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

Running title: Fast and adaptive Z chromosome evolution

Andrew J. Mongue[1,2,*], Megan E. Hansen[1], and James R. Walters[1]

[1]University of Kansas, Department of Ecology and Evolutionary Biology, 1200 Sunnyside Ave, Lawrence, KS, USA, 66045

[2]University of Edinburgh, Institute of Evolutionary Biology, Charlotte Auebach Road, Edinburgh, UK, EH93FL

*correspondence contact: ajmongue@gmail.com

Conflicts of interest

The authors declare no conflicts of interest.

Funding sources

This work was funded by NSF award DEB-1701931 to AJM and NSF award ABI-1661454 to JRW.

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.

Acknowledgements

The authors wish to acknowledge Wesley Mason and Michael 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 Chip Taylor and Ann Ryan from MonarchWatch for access to butterflies and to Clyde Sorenson for logistical and moral support in obtaining sphinx moths. Thanks, as well, to Alex Mackintosh and other frequenters of the Darwin Dance Hall for conversations during the interpretation of these results.

Data accessibility

Manduca sexta resequencing data can be found on NCBI's Sequence Read Archive with the following accessions: SRP144217, 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 supplement to this manuscript for specific samples used in these analyses. Analysis scripts and input data files can be found at https://github.com/WaltersLab/FastZ.

Abstract

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, though 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.

Introduction

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 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 census 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), though 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 four-fold degenerate sizes). 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) while 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 one 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 allows 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 *Manduca 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 population 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), while heterospecific reads were aligned to the same reference with *stampy* v.1.0.22 with an increased allowance for mismatches to better align divergent data (default parameters with the exception of substitutionrate = 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 & Fisher Strand-bias < 60 & Mapping Quality > 40); from the remaining single nucleotide variants, we classified each as synonymous or non-synonymous using *SNPeff* (v. 4.2, Cingolani et al. 2012) and normalized variant counts by the number of non-synonymous 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, though only via synteny (Kanost et al. 2016). To directly assess Z-linkage via sequencing coverage differences, we generated new ~16x 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 two-fold greater sequencing depth in males than females), while autosomal scaffolds cluster around 0 (equal coverage between the sexes). Formally, we took all scaffolds above the N90 length with a $\log_2(M:F) > 0.75$ 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. While 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 data set 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 supplement 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 non-zero 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 supplement we present analyses for all genes with FPKM > 5 in at least one sex, similar to Assis et al. (who used FPKM > 4, 2012). For the genes under consideration, we calculated SPM as the square of expression in one sex divided

240 by the sum of squared expression in both sexes. This resulted in specificity values ranging from
241 0 to 1, inclusive, indicating what proportion of a given gene's expression was unique to one sex.
242 As implemented here, an SPM = 1 indicates completely female-specific expression, SPM = 0
243 indicates male-specific expression, and SPM = 0.5 reflects unbiased expression between the
244 sexes.

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

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

259 To further show that the SPM approach provides a valid and informative assessment of sex-
260 biased expression, we performed a typical differential expression analysis on read counts from
261 the *D. plexippus* RNA-seq data with DESeq2, using an adjusted p-value cutoff of < 0.1 to define
262 significantly sex-biased genes. (Love et al. 2014). All other genes which 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 supplement.

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 X^2 tests of independence to assess whether or not 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 (four-fold degenerate) sites across the genome, and used the genomics tool ANGSD (Korneliussen et al. 2014) to estimate heterozygosity (Watterson's Θ) for all four-fold 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 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 1 degree of freedom in *M. sexta* or 2 degrees of freedom in *D. plexippus*

to account for 3 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 α, the proportion of substitutions driven by adaptive evolution. 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 α (= 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 α for each, 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 α for these two random gene sets for 10,000 permutations. In doing so, we built a distribution of differences in point estimates of α 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 π and Tajima's D at both four-fold 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 (ρ^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 non-parametric tests, specifically the Mann-Whitney-Wilcoxon test (the pairwise equivalent of the Kruskal-Wallis test).

Results

Assignment of sex-linkage in *Manduca sexta*

Based on previous synteny analyses comparing *Manduca 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 9 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 $\log_2(M:F) > 0.75$ 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 supplemental Figure S1. We recovered all previously annotated 27 scaffolds as Z-linked and identified an additional 9 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 (≤ 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 supplementary datasheet.

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* ($X^2_2 = 47.37$, $p = 5.2*10^{-11}$) and *D. plexippus* ($X^2_2 = 30.04$, $p = 3.0*10^{-7}$). In both species, this difference comes from an excess 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 supplement). 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* ($X^2_1 = 6.89$, $p = 0.009$, Figure 1A, Table 2) and *D. plexippus* ($X^2_2 = 9.72$, $p = 0.008$). 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 ($p = 0.006$,

Figure 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 ($X^2_2 = 1.12$, $p = 0.571$, Figure 1C). On the autosomes however, there was a clear effect of sex-biased expression ($X^2_2 = 26.26$, $p = 1.98*10^{-6}$). Post-hoc testing revealed this to be driven largely by male-biased genes, which have higher divergence rates than unbiased ($p = 8.1*10^{-6}$) or female-biased genes ($p = 4.2*10^{-5}$). 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 ($X^2_2 = 249$, $p < 1.0 * 10^{-10}$, Figure 1G). Unlike *M. sexta* however, the effect of sex-bias did not differ between sexes. Both male-biased ($p < 1.0*10^{-10}$) and female-biased genes ($p < 1.0 * 10^{-10}$) 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 ($X^2_2 = 9.99$, $p = 0.007$) and neo ($X^2_2 = 11.85$, $p = 0.003$, Figure 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.

