418>.

chromosomes has been observed repeatedly (Baines et al. 2008; Meisel and Connallon 2013; Kousathanas et al. 2014; Hayes et al. 2020), discerning between drift and selection as the primary cause of this pattern remains an outstanding challenge in evolutionary genomics.

A further complication to understanding sex chromosome evolution is the sex-biased expression of many genes on the sex chromosomes. Because selection can only act on expressed phenotypes, sex-biased genes should be shielded from selection in one sex and experience increased divergence due to drift (Gershoni and Pietrokovski 2014; Dapper and Wade 2016). However, as mentioned above, haploid selection could counter this reduced selection, but (assuming both copies of the X/Z are expressed in the homogametic sex) this benefit will only apply to male-biased genes on the X or female-biased genes on the Z. As such, the importance of haploid selection compared to drift on the sex chromosomes should depend on the gene content of the chromosomes (e.g., more efficient selection of female-biased genes on the Z may have little overall impact on the chromosome if the vast majority of Z-linked genes are male-biased in expression).

The X spends more time in females than males (and vice versa for the Z), which generates the expectation that sex-biased genes will accumulate on the sex chromosomes, although whether male- or female-biased genes accumulate is thought to be dependent on whether the average new mutation in these genes is dominant or recessive (Rice 1984; Chapman et al. 2003). In practice, however, the X is often found to be enriched for female-biased genes and the Z is commonly observed to be male-biased in composition (Walters and Hardcastle 2011; Meisel et al. 2012; Wright et al. 2012; Mank et al. 2014; Mongue and Walters 2017). In other words, the composition of the sex chromosomes tends to be biased *against* the class of genes that could drive adaptation through haploid selection (Baines et al. 2008). So although faster-Z adaptation should be most apparent for female-biased genes (Parsch and Ellegren 2013; Sackton et al. 2014), the relative scarcity of Z-linked female-biased genes may limit both the importance of this adaptation and our ability to detect it.

Finally, all of the above processes of increased drift or enhanced selection relative to the autosomes exist within the bounds of the focal organism's demography and biology. In X chromosome systems, evidence for more adaptive evolution tends to be associated with species with larger absolute effective population sizes and consequently more efficient selection across the genome (typically invertebrates, reviewed in Meisel and Connallon 2013; but see also Whittle et al. 2020 for a lack of faster-X in a beetle). Relative effective population sizes between the sex chromosomes and autosomes can vary between species as well, adding complexity. Males of many species have higher variance in reproductive success than females (Bateman 1948), meaning that the number of successful male alleles is lower than the cen-

sus count; the degree of this difference depends on how much male-male competition exists in a population. For the sex chromosomes, especially the male-biased Z, this can mean a further reduction in the effective population size and a greater role for drift (Vicoso and Charlesworth 2009).

Compared to X chromosome systems, Z chromosome systems are less well-studied, with results coming mostly from the single-origin Z chromosome of birds (Griffiths et al. 1998). These studies indicate Z-linked genes diverge faster primarily due to increased genetic drift, not adaptation (Mank et al. 2009; Wang et al. 2014; Wright et al. 2015; Xu et al. 2019; Hayes et al. 2020), although one study did show increased adaptive divergence on the Z by looking at expression differences rather than sequence divergence (Dean et al. 2015). The relative consistency of Z chromosome evolution in birds may be driven by the relatively low genome-wide effective population size of these vertebrates (compared to invertebrates) or by other idiosyncratic biology of birds. Most prominently, birds lack dosage compensation of the sex chromosomes; in other words, genes on the single copy Z in females are generally expressed at a lower level than genes expressed on the Z chromosomes of males (Ellegren et al. 2007; reviewed in Gu and Walters 2017), which could reduce the selective advantage of beneficial alleles expressed primarily in females (Charlesworth et al. 1987) and hinder adaptive evolution. As such, the generalizability of a faster-Z driven primarily by drift is in question. If larger effective population sizes yield greater adaptation on the sex chromosomes and dosage compensation supports selection, then the strongest test for adaptive Z evolution should come from ZW systems with large natural populations and dosage compensation of the sex chromosomes.

Butterflies and moths (Lepidoptera) are one of the oldest female-heterogametic groups and often have estimated effective population sizes that are orders of magnitude larger than most vertebrates (Mongue et al. 2019; based on nucleotide diversity at fourfold degenerate sites). So even the Z chromosome, with three fourths the population size of autosomes, should have much more efficient selection than found in most vertebrate species. Generally speaking, the lepidopteran Z chromosome's expression is balanced such that expression is equal between the sexes (Gu and Walters 2017), removing one of the complications to untangling Z evolution in birds. Moreover, recombination takes place in spermatogenesis but not oogenesis in Lepidoptera (Turner and Sheppard 1975). As such, in a given generation, two thirds of the Z chromosomes will recombine (those found in males), whereas only half of the autosomes (being found equally in males and females) will undergo recombination. This increased rate of recombination could help overcome the smaller population size of the Z relative to autosomes as it should decrease the linkage disequilibrium between loci and allow for more efficient selection. Yet

in spite of this confluence of factors, evidence for a lepidopteran faster-Z effect is mixed at best.

One study found faster rates of evolution on the Z (Sackton et al. 2014) but two others did not (Rousselle et al. 2016; Pinharanda et al. 2019). Likewise, evidence for a more adaptive Z than autosomes is conflicting, with two of the previous studies finding more adaptation (Sackton et al. 2014; Pinharanda et al. 2019) and the third finding the opposite: increased purifying selection (Rousselle et al. 2016). These contradictory results are particularly baffling given that all Lepidoptera share a single-origin Z chromosome (Fraïsse et al. 2017) and high levels of synteny (i.e., conserved gene order) across their phylogeny (Ahola et al. 2014; Davey et al. 2016; Kanost et al. 2016). Thus, differences in observed evolution may be attributable to a mixture of methodology and lineage-specific effects (e.g., mating systems skewing effective population sizes).

Here, we combine genomic data with gene expression analysis in a pair of distantly related Lepidoptera to place existing studies in context and better understand whether and why the Z chromosome evolves faster than autosomes. We take advantage of robust sequencing data in two species with estimated autosomal effective population sizes greater than 1 million (Mongue et al. 2019): the Carolina sphinx moth, *Manduca sexta*, and the monarch butterfly, *Danaus plexippus*. Of particular importance, the monarch possesses a recent Z-autosome fusion, creating a Z chromosome with roughly twice the number of genes found in the ancestral karyotype. The exact age of this fusion is still unknown, but it is shared by all members of the genus *Danaus* but no other members of the family Nymphalidae, to which *Danaus* belongs. Although these butterflies also appear to possess a neo-W chromosome, there remains no detectable sequence homology between the neo-Z and neo-W (as evidenced by in situ hybridization: Mongue et al. 2017; and a lack of heterozygosity on the neo-Z of females: Gu et al. 2019). As a result, many previously autosomal genes have become sex-linked and haploid expressed in these butterflies. Thus, the gene content, distribution, and differentiation of the neo-Z allow us to examine how relatively newly sex-linked sequence evolves once it becomes haploid in one sex.

## Materials and Methods

### POPULATION RESEQUENCING, POLYMORPHISM, AND DIVERGENCE

For *M. sexta*, the within-species variation dataset came from published whole-genome resequencing of 12 wild North Carolinian males and sequence divergence came from comparison of *M. sexta* to a *Manduca quinquemaculata* male (Mongue et al. 2019). For *Danaus plexippus*, polymorphisms and divergences came from a resequencing project (Zhan et al. 2014), from which we selected 12 males from the North American migratory popu-

lation of *D. plexippus* and one *Danaus gilippus* male. Note that *D. gilippus* shares the neo-Z with *D. plexippus*, allowing for an equivalent comparison in divergence rates across the genome. Polymorphism and expression analyses both used as a reference *D. plexippus* genome assembly version 3 and gene set version 2 (OGS2.0) (Zhan and Reppert 2013).

For each gene, we took the whole-genome Illumina data described above through a variant-calling pipeline described in Mongue et al. (2019). Briefly, we took adapter-removed, quality-trimmed data through the *Genome Analysis Toolkit* (version 3.7) pipeline (McKenna et al. 2010) to generate a set of high-quality variants. Within-species reads were aligned to the reference genome using *Bowtie2*'s very-sensitive-local aligner (Langmead and Salzberg 2012), whereas heterospecific reads were aligned to the same reference with *stampy* version 1.0.22 with an increased allowance for mismatches to better align divergent data (default parameters with the exception of substitution rate = 0.1, Lunter and Goodson 2011). Variant call files were hard-filtered to remove low-quality variants (specific filtering parameters: Quality by Depth >2.0 and Fisher Strand-bias <60 and Mapping Quality > 40); from the remaining single nucleotide variants, we classified each as synonymous or nonsynonymous using *SNPeff* (version 4.2, Cingolani et al. 2012) and normalized variant counts by the number of nonsynonymous or synonymous sites in each gene using R scripts in version 3.3.3 to annotate and sum the degeneracy of each amino acid coding site per gene (R Core Team 2017).

### ASSIGNMENT OF SEX LINKAGE

Z-linkage in *D. plexippus*, including the presence of a neo-Z segment, was previously characterized using a combination of synteny with other Lepidoptera and differential sequencing coverage between males and females (Mongue et al. 2017). Z-linkage in *M. sexta* has also been previously assessed, although only via synteny (Kanost et al. 2016). To directly assess Z-linkage via sequencing coverage differences, we generated new  $\sim 16\times$  coverage Illumina sequencing from a female *M. sexta* and compared coverage with a male sample with comparable sequencing depth (S35, from Mongue et al. 2019) by aligning to a repeat masked version of the reference. We used *BEDtools* (Quinlan and Hall 2010) to calculate the median coverage for each scaffold (to avoid the skewing effect of read pile-ups around repetitive sequence that can bias the mean values) and normalized scaffold medians for each sample by dividing by the mean of all medians. We then assessed linkage by taking the  $\log_2$  of the male:female coverage ratio for each scaffold. Under this metric, Z-linked scaffolds are expected to group around 1 (indicating a twofold greater sequencing depth in males than females), whereas autosomal scaffolds cluster around 0 (equal coverage between the sexes). Formally, we took all scaffolds above the N90 length with a  $\log_2(\text{M:F}) > 0.75$  to be Z-linked.

**GENE EXPRESSION AND ASSESSMENT OF SEX-BIAS**

Gene expression levels (FPKM) for *M. sexta* were used as published from a large RNA-seq dataset with numerous tissue-specific samples (Cao and Jiang 2017). We limited our analysis to tissues with comparable male and female data: adult heads and antennae, as well as adult and pupal gonads. Although heads had four replicate observations, all other tissues were represented by a single replicate.

Gene expression analysis in *D. plexippus* was based on RNA-seq data we previously generated, only some of which has been reported in previous publications (Gu et al. 2019; Mongue et al. 2019). The complete dataset employed here consists of triplicate samples from adults of both sexes generated from a single outbred laboratory population for head, midgut, thorax, gonad, and accessory glands (male only); see the Supporting Information for accessions of all samples. RNA extraction and library construction were performed contemporaneously for all samples, with details as reported in Gu et al. (2019). Using the OGS2 annotation, we aligned and quantified read counts with RSEM (Li and Dewey 2011), then normalized to FPKM values with Trinity using a TMM scaling factor (Grabherr et al. 2011). We averaged the three replicates to give a single expression value per tissue and sex.

The sampling structure for expression data from these two species was heterogeneous. In particular, the lack of replication for many of the *M. sexta* samples substantially limited gene-wise statistical assessments of differential expression between sexes. To accommodate this heterogeneous sampling while also aiming to employ comparable approaches between species, we assessed sex-bias using a tissue-aggregated measure of expression specificity. Namely, we calculated the specificity metric (SPM) for male versus female expression for each annotated gene (Kryuchkova-Mostacci and Robinson-Rechavi 2017). We summed FPKM in each sex and divided by the number of replicates for that tissue in that sex to obtain a mean value for each sex and tissue combination. In the main results, we present analyses on all annotated genes with nonzero expression, but we also confirmed that our results were not driven by spurious assignment of sex-bias in genes with very low expression. In the Supporting Information, we present analyses for all genes with FPKM >5 in at least one sex, similar to Assis et al. (2012) (who used FPKM >4). For the genes under consideration, we calculated SPM as the square of expression in one sex divided by the sum of squared expression in both sexes. This resulted in specificity values ranging from 0 to 1, inclusive, indicating what proportion of a given gene's expression was unique to one sex. As implemented here, an SPM = 1 indicates completely female-specific expression, SPM = 0 indicates male-specific expression, and SPM = 0.5 reflects unbiased expression between the sexes.

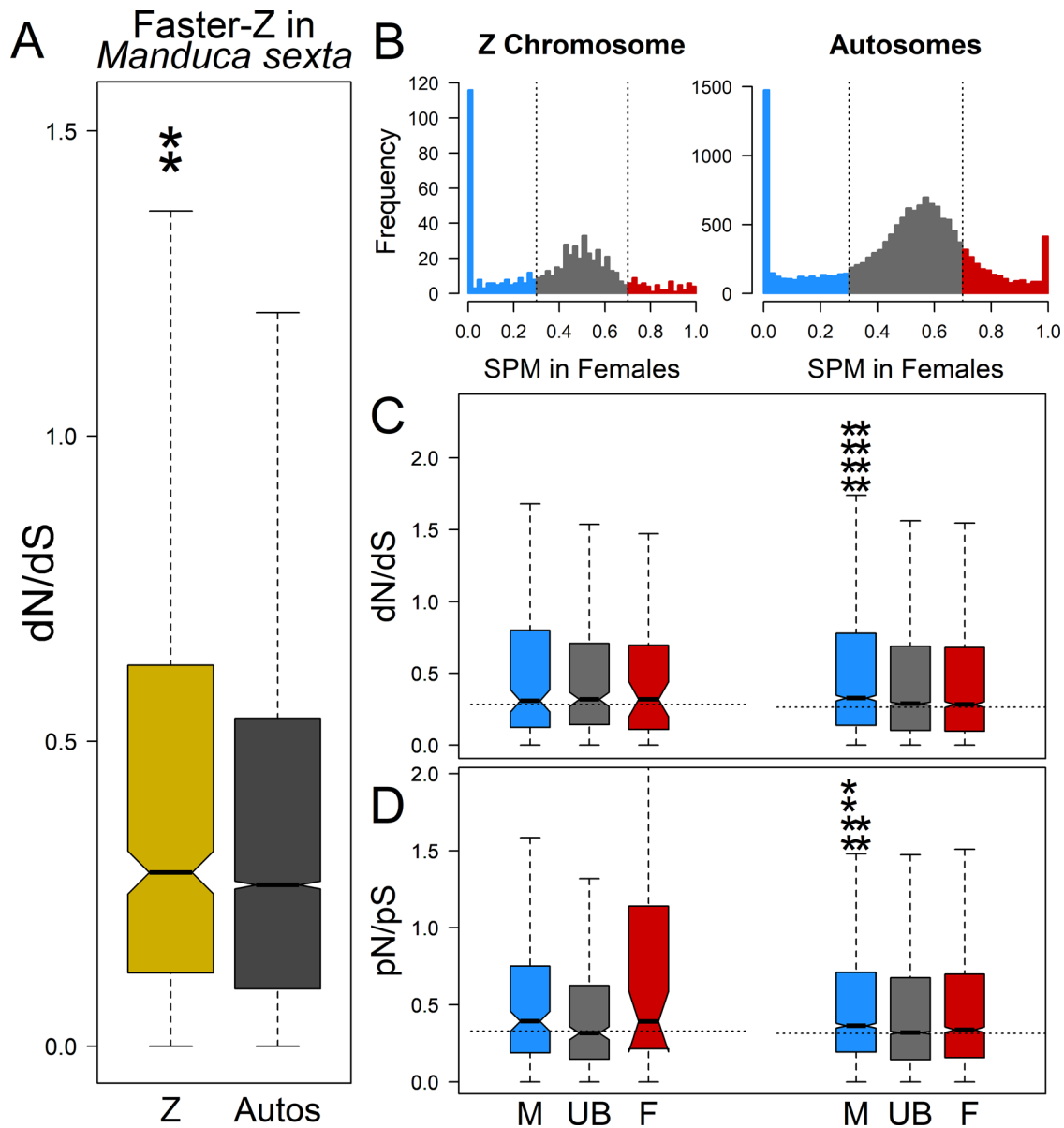
We sought to make our methodology comparable to existing studies that use fold-change in expression to delineate sex-biased genes. In those analyses, sex-bias cutoffs are typically  $1.5\times$  difference in expression between males and females (e.g., in Pinharanda et al. 2019). This difference corresponds to a 70–30 bias in SPM. Thus, we classified female-biased genes as those with SPM >0.7 in females, male-biased genes with SPM <0.3, and unbiased genes that fell within the range of 0.3–0.7 (see Fig. 1B, F for visualizations of these categories).

Although this SPM approach flexibly accommodates the heterogeneous structure of available samples, one potential weakness is that it does not provide an assessment of statistical significance for sex-bias (i.e., differential expression between sexes). To increase confidence in the patterns we report for evolution of the Z chromosome, we verified that our results were robust to the chosen SPM thresholds by re-analyzing the sex-bias data using a much stricter bias, requiring 85% of a gene's expression to be limited to one sex to classify it as sex-biased. These results were qualitatively the same as the more permissive bias cutoff, so we only present the former here and the latter in the supplement.

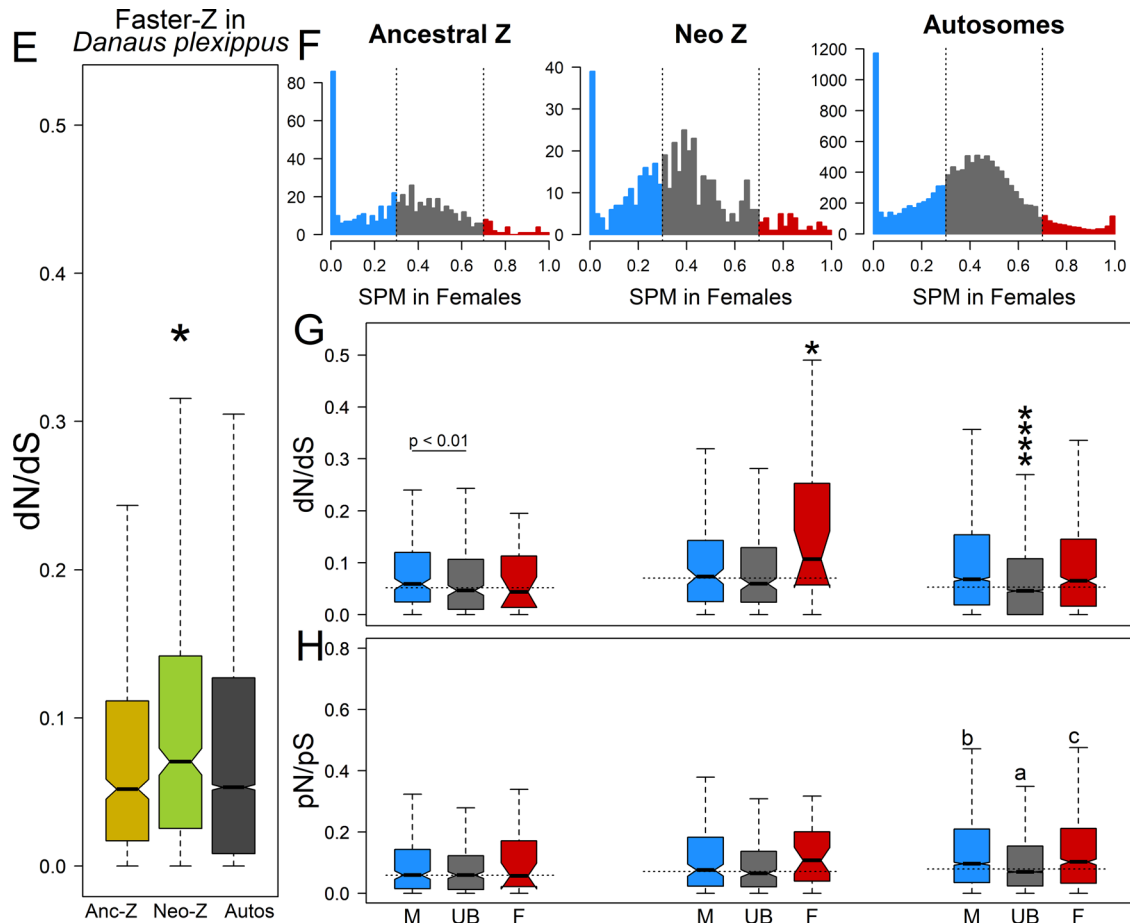
To further show that the SPM approach provides a valid and informative assessment of sex-biased expression, we performed a typical differential expression analysis on read counts from the *D. plexippus* RNA-seq data with DESeq2, using an adjusted *P*-value cutoff of <0.1 to define significantly sex-biased genes (Love et al. 2014). All other genes that passed the expression minimum but were not significantly biased were labeled unbiased. We found strong agreement in categorization of sex-biased genes between SPM and DESeq, with the caveat that the latter is more conservative in defining sex-biased genes. Crucially, the two methods give equivalent results when used to test adaptive evolution of sex-biased genes. A more detailed explanation of this comparison can be found in the Supporting Information.

It has been shown previously that both the *D. plexippus* and *M. sexta* Z chromosomes are masculinized based on distributions of genes encoding sperm proteins (Mongue and Walters 2017), but this expression dataset affords the opportunity to validate those results with a more complete set of sex-biased genes identified above. We used  $\chi^2$  tests of independence to assess whether the proportion of sex-biased genes differed between the autosomes and (neo-)Z chromosomes.

Finally, it is possible that the effective population size of the Z is smaller than its census size in the population (Vicoso and Charlesworth 2009). To investigate this, we identified putatively neutral (fourfold degenerate) sites across the genome, and used the genomics tool ANGSD (Korneliussen et al. 2014) to estimate heterozygosity (Watterson's  $\Theta$ ) for all fourfold degenerate sites on the Z and autosomes separately. We then took the ratio of the mean per-site heterozygosity of the two regions as our estimator



**Figure 1.** Faster-Z evolution in *Manduca sexta* and *Danaus plexippus*. Throughout, asterisks represent statistical differences of one group from all others to which it is compared, with the number of asterisks indicating the level of significance (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$  and below). Horizontal lines with significance annotations are given for significant pairwise differences. (A) The Z evolves faster than the autosomes in *M. sexta*. (B) The distributions of sex-bias for both Z-linked (left) and autosomal (right) genes are plotted with dashed lines to indicate the traditional cutoff points for sex-bias analysis. Bias is plotted such that higher SPM values are more female biased in expression, whereas values closer to 0 are male-biased. (C) Rates of divergence for genes in each sex-bias class (M: male-biased, UB: unbiased, F: female-biased). In *M. sexta*, only autosomal genes show differences between rates of evolution of genes with different sex-bias. (D) Likewise, male-biased genes have higher pN/pS than on other bias classes, but only on the autosomes. (E) The neo-Z is the source of a faster-Z signal in *D. plexippus*. (F) Again we plot distributions of sex-bias categories for genes on the ancestral Z (left), neo-Z (middle), and autosomes (right). (G) Male-biased genes evolve more quickly on the ancestral Z. Female-biased genes evolve more quickly on the neo-Z, and unbiased genes evolve more slowly on the autosomes. (H) Finally, sex-biased genes hold different levels of polymorphism on the autosomes, with unbiased genes having the lowest pN/pS, followed by male-biased, then female-biased with the highest (graphically represented as a  $< b < c$ ).



**Figure 1.** Continued.

for the difference in effective population size between the sex chromosome and autosomes of each species.

### STATISTICAL ANALYSIS OF MOLECULAR EVOLUTION

Because divergence and polymorphism rates are not normally distributed, we analyzed molecular evolution with a series of non-parametric tests. Initially, we tested for a faster-Z effect by comparing the scaled rate of divergence ( $dN/dS$ ) of autosomal and Z-linked genes using Kruskal–Wallis tests with either one degree of freedom in *M. sexta* or two degrees of freedom in *D. plexippus* to account for three potential classes of linkage (autosomal, ancestral Z, and neo-Z). Next, we assessed the effect of sex-biased gene expression (e.g. male- or female-limited expression) on rates of evolution with another set of Kruskal–Wallis tests to determine if there was an effect of sex-bias. In the case of significant results, we investigated pair-wise post hoc differences with a Nemenyi test (Nemenyi 1962; Pohlert 2014). Equivalent tests examining the effects of sex bias and sex linkage on scaled rates of polymorphism ( $pN/pS$ ) were performed for the within-species data.

We combined the polymorphism and divergence data to calculate  $\alpha$ , the proportion of substitutions driven by adaptive evo-

lution. Specifically, we used a calculation of the neutrality index (NI; Stoletzki and Eyre-Walker 2011) for each class of genes to give us a point estimate of  $\alpha$  ( $= 1 - NI$ ) summed across genes within a bias class and linkage group. We assessed significance via a permutation test framework, as in Mongue et al. (2019). We compared evolution of two gene classes, calculated the point estimate  $\alpha$  for each, and then took the absolute value of the difference of these estimates as our permutation test statistic. Next, we combined the two gene sets and randomly drew two permuted classes of sizes equal to the true classes without replacement. We calculated the absolute difference in  $\alpha$  for these two random gene sets for 10,000 permutations. In doing so, we built a distribution of differences in point estimates of  $\alpha$  that could be expected by chance alone. We then compared our true value to this distribution and took the  $P$ -value to be the proportion of times we observed a greater value in the permuted distribution than the true value.

To verify our inferences based on SNP calling, we also used ANGSD to estimate  $\pi$  and Tajima's  $D$  at both fourfold and zero-fold degenerate sites across the genome. We examined differences between the autosomes and the Z in both species but did

not further partition the genomic regions by sex-bias owing to limitations in ANGSD's ability to generate meaningful priors for small portions of the genome. Finally, we assessed the potential for differences in linkage disequilibrium across the genome using the `-geno-r2` option in VCFtools (Danecek et al. 2011) to assess the correlation coefficient ( $\rho^2$ ) between unphased genotypes in 50 base-pair windows along the genome. For all of these ancillary population genetic statistics, we tested for differences between (parts of) the Z and the autosomes using nonparametric tests, specifically the Mann–Whitney–Wilcoxon test (the pairwise equivalent of the Kruskal–Wallis test).

## Results

### ASSIGNMENT OF SEX-LINKAGE IN *M. sexta*

Based on previous synteny analyses comparing *M. sexta* to *Bombyx mori*, 27 scaffolds were annotated as Z-linked in the *M. sexta* assembly (Kanost et al. 2016). By using previously sequenced male and newly sequenced female genomic DNA to calculate male-female coverage differences, we validated these previously annotated scaffolds and identified nine additional scaffolds as Z-linked. We considered only scaffolds above the genome N90 (45Kb) to avoid coverage differences that could arise by chance on short sequences. We considered all scaffolds with  $\log_2(\text{M:F}) > 0.75$  as z-linked. The data showed no ambiguous scaffolds by coverage, with two clearly separated distributions, one centered around 0 (autosomes) and another, smaller set of scaffolds centered around 1 [Z-linked scaffold range: (0.80, 1.20)]. The visualization of these distributions can be seen in Figure S1. We recovered all previously annotated 27 scaffolds as Z-linked and identified an additional nine Z-linked scaffolds, spanning 2.1Mb and containing an additional 43 annotated genes. Seven of these newly identified scaffolds were previously not assigned to any chromosome owing to unclear sequence homology. The remaining two were previously annotated autosomal based on linkage of *Bombyx* orthologs but are clearly Z-linked in coverage bias. These scaffolds are relatively gene-poor ( $\leq 10$  annotated genes each) and may represent small-scale gene trafficking events between *Manduca* and *Bombyx* but are unlikely to be the product of a large-scale fusion. This updated linkage information is included as a datasheet in the Supporting Information.

### SEX-BIAS ON THE Z CHROMOSOMES

Based on the assignment of sex-biased genes from the RNA-sequencing data (head, antennae, and gonad in *M. sexta*; head, thorax, midgut, and gonad in *D. plexippus*), the gene content differs between the Z and autosomes in both *M. sexta* ( $\chi^2_2 = 47.37$ ,  $P = 5.2 \times 10^{-11}$ ) and *D. plexippus* ( $\chi^2_2 = 30.04$ ,  $P = 3.0 \times 10^{-7}$ ). In both species, this difference comes from an ex-

cess of male-biased genes on the Z chromosome, as well as a paucity of female-biased genes on the *M. sexta* Z and unbiased genes on the *D. plexippus* ancestral Z (Table 1). These results hold for both traditional cutoffs for sex bias and for stricter criteria (see the Supporting Information). It is worth noting that the excess of male-biased genes on the Z chromosome is not the result of dosage effects, as both *M. sexta* and *D. plexippus* have been shown to have sex-balanced expression on the Z (Smith et al. 2014; Gu et al. 2019).

### DIVERGENCE BETWEEN SPECIES

The Z chromosome has higher scaled divergence than the autosomes in both species: *M. sexta* ( $\chi^2_1 = 6.89$ ,  $P = 0.009$ ; Fig. 1A and Table 2) and *D. plexippus* ( $\chi^2_2 = 9.72$ ,  $P = 0.008$ ). For *D. plexippus*, we further classified the Z into the ancestral (i.e., long-term sex-linked) and neo-Z (the Z sequence resulting from an autosomal fusion). Based on the significant chromosomal linkage effect, we conducted post hoc testing and found that the signal for faster-Z evolution comes primarily from the neo-Z, which diverges distinctly faster than the autosomes ( $P = 0.006$ ; Fig. 1E) and marginally faster than the ancestral Z ( $P = 0.048$ ; Table 2). On its own, the ancestral Z is not faster evolving than the autosomes ( $P = 0.99$ ).

In *M. sexta*, divergence rates did not differ between genes with differing sex-bias patterns on the Z chromosome ( $\chi^2_2 = 1.12$ ,  $P = 0.571$ ; Fig. 1C). On the autosomes, however, there was a clear effect of sex-biased expression ( $\chi^2_2 = 26.26$ ,  $P = 1.98 \times 10^{-6}$ ). Post hoc testing revealed this to be driven largely by male-biased genes, which have higher divergence rates than unbiased ( $P = 8.1 \times 10^{-6}$ ) or female-biased genes ( $P = 4.2 \times 10^{-5}$ ). Female-biased genes do not evolve at a different rate than unbiased genes ( $P = 0.63$ ).

In *D. plexippus* as well, evolutionary rates of autosomal loci varied with sex-bias class ( $\chi^2_2 = 249$ ,  $P < 1.0 \times 10^{-10}$ ; Fig. 1G). Unlike *M. sexta*, however, the effect of sex-bias did not differ between sexes. Both male-biased ( $P < 1.0 \times 10^{-10}$ ) and female-biased genes ( $P < 1.0 \times 10^{-10}$ ) evolve faster than unbiased genes, but male-biased and female-biased genes do not evolve differently from each other ( $P = 0.75$ ).

Considering the *D. plexippus* Z chromosome, both the ancestral ( $\chi^2_2 = 9.99$ ,  $P = 0.007$ ) and neo ( $\chi^2_2 = 11.85$ ,  $P = 0.003$ ; Fig. 1G) segments showed a sex-bias effect. For the ancestral Z, this difference is driven solely by faster evolution of male-biased genes compared to unbiased genes ( $P = 0.005$ ); evolutionary rates of female-biased genes did not differ significantly from the unbiased nor male-biased genes on the ancestral Z. On the neo-Z, female-biased genes evolve faster than both male-biased ( $P = 0.044$ ) and unbiased genes ( $P = 0.002$ ); divergence of male-biased genes did not differ from unbiased on the neo-Z.

**Table 1.** Sex bias of the Z chromosomes in the two species studied with gene counts and proportions in parentheses.

	Carolina sphinx moth ( <i>Manduca sexta</i> )		Monarch butterfly ( <i>Danaus plexippus</i> )		
	Autosomes	Z	Autosomes	Ancestral Z	Neo-Z
Male-biased	2477 (0.21)	177 (0.34)	4721 (0.35)	279 (0.47)	184 (0.39)
Unbiased	7219 (0.63)	295 (0.56)	7529 (0.56)	278 (0.46)	243 (0.52)
Female-biased	1856 (0.16)	55 (0.10)	1248 (0.09)	44 (0.07)	41 (0.09)

Sex bias is based on expression analysis of heads, antennae, and gonads in *M. sexta* and heads, thoraces, midguts, and gonads in *D. plexippus*. In both species, composition of the Z differs from composition of the autosomes due to an increased proportion of male-biased Z-linked genes (based on  $\chi^2$  *P*-values  $< 1.0 \times 10^{-6}$ ; note that this significant result holds in *D. plexippus* whether the Z is considered as one category or two [i.e., neo and ancestral]). The *M. sexta* Z is depleted for female-biased genes, whereas the monarch (ancestral)-Z is depleted for unbiased genes.

**Table 2.** Population genetic parameters across the genomes of both Lepidoptera.

	<i>M. sexta</i>		<i>D. plexippus</i>		
	Autosomes	Z	Autosomes	Ancestral Z	Neo-Z
dN	0.0037 ( $\pm 0.016$ )	<b>0.0049</b> ( $\pm 0.008$ )	0.0016 ( $\pm 0.006$ )	<b>0.0032</b> ( $\pm 0.006$ )	<b>0.0044</b> ( $\pm 0.009$ )
dS	0.0158 ( $\pm 0.028$ )	<b>0.0181</b> ( $\pm 0.034$ )	0.0347 ( $\pm 0.045$ )	<b>0.0667</b> ( $\pm 0.048$ )	<b>0.0750</b> ( $\pm 0.054$ )
dN/dS	0.2589 ( $\pm 0.665$ )	<b>0.2805</b> ( $\pm 0.817$ )	0.0583 ( $\pm 0.366$ )	0.0570 ( $\pm 0.255$ )	<b>0.0757</b> ( $\pm 0.269$ )
pN	<b>0.0068</b> ( $\pm 0.022$ )	0.0034 ( $\pm 0.019$ )	<b>0.0043</b> ( $\pm 0.012$ )	0.0020 ( $\pm 0.006$ )	<b>0.0037</b> ( $\pm 0.007$ )
pS	<b>0.0232</b> ( $\pm 0.050$ )	0.0104 ( $\pm 0.069$ )	<b>0.0545</b> ( $\pm 0.053$ )	0.0346 ( $\pm 0.039$ )	<b>0.0522</b> ( $\pm 0.040$ )
pN/pS	0.3056 ( $\pm 0.282$ )	0.3188 ( $\pm 0.325$ )	<b>0.0908</b> ( $\pm 0.245$ )	0.0678 ( $\pm 0.203$ )	0.0776 ( $\pm 0.115$ )
$\pi_N$	<b><math>8.82 \times 10^{-9}</math></b> ( $\pm 0.030$ )	$4.68 \times 10^{-9}$ ( $\pm 0.022$ )	<b><math>1.83 \times 10^{-5}</math></b> ( $\pm 0.029$ )	$7.80 \times 10^{-6}$ ( $\pm 0.022$ )	<b><math>1.03 \times 10^{-5}</math></b> ( $\pm 0.024$ )
$\pi_S$	<b><math>5.96 \times 10^{-8}</math></b> ( $\pm 0.082$ )	$2.23 \times 10^{-8}$ ( $\pm 0.055$ )	<b><math>1.36 \times 10^{-4}</math></b> ( $\pm 0.080$ )	$5.19 \times 10^{-5}$ ( $\pm 0.061$ )	<b><math>1.16 \times 10^{-4}</math></b> ( $\pm 0.080$ )
Tajima's $D_0$	-0.0795 ( $\pm 0.433$ )	<b>-0.0419</b> ( $\pm 0.319$ )	<b>-0.3341</b> ( $\pm 0.673$ )	<b>-0.2749</b> ( $\pm 0.613$ )	-0.3692 ( $\pm 0.696$ )
Tajima's $D_4$	-0.0259 ( $\pm 0.510$ )	<b>-0.0171</b> ( $\pm 0.392$ )	<b>-0.2702</b> ( $\pm 0.631$ )	<b>-0.2734</b> ( $\pm 0.643$ )	<b>-0.2584</b> ( $\pm 0.629$ )
$\rho^2$	0.3546 ( $\pm 0.396$ )	<b>0.3577</b> ( $\pm 0.398$ )	0.1364 ( $\pm 0.308$ )	0.1437 ( $\pm 0.317$ )	0.1430 ( $\pm 0.316$ )

Median values are given for divergence and polymorphism estimates (to avoid skew from outliers), whereas means are reported for Tajima's *D* (as in every case, the median value is centered on zero). Mean linkage disequilibrium ( $\rho^2$ ) reported for 50 basepair windows. Standard deviations appear in parentheses. Bolded numbers are significantly higher than the other other(s) in their category; see results for exact *P*-values. Bolded and underlined numbers are higher than both others (e.g., dN on *D. plexippus* neo Z > ancestral Z > autosomes). Patterns are consistent with reduced within-population variation on the *Manduca* Z and *Danaus* ancestral Z relative to the autosomes at both putatively neutral and selected sites. The neo-Z however holds roughly as much variation as the *Danaus* autosomes. The neo-Z is also notable in having the most negative Tajima's *D* value in the *D. plexippus* genome at selected sites.

### GENETIC VARIATION WITHIN SPECIES

In *M. sexta*, the scaled levels of nonsynonymous polymorphism did not differ between the Z and autosomes ( $\chi^2_1 = 2.57$ ,  $P = 0.110$ ). However, separately both silent (pS:  $W = 3,243,400$ ,  $P < 1.0 \times 10^{-10}$  and  $\pi_S$ :  $W = 2,635,900,000$ ,  $P < 1.0 \times 10^{-10}$ ) and nonsilent (pN:  $W = 3,194,500$ ,  $P < 1.0 \times 10^{-10}$  and  $\pi_N$ :  $W = 44,765,000,000$ ,  $P < 1.0 \times 10^{-10}$ ) were significantly lower for the Z than the autosomes (Table 2). Scaled polymorphism differed between the different sex bias classes (Figure 1D;  $\chi^2_2 = 43.45$ ,  $P = 3.7 \times 10^{-10}$ ). Here again, male-biased genes showed increased nonsynonymous variation compared to unbiased genes ( $P = 1.4 \times 10^{-10}$ ) and female-biased genes ( $P = 0.002$ ). Female-biased and unbiased genes did not significantly differ from each other ( $P = 0.14$ ).

In *D. plexippus*, polymorphism strongly differed between the Z and autosomes (Figure 1H;  $\chi^2_2 = 34.18$ ,  $P = 38 \times 10^{-8}$ ). Both the ancestral Z ( $P = 3.9 \times 10^{-7}$ ) and neo-Z ( $P = 0.02$ ) had lower levels of scaled nonsynonymous polymorphism than the autosomes, but the two portions of the Z did not differ from each other ( $P = 0.27$ ). Individually, pN and pS were both higher on the autosomes than the ancestral Z (pN:  $P < 1.0 \times 10^{-10}$ , pS:  $P < 1.0 \times 10^{-10}$ ), as well as the neo-Z compared to the ancestral Z (pN:  $P = 7.3 \times 10^{-8}$ , pS:  $P < 1.0 \times 10^{-10}$ ) but the autosomes and neo-Z did not differ from one another by these metrics (pN:  $P = 0.10$ , pS:  $P = 0.86$ ). When considering only  $\pi$  at either zero-fold or fourfold degenerate sites, however, we recovered the pattern that the autosomes had the highest levels of variation ( $\pi_N$ :  $P < 1.0 \times 10^{-10}$  vs. neo-Z and ancestral Z,  $\pi_S$ :  $P < 1.0 \times 10^{-10}$  vs.



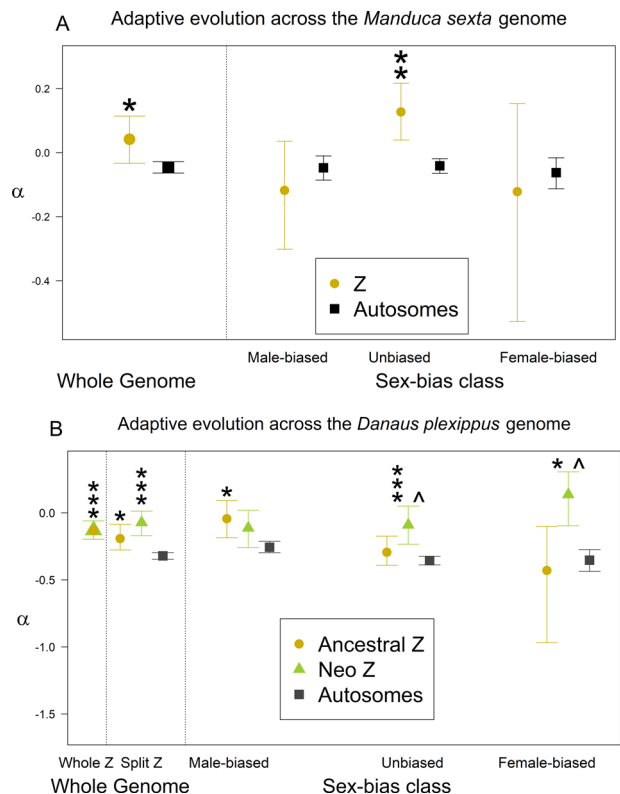
neo-Z and ancestral Z) followed by the neo-Z ( $\pi_N$ :  $P < 1.0 \times 10^{-10}$  vs. ancestral Z,  $\pi_S$ :  $P < 1.0 \times 10^{-10}$  vs. ancestral Z), then the ancestral Z (see Table 2 for point estimates).

Genes of differing sex-bias class did not vary in rates of polymorphism on either part of the Z (ancestral:  $X^2_2 = 2.70$ ,  $P = 0.259$ ; neo:  $X^2_2 = 5.75$ ,  $P = 0.06$ ). In contrast, autosomal genes did differ: female-biased genes showed the highest rates of polymorphism, higher than male-biased ( $P = 1.8 \times 10^{-10}$ ) or unbiased genes ( $P < 1.0 \times 10^{-10}$ ); male-biased genes had elevated rates of polymorphism compared to unbiased genes ( $P < 1.0 \times 10^{-10}$ ).

### EVIDENCE FOR ADAPTIVE EVOLUTION

To examine rates of adaptation, we estimated the proportion of adaptive substitutions ( $\alpha$ ) first for Z versus autosomes as a whole, then further partitioning loci by sex-biased expression. In *M. sexta*, the Z overall showed more adaptive evolution than the autosomes ( $P = 0.039$ ), in spite of slightly, albeit significantly, higher Tajima's D values at both non-silent ( $W = 2737900000$ ,  $P = 0.0002$ ) and silent ( $W = 2920800000$ ,  $P = 1.6 \times 10^{-9}$ ) sites (Table 2). Adaptation of male-biased ( $P = 0.340$ ) and female-biased genes ( $P = 0.812$ ) did not differ based on genomic location, but genes with unbiased expression showed higher rates of adaptive evolution ( $\alpha$ ) on the Z chromosome than the autosomes ( $P = 0.007$ ; Fig. 2A), in spite of non-significant differences in dN/dS for unbiased genes. Instead, this result stems from a marginally higher dN/dS and marginally lower pN/pS in combination.

In *D. plexippus*, the Z also exhibited increased rates of adaptation compared to autosomes ( $P = 0.0004$ ; Fig. 2B, left). Considered separately, both the ancestral and neo-Z segments evolved more adaptively than the autosomes (ancestral-Z vs. autosomes:  $P = 0.0338$ , neo-Z vs. autosomes:  $P = 0.0005$ ). The neo-Z segment trended towards more adaptive evolution than the ancestral Z, but not strongly ( $P = 0.079$ ). Estimates of Tajima's D also reinforce the notion of stronger directional selection on the neo-Z, where D values at zero-fold degenerate sites were significantly more negative than the autosomes ( $W = 547230000$ ,  $P < 1.0 \times 10^{-10}$ ) or the ancestral Z ( $W = 1078300000$ ,  $P = 0.0052$ , Table 2). Regarding sex-bias, we found that male-biased genes evolved more adaptively on the ancestral Z than the autosomes ( $P = 0.0177$ ) but that differences in adaptation could not be distinguished between the neo-Z and the rest of genome (autosomal vs. neo-Z  $P = 0.318$ , ancestral vs. neo-Z  $P = 0.500$ ). In contrast, female-biased genes evolved more adaptively on the neo-Z than the autosomes ( $P = 0.0474$ ) or ancestral Z ( $P = 0.008$ ). Additionally, ancestrally Z-linked female-biased genes did not evolve differently than their autosomal counterparts ( $P = 0.539$ ). Furthermore, unbiased genes on the neo-Z showed greater rates of



**Figure 2.** Adaptive evolution across the genomes of the two Lepidoptera considered in this study. In each panel, coarse-scale comparison of the Z chromosome to autosomes is plotted left of the dotted lines. Points are the point estimate of the  $\alpha$  statistic and error bars represent 95% confidence intervals for each point estimate obtained by parametric bootstrapping. Significant differences are noted with an \* for differences between the Z and autosomes, and a  $\Delta$  for differences between parts of the Z in *D. plexippus*. In *M. sexta* (A), the Z evolves more adaptively than the autosomes overall (left of dash). This pattern appears to be driven by unbiased genes (right of dash). In *D. plexippus* (B), the whole Z is more adaptively evolving than the autosomes (leftmost), and both the ancestral and neo-Z segments show elevated  $\alpha$  compared to the autosomes (middle). For the ancestral Z, male-biased genes drive the increase in adaptation; in contrast, unbiased and female-biased genes are more adaptively evolving on the neo-Z.

adaptation than unbiased genes on the autosomes ( $P = 0.018$ ) or ancestral Z ( $P = 0.048$ ).

### THE EFFECTIVE POPULATION SIZE OF THE Z CHROMOSOME

Under simple biological conditions, we expect the ratio of Z:Autosomes population sizes to be 0.75 (Wilson Sayres 2018); however, because female Lepidoptera have achiasmatic meiosis (i.e., chromosomes do not undergo recombination), this expectation may be naïve (Turner and Sheppard 1975). We examined levels of diversity (Watterson's  $\Theta$ ) at fourfold degenerate (i.e.,

putatively neutral) sites on the Z and autosomes and took the ratio of the means of these two classes to be an estimator for the difference in effective population size. We found that, in practice, this ratio for *M. sexta* is much lower than expected ( $Ne_Z:Ne_A = 0.44$ ). For *D. plexippus*, the difference in population sizes is less skewed ( $Ne_Z:Ne_A = 0.66$ ). Intriguingly, this difference is not uniform across the *D. plexippus* Z. The ancestral portion of the Z has a lower population size,  $Ne_{Z\_Anc}:Ne_A = 0.58$ , but the neo-Z holds essentially as much diversity as the autosomes,  $Ne_{Z\_Neo}:Ne_A = 0.98$ . In line with these expectations, we found that linkage disequilibrium across 50 base-pair windows was higher on the Z than the autosomes in *M. sexta* ( $\rho^2 = 0.358$  vs.  $0.355$ ,  $W = 7.13 \times 10^{12}$ ,  $P < 1.0 \times 10^{-10}$ ); conversely, for *D. plexippus*, linkage disequilibrium did not differ across the Z or autosomes ( $Z_{Anc}$  vs.  $Z_{Neo}$ :  $W = 5.59 \times 10^{10}$ ,  $P = 0.5147$ ;  $Z_{Anc}$  vs. Autos:  $W = 1.62 \times 10^{12}$ ,  $P = 0.3904$ ;  $Z_{Neo}$  vs. Autos:  $W = 1.76 \times 10^{12}$ ,  $P = 0.06867$ ). Comparing between species,  $\rho^2$  was consistently lower in *D. plexippus* ( $Z_{Anc}$ : 0.144,  $Z_{Neo}$ : 0.143, Autos: 0.136) than in *M. sexta*.

## Discussion

### NEW EVIDENCE FOR A FASTER-Z

Although previous evidence for faster-Z evolution in Lepidoptera has been mixed, we found that the Z chromosome is faster evolving (i.e., has elevated dN/dS) than the autosomes in two distantly related Lepidoptera: *M. sexta* and *Danaus plexippus*. At first pass, our results seemingly suggest a long-term faster-Z evolution, bolstered by similar results in silkmoths (Sackton et al. 2014), but at odds with other studies in butterflies (Rousselle et al. 2016; Pinharanda et al. 2019). However, a more nuanced consideration indicates some congruence with both sets of studies. *Danaus plexippus* shows an overall faster Z, but this result is driven by the neo-Z portion of the chromosome evolving faster than the autosomes. Considering only the ancestral portion, which is homologous to the Z of the butterflies previously studied, there is no evidence for increased divergence on the ancestral-Z in *D. plexippus*. Nevertheless, evidence for higher rates of adaptive evolution ( $\alpha$ ) on the Z is less ambiguous in our insects; both *M. sexta* and *D. plexippus* showed overall more adaptation for Z-linked genes, as reported in *B. mori*.

Beginning with the simpler case in *M. sexta*, we found that increased adaptation on the Z chromosome is driven by genes with unbiased expression. These genes are haploid expressed in females and should experience more efficient selection than unbiased genes on the autosomes (which are always diploid in expression). Female-biased genes should follow this pattern as well, but the lack of a clear signal might be attributable to the small number of female-biased genes on the Z, which reduces our power to detect differences. Moreover, the effective population size of the *M.*

*sexta* Z compared to the autosomes is much lower than the neutral expectation (0.44 vs. 0.75). With such a decrease in the population of Z chromosomes, selection is predicted to be less efficient (Vicoso and Charlesworth 2009) and may further limit the adaptive evolution of female-biased genes. These lines of reasoning track well with Tajima's *D*, which is less negative for selected sites on the Z than the autosomes in this species. Thus, a combination of weakened positive selection for male-biased genes and less population growth on the sex chromosome overall appears to explain the differences with the autosomes in *M. sexta*.

Sex chromosome evolution in *Danaus* presents a more complicated case than that of *M. sexta*, owing to the Z-autosomal fusion in this genus (Mongue et al. 2017). This fusion event added a large number of previously autosomal genes to the Z. Intriguingly, it is the neo-Z that best fits with predictions for adaptive Z evolution; increased adaptation is concentrated in unbiased and female-biased genes, which are more abundant on the neo-Z than the ancestral Z. Similarly, Tajima's *D* is at its most negative in the genome on the neo-Z. In principle, this could arise through either recurrent positive selection, as suggested by differences in  $\alpha$ , or a large expansion in population size relative to the autosomes. It is worth noting that the neo-Z has an inferred effective population size nearly equal to that of the autosomes ( $Ne_{Z\_neo}:Ne_{Autos} = 0.98$ ). This is an unexpected result that is difficult to explain.

To begin with, the parity cannot be attributed to sequence homology with a neo-W. Any existing neo-W chromosome must be highly divergent from the neo-Z because neither alignment of sequencing data nor in situ hybridization of labeled probes indicates any conserved sequence between the Z and W or remaining autosomes (see Mongue et al. 2017 for details; Gu et al. 2019), so there is no evidence for anything like a W-linked "pseudo-autosomal region" to explain comparable Z versus autosomal heterozygosity. Such parity may also arise due to biased sex ratios or greater variance in the reproductive success of the heterogametic sex, as seen in other taxa (Hedrick 2007; Ellegren 2009). A skewed sex ratio seems unlikely in this case, as only a male-biased population would restore parity to the Z:A ratio. *Danaus plexippus* has one of the most closely monitored populations of any insect (Oberhauser and Solensky 2004), and no such dynamics have been observed (on the contrary, another *Danaus* species is known for male-killing genetic elements [Smith et al. 2016]). High variance in female reproductive success could generate similar effective population sizes for the Z and autosomes (Vicoso and Charlesworth 2009). However, available evidence indicates that *Danaus* females very rarely fail to mate in the wild, so variance in female reproduction also does not explain the observed neo-Z versus autosomal population sizes (Pliske 1973). Ultimately, both sex ratio bias and female mating variance, even if they occurred, should theoretically affect the ancestral and neo-

Z equally, and thus would not explain the discrepancy observed between the two portions of the Z.

A more plausible, albeit complicated, explanation involves the lack of recombination in female Lepidoptera, leaving male meiosis as the only opportunity for recombination (Turner and Sheppard 1975). With equal sex ratios, in a given generation only one half of the autosomes will recombine, but two thirds of the Z chromosomes undergo recombination. This elevation in relative recombination rate can aid in adaptive evolution by decoupling deleterious alleles and bringing together beneficial variants; as such, linkage disequilibrium should decay faster on the Z than the autosomes, leading to less of a reduction in effective population size associated with selective events. In other words, the default prediction for the lepidopteran Z to autosome ratio might be closer to 1 than 0.75. In a similar vein, population growth has been shown to impact genetic diversity on the sex chromosomes more than the autosomes (Pool and Nielsen 2007), and *D. plexippus* has apparently undergone recent population expansion in North America (Zhan et al. 2014; Mongue et al. 2019). Under this paradigm, the neo-Z fits the expectation, but the ancestral-Z has much lower-than-expected diversity. This observation, along with the male-biased composition of the ancestral Z, fits with the observed strong purifying selection on male-biased genes (as observed on the autosomes of *D. plexippus* in Mongue et al. 2019). Purifying selection on male-biased genes on the ancestral Z, positive selection of novel beneficial female-biased variants on the neo-Z, and the relatively high recombination of the Z may act to decouple the effective population sizes of the neo- and ancestral Z. Lepidoptera are generally observed to have one crossover event per chromosome per male meiosis (linkage maps from two highly diverged species both estimate the average chromosome size at about 50 centimorgans: Yamamoto et al. 2008; Davey et al. 2017), which would be enough to separate the evolutionary trajectory of the two halves of the Z.

Examining patterns of linkage disequilibrium, we found that linkage was comparable across both halves of the *D. plexippus* Z and the autosomes. In absolute terms, linkage disequilibrium was much lower in *D. plexippus* than in *M. sexta*. These results suggest that linkage should decay quickly enough on the neo-Z to separate it from linked selection on the ancestral Z, but they also point to lineage-specific effects that differentiate the two species we study here. One obvious difference is that, although both are broadly distributed North American insects, migratory *D. plexippus* form a massive panmictic population across the continent (Lyons et al. 2012) but *M. sexta* populations are geographically structured, with at least one segregating Z-linked inversion (Mongue and Kawahara 2020), meaning that starting pool of recombining alleles should be much larger in *D. plexippus*.

Whatever the cause of this difference, the high effective population size of the neo-Z should permit selection to remove dele-

terious variation more efficiently on the neo-Z than on the autosomes for all dominance coefficients of mutations (Vicoso and Charlesworth 2009). Moreover, the dosage of the neo-Z is compensated differently to that of the ancestral Z. Although the ancestral Z is downregulated in males such that expression is balanced between the sexes ( $ZZ\downarrow = Z$ ), the neo-Z is upregulated in females to create balance ( $Z\uparrow = ZZ$ , Gu et al. 2019). If, as theory predicts, the selective importance of variants is related to the level of their expression (Vicoso and Charlesworth 2009), then the relatively higher expression of the neo-Z and lower expression of the ancestral Z also help explain the differing rates of molecular evolution across the *D. plexippus* Z chromosome.

### RECONCILING EXISTING INVESTIGATIONS OF LEPIDOPTERAN Z CHROMOSOME EVOLUTION

Our results most strongly agree with existing work from the silkworm genus *Bombyx* (Sackton et al. 2014), which found both fast and adaptive Z effects. Efforts in other butterflies have found no faster-Z effect. In the case of satyr butterflies, this negative result may be attributable to “noisy” sequence data (de novo transcriptome assemblies) and potential uncertainty in Z-linkage (which was inferred from sequence homology alone) (Rousselle et al. 2016). In the case of *Heliconius* butterflies, it is worth noting that point estimates for  $\alpha$  and dN/dS largely fit predictions for a fast and adaptive Z, but results did not differ significantly between the Z and autosomes thanks to high variance in these estimates, especially on the Z chromosomes (Pinharanda et al. 2019). In this case, the use of a relatively small RNA-sequencing dataset limited the number of sex-biased genes with which to work; only 200 of ~700 total Z-linked genes were analyzed.

Nonetheless, these lepidopteran faster-Z studies suggest a phylogenetic signal for Z chromosome evolution. *Bombyx* and *Manduca* are from sister families of moths (Kawahara and Breinholt 2014) and share patterns of faster and more adaptive Z evolution. Satyrs, *Heliconius*, and *Danaus* butterflies all fall within the family Nymphalidae and show mixed to negative evidence for increased divergence and adaptation on the (ancestral) Z. In other words, there is more agreement for Z chromosome evolution for more closely related species. These observations demonstrate that sex-linkage per se does not lead to consistent evolutionary outcomes for the genes involved. Instead faster-Z evolution likely depends on the demographic history or degree of sex-bias of the Z chromosomes examined. This is illustrated by the relatively young neo-Z in *Danaus*, which is not masculinized like the ancestral Z and instead appears comparable to autosomes in the proportion of unbiased and female-biased genes (Mongue and Walters 2017). The neo-Z fits completely within the theoretical prediction for adaptive faster-Z evolution, evolving faster due to increased adaptation of unbiased and female-biased genes that are subject to haploid selection (Charlesworth et al. 1987).

These observations raise the possibility that faster-Z dynamics may be transient rather than perpetual. Adaptive evolution of the sex chromosomes is thought to be driven by the hemizygous expression of some genes in one sex (Charlesworth et al. 1987), but depending on the dominance of gene expression, genes benefitting the opposite sex are predicted to accumulate on that sex chromosome (Rice 1984). As such, if the sex chromosomes change composition over evolutionary time, they may bias toward alleles benefitting the homogametic sex (e.g., male-benefit, male-biased genes on the Z). Genes with haploid expression (e.g., unbiased or female-biased genes) will become less abundant and thus less important to the evolution of the chromosome. Moreover, if sexual selection produces high variance in male reproductive success, the effective population size of Z chromosomes can be depressed below the census size, further limiting the role of positive selection on the few unbiased or female-biased left on the Z. Particularly old sex chromosomes should be more likely to experience these effects.

This dynamic may also explain the abundance of neo-Z chromosomes in Lepidoptera (Nguyen et al. 2013; Nguyen and Paladino 2016; Mongue et al. 2017). Conserved synteny implies that small-scale gene trafficking events are rare (but evidence is somewhat contradictory here as well; see Troups et al. 2011; Wang et al. 2012) and fusion-fission events may be the key source of linkage shuffling. For a highly masculinized Z chromosome, a sudden influx of unbiased and female-biased genes can create strong positive selection and favor these fused chromosomes, even at initially low frequencies. Under this paradigm, even the seemingly contradictory findings on Z chromosome evolution can be reconciled as being the product of lineage-specific differences in sex-biased gene content and chromosomal history. If this line of reasoning is accurate, it should be borne out by investigating other, independently evolved neo-Z chromosomes.

#### AUTHOR CONTRIBUTIONS

AJM conceived of this study, performed analyses, and wrote the manuscript text. MEH contributed expression data analyses and edited the manuscript. JRW provided theoretical expertise, analysis guidance, additional expression analyses, and edited the manuscript.

#### ACKNOWLEDGMENTS

The authors wish to acknowledge W. Mason and M. Hulet and the rest of the Information and Telecommunication Technology Center (ITTC) staff at the University of Kansas for their support with our high-performance computing. We are continually grateful to C. Taylor and A. Ryan from MonarchWatch for access to butterflies and to C. Sorenson for logistical and moral support in obtaining sphinx moths. Thanks, as well, to A. Mackintosh and other frequenters of the Darwin Dance Hall for conversations during the interpretation of these results. This work was funded by National Science Foundation award DEB-1701931 to AJM and ABI-1661454 to JRW.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### DATA ARCHIVING

*Manduca sexta* resequencing data can be found on NCBI's Sequence Read Archive with the following accessions: SRP144217 and PRJNA639154. *Danaus plexippus* RNA sequencing can be found with PRJNA522622. The *M. sexta* expression data can be found as a supplementary table in Cao and Jiang (2017) (<https://doi.org/10.1186/s12864-017-4147-y>). The *D. plexippus* sequencing data can be found in Zhan et al. (2014), (<https://doi.org/10.1038/nature13812>). Please see the Supporting Information of this manuscript for specific samples used in these analyses. Analysis scripts and input data files can be found at <https://github.com/WaltersLab/FastZ>.

#### LITERATURE CITED

- Ahola, V., R. Lehtonen, P. Somervuo, L. Salmela, P. Koskinen, P. Rastas, N. Välimäki, L. Paulin, J. Kvist, N. Wahlberg, et al. 2014. The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. *Nat. Commun.* 5:1–9.
- Assis, R., Q. Zhou, and D. Bachtrog. 2012. Sex-biased transcriptome evolution in *Drosophila*. *Genome Biol. Evol.* 4:1189–1200.
- Bachtrog, D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat. Rev. Genet.* 14:113–124.
- Baines, J. F., S. A. Sawyer, D. L. Hartl, and J. Parsch. 2008. Effects of X-linkage and sex-biased gene expression on the rate of adaptive protein evolution in *Drosophila*. *Mol. Biol. Evol.* 25:1639–1650.
- Bateman, A. J. 1948. Intra-sexual selection in *Drosophila*. *Heredity* 2:349–368.
- Cao, X., and H. Jiang. 2017. An analysis of 67 RNA-seq datasets from various tissues at different stages of a model insect, *Manduca sexta*. *BMC Genomics*. 18:796.
- Chapman, T., G. Arnqvist, J. Bangham, and L. Rowe. 2003. Sexual conflict. *Trends Ecol. Evol.* 18:41–47.
- Charlesworth, B., and D. Charlesworth. 2000. The degeneration of Y chromosomes. *Philos. Trans. R Soc. B Biol. Sci.* 355:1563–1572.
- Charlesworth, B., J. A. Coyne, and N. H. Barton. 1987. The relative rates of evolution of sex chromosomes and autosomes. *Am. Nat.* 130:113–146.
- Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w 1118; iso-2; iso-3. *Fly* 6:80–92.
- Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, and S. T. Sherry. 2011. The variant call format and VCFtools. *Bioinformatics* 27:2156–2158.
- Dapper, A. L., and M. J. Wade. 2016. The evolution of sperm competition genes: the effect of mating system on levels of genetic variation within and between species. *Evolution* 70:502–511.
- Davey, J. W., M. Chouteau, S. L. Barker, L. Maroja, S. W. Baxter, F. Simpson, M. Joron, J. Mallet, K. K. Dasmahapatra, and C. D. Jiggins. 2016. Major improvements to the *Heliconius melpomene* genome assembly used to confirm 10 chromosome fusion events in 6 million years of butterfly evolution. *G3* 6:695–708.
- Davey, J. W., S. L. Barker, P. M. Rastas, A. Pinharanda, S. H. Martin, R. Durbin, W. O. McMillan, R. M. Merrill, and C. D. Jiggins. 2017. No evidence for maintenance of a sympatric *Heliconius* species barrier by chromosomal inversions. *Evol. Lett.* 1:138–154.

- Dean, R., P. W. Harrison, A. E. Wright, F. Zimmer, and J. E. Mank. 2015. Positive selection underlies faster-Z evolution of gene expression in birds. *Mol. Biol. Evol.* 32:2646–2656.
- Ellegren, H. 2009. The different levels of genetic diversity in sex chromosomes and autosomes. *Trends Genet.* 25:278–284.
- Ellegren, H., L. Hultin-Rosenberg, B. Brunström, L. Dencker, K. Kultima, and B. Scholz. 2007. Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. *BMC Biol.* 5:40–52.
- Fraïsse, C., M. A. Picard, and B. Vicoso. 2017. The deep conservation of the Lepidoptera Z chromosome suggests a non-canonical origin of the W. *Nat. Commun.* 8:1486.
- Gershoni, M., and S. Pietrokovski. 2014. Reduced selection and accumulation of deleterious mutations in genes exclusively expressed in men. *Nat. Commun.* 5:4438.
- Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, and Q. Zeng. 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* 29:644–652.
- Griffiths, R., M. C. Double, K. Orr, and R. J. G. Dawson. 1998. A DNA test to sex most birds. *Mol. Ecol.* 7:1071–1075.
- Gu, L., and J. R. Walters. 2017. Evolution of sex chromosome dosage compensation in animals: a beautiful theory, undermined by facts and be-deviled by details. *Genome Biol. Evol.* 9:2461–2476.
- Gu, L., P. F. Reilly, J. J. Lewis, R. D. Reed, P. Andolfatto, and J. R. Walters. 2019. Dichotomy of dosage compensation along the neo Z chromosome of the monarch butterfly. *Curr. Biol.* 29:4071–4077.e3.
- Hayes, K., H. J. Barton, and K. Zeng. 2020. A study of faster-Z evolution in the great tit (*Parus major*). *Genome Biol. Evol.* 12:210–222.
- Hedrick, P. W. 2007. Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. *Evolution* 61:2750–2771.
- Kanost, M. R., E. L. Arrese, X. Cao, Y.-R. Chen, S. Chellapilla, M. Goldsmith, E. Grosse-Wilde, D. G. Heckel, N. Herndon, and H. Jiang. 2016. Multifaceted biological insights from a draft genome sequence of the tobacco hornworm moth, *Manduca sexta*. *Insect Biochem. Mol. Biol.* 76:118–147.
- Kawahara, A. Y., and J. W. Breinholt. 2014. Phylogenomics provides strong evidence for relationships of butterflies and moths. *Proc. R Soc B: Biol. Sci.* 281:20140970.
- Korneliussen, T. S., A. Albrechtsen, and R. Nielsen. 2014. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics* 15:356.
- Kousathanas, A., D. L. Halligan, and P. D. Keightley. 2014. Faster-X adaptive protein evolution in house mice. *Genetics* 196:1131–1143.
- Kryuchkova-Mostacci, N., and M. Robinson-Rechavi. 2017. A benchmark of gene expression tissue-specificity metrics. *Brief. Bioinform.* 18:205–214.
- Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* 9:357–359.
- Li, B., and C. N. Dewey. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323.
- Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:1–21.
- Lunter, G., and M. Goodson. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res.* 21:936–939.
- Lyons, J. I., A. A. Pierce, S. M. Barribeau, E. D. Sternberg, A. J. Mongue, and J. C. de Roode. 2012. Lack of genetic differentiation between monarch butterflies with divergent migration destinations. *Mol. Ecol.* 21:3433–3444.
- Mank, J. E., K. Nam, and H. Ellegren. 2009. Faster-Z evolution is predominantly due to genetic drift. *Mol. Biol. Evol.* 27:661–670.
- Mank, J. E., D. J. Hosken, and N. Wedell. 2014. Conflict on the sex chromosomes: cause, effect, and complexity. *Cold Spring Harb. Perspect. Biol.* 6:a017715.
- McKenna, A. H., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernysky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, et al. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20:1297–1303.
- Meisel, R. P., and T. Connallon. 2013. The faster-X effect: integrating theory and data. *Trends Genet.* 29:537–544.
- Meisel, R. P., J. H. Malone, and A. G. Clark. 2012. Disentangling the relationship between sex-biased gene expression and X-linkage. *Genome Res.* 22:1255–1265.
- Mongue, A. J., and A. Y. Kawahara. 2020. Population differentiation and structural variation in the genome of *Manduca sexta* across the United States. *bioRxiv* <https://doi.org/10.1101/2020.11.01.364000>.
- Mongue, A. J., and J. Walters. 2017. The Z chromosome is enriched for sperm proteins in two divergent species of Lepidoptera. *Genome* 61:248–253.
- Mongue, A. J., P. Nguyen, A. Volenikova, and J. R. Walters. 2017. Neo-sex chromosomes in the Monarch butterfly, *Danaus plexippus*. *G3* 7:3281–3294.
- Mongue, A. J., M. E. Hansen, L. Gu, C. E. Sorenson, and J. R. Walters. 2019. Nonfertilizing sperm in Lepidoptera show little evidence for recurrent positive selection. *Mol. Ecol.* 28:2517–2530.
- Nemenyi, P. 1962. Distribution-free multiple comparisons. Ph.D. diss. Princeton University, Princeton, NJ.
- Nguyen, P., and L. C. Paladino. 2016. On the neo-sex chromosomes of Lepidoptera. Pp. 171–185 in P. Pontarotti, ed. *Evolutionary biology*. Springer, Berlin.
- Nguyen, P., M. Sýkorová, J. Šíchová, V. Kůta, M. Dalíková, R. Čapková Frydrychová, L. G. Neven, K. Sahara, and F. Marec. 2013. Neo-sex chromosomes and adaptive potential in tortricid pests. *Proc. Natl Acad. Sci. USA* 110:6931–6936.
- Oberhauser, K. S., and M. J. Solensky. 2004. *The monarch butterfly: biology and conservation*. Cornell Univ. Press, Ithaca, NY.
- Ohta, T. 1992. The nearly neutral theory of molecular evolution. *Annu. Rev. Ecol. Syst.* 23:263–286.
- Parsch, J., and H. Ellegren. 2013. The evolutionary causes and consequences of sex-biased gene expression. *Nat. Rev. Genet.* 14:83–87.
- Pinharanda, A., M. Rousselle, S. H. Martin, J. J. Hanly, J. W. Davey, S. Kumar, N. Galtier, and C. D. Jiggins. 2019. Sexually dimorphic gene expression and transcriptome evolution provide mixed evidence for a fast-Z effect in *Heliconius*. *J. Evol. Biol.* 32:194–204.
- Pliske, T. E. 1973. Factors determining mating frequencies in some New World butterflies and skippers. *Ann. Entomol. Soc. Am.* 66:164–169.
- Pohlert, T. 2014. The pairwise multiple comparison of mean ranks package (PMCMR). *R Package* 27:10.
- Pool, J. E., and R. Nielsen. 2007. Population size changes reshape genomic patterns of diversity. *Evolution* 61:3001–3006.
- Quinlan, A. R., and I. M. Hall. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842.
- R Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available via <http://www.R-project.org/>.
- Rice, W. R. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* 38:735–742.
- Rousselle, M., N. Faivre, M. Ballenghien, N. Galtier, and B. Nabholz. 2016. Hemizygoty enhances purifying selection: lack of fast-Z evolution in two satyrine butterflies. *Genome Biol. Evol.* 8:3108–3119.

- Sackton, T. B., R. B. Corbett-Detig, J. Nagaraju, L. Vaishna, K. P. Arunkumar, and D. L. Hartl. 2014. Positive selection drives faster-Z evolution in silkworms. *Evolution* 68:2331–2342.
- Smith, D. A. S., I. J. Gordon, W. Traut, J. Herren, S. Collins, D. J. Martins, K. Saitoti, and P. Ireri. 2016. A neo-W chromosome in a tropical butterfly links colour pattern, male-killing, and speciation. *Proc. R. Soc. B Biol. Sci.* 283:20160821.
- Smith, G., Y. R. Chen, G. W. Blissard, and A. D. Briscoe. 2014. Complete dosage compensation and sex-biased gene expression in the moth *Manduca sexta*. *Genome Biol. Evol.* 6:526–537.
- Stoletzki, N., and A. Eyre-Walker. 2011. Estimation of the neutrality index. *Mol. Biol. Evol.* 28:63–70.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Toups, M. A., J. B. Pease, and M. W. Hahn. 2011. No Excess gene movement is detected off the avian or lepidopteran Z chromosome. *Genome Biol. Evol.* 3:1381–1390.
- Turner, J. R. G., and P. M. Sheppard. 1975. Absence of crossing-over in female butterflies (*Heliconius*). *Heredity* 34:265–269.
- Vicoso, B., and B. Charlesworth. 2009. Effective population size and the faster-X effect: an extended model. *Evolution*. 63:2413–2426.
- Walters, J. R., and T. J. Hardcastle. 2011. Getting a full dose? Reconsidering sex chromosome dosage compensation in the silkworm, *Bombyx mori*. *Genome Biol. Evol.* 3:491–504.
- Wang, J., M. Long, and M. D. Vibranovski. 2012. Retrogenes moved out of the Z chromosome in the silkworm. *J. Mol. Evol.* 74:113–126.
- Wang, Z., J. Zhang, W. Yang, N. An, P. Zhang, G. Zhang, and Q. Zhou. 2014. Temporal genomic evolution of bird sex chromosomes. *BMC Evol. Biol.* 14:250.
- Whittle, C. A., A. Kulkarni, and C. G. Extavour. 2020. Absence of a faster-X effect in beetles (*Tribolium*, Coleoptera). *G3* 10:1125–1136.
- Wilson Sayres, M. A. 2018. Genetic diversity on the sex chromosomes. *Genome Biol. Evol.* 10:1064–1078.
- Wright, A. E., H. K. Moghadam, and J. E. Mank. 2012. Trade-off between selection for dosage compensation and masculinization on the avian Z chromosome. *Genetics* 192:1433–1445.
- Wright, A. E., P. W. Harrison, F. Zimmer, S. H. Montgomery, M. A. Pointer, and J. E. Mank. 2015. Variation in promiscuity and sexual selection drives avian rate of Faster-Z evolution. *Mol. Ecol.* 24:1218–1235.
- Xu, L., S. Y. Wa Sin, P. Grayson, S. V. Edwards, and T. B. Sackton. 2019. Evolutionary dynamics of sex chromosomes of paleognathous birds. *Genome Biol. Evol.* 11:2376–2390.
- Yamamoto, K., J. Nohata, K. Kadono-Okuda, J. Narukawa, M. Sasanuma, S. Sasanuma, H. Minami, M. Shimomura, Y. Suetsugu, and Y. Banno. 2008. A BAC-based integrated linkage map of the silkworm *Bombyx mori*. *Genome Biol.* 9:1–14.
- Zhan, S., and S. M. Reppert. 2013. MonarchBase: the monarch butterfly genome database. *Nucleic Acids Res.* 41:D758–D763.
- Zhan, S., W. Zhang, K. Niitepöld, J. Hsu, J. F. Haeger, M. P. Zalucki, S. Altizer, J. C. de Roode, S. M. Reppert, and M. R. Kronforst. 2014. The genetics of monarch butterfly migration and warning colouration. *Nature* 514:317–321.

Associate Editor: D. Filatov  
Handling Editor: T. Chapman

## Supporting Information

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

**Table S1.** Sex bias of the Z chromosomes in the two species studied with a stricter definition of sex-biased expression.

**Table S2.** Sex bias of the Z chromosomes in the two species studied with a higher minimum expression cutoff (FPKM >5).

**Table S3.** Comparison of sex-bias labeling from traditional differential expression analysis and SPM.

**Table S4.** Sex-bias of the Z chromosome using DEseq definitions.

**Figure S1.** Sex-linkage assignment by difference in male:female coverage of genomic scaffolds.