Genetic variation within species

In *M. sexta*, the scaled levels of non-synonymous polymorphism did not differ between the Z and autosomes ($X^2_1 = 2.57$, $p = 0.110$). However, separately both silent (pS: $W = 3243400$, $p < 1.0 * 10^{-10}$ and π_S: $W = 2635900000$, $p < 1.0 * 10^{-10}$) and non-silent (pN: $W = 3194500$, $p < 1.0 * 10^{-10}$ and π_N: $W = 44765000000$, $p < 1.0 * 10^{-10}$), were significantly lower for the Z than the autosomes (Table 2). Scaled polymorphism differed between the different sex bias classes (Figure 1D, $X^2_2 = 43.45$, $p = 3.7 * 10^{-10}$). Here again, male-biased genes showed increased non-synonymous variation compared to unbiased genes ($p = 1.4 * 10^{-10}$) 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, $X^2_2 = 34.18$, $p = 38 * 10^{-8}$). Both the ancestral Z ($p = 3.9 * 10^{-7}$) and neo-Z ($p = 0.02$) had lower levels of scaled non-synonymous 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: $p < 1.0 * 10^{-10}$, pS: $p < 1.0 * 10^{-10}$), as well as the neo-Z compared to the ancestral Z (pN: $p = 7.3*10^{-8}$, pS: $p < 1.0 * 10^{-10}$) 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 π at either zero-fold or four-fold degenerate sites however, we recovered the pattern that the autosomes had the highest levels of variation (π_N: $p < 1.0 * 10^{-10}$ vs neo-Z and ancestral Z , π_S: $p < 1.0 * 10^{-10}$ vs neo-Z and ancestral Z) followed by the neo-Z (π_N: $p < 1.0 * 10^{-10}$ vs ancestral Z , π_S: $p < 1.0 * 10^{-10}$ 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 ($p = 1.8*10^{-10}$) or unbiased genes ($p < 1.0 * 10^{-10}$); male-biased genes had elevated rates of polymorphism compared to unbiased genes ($p < 1.0 * 10^{-10}$).

Evidence for adaptive evolution

To examine rates of adaptation, we estimated the proportion of adaptive substitutions (α) 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 ($W = 2920800000$, $p = 1.6*10^{-9}$) 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 (α) on the Z chromosome than the autosomes ($p = 0.007$; Figure 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 ($p = 0.0004$; Figure 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 ($W = 547230000$, $p < 1.0 * 10^{-10}$) 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 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 Θ) at four-fold 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* (ρ^2 = 0.358 vs 0.355, W = $7.13*10^{12}$, p < $1.0 * 10^{-10}$); conversely, for *D. plexippus*, linkage disequilibrium did not differ across the Z or autosomes (Z_{Anc} vs Z_{Neo}: W = $5.59*10^{10}$, p = 0.5147; Z_{Anc} vs Autos: W =

442 $1.62*10^{12}$, p = 0.3904; Z_{Neo} vs Autos: W = $1.76*10^{12}$, p = 0.06867). Comparing between species,
443 ρ^2 was consistent lower in *D. plexippus* (Z_{Anc}: 0.144, Z_{Neo}: 0.143, Autos: 0.136).

444 ## Discussion

445 ### New evidence for a faster-Z

446 While previous evidence for faster-Z evolution in Lepidoptera has been mixed, we found that the
447 Z chromosome is faster evolving (*i.e.* has elevated dN/dS) than the autosomes in two distantly
448 related Lepidoptera: *Manduca sexta* and *Danaus plexippus*. At first pass, our results seemingly
449 suggest a long-term faster-Z evolution, bolstered by similar results in silkmoths (Sackton et al.
450 2014), but at odds with other studies in butterflies (Rousselle et al. 2016; Pinharanda et al. 2019).
451 However, a more nuanced consideration indicates some congruence with both sets of studies. *D.*
452 *plexippus* shows an overall faster Z, but this result is driven by the neo-Z portion of the
453 chromosome evolving faster than the autosomes. Considering only the ancestral portion, which is
454 homologous to the Z of the butterflies previously studied, there is no evidence for increased
455 divergence on the ancestral-Z in *D. plexippus*. Nevertheless, evidence for higher rates of adaptive
456 evolution (α) on the Z is less ambiguous in our insects; both *M. sexta* and *D. plexippus* showed
457 overall more adaptation for Z-linked genes, as reported in *Bombyx mori*.

458 Beginning with the simpler case in *M. sexta*, we found that increased adaptation on the Z
459 chromosome is driven by genes with unbiased expression. These genes are haploid expressed in
460 females and should experience more efficient selection than unbiased genes on the autosomes
461 (which are always diploid in expression). Female-biased genes should follow this pattern as well,
462 but the lack of a clear signal might be attributable to the small number of female-biased genes on
463 the Z, which reduces our power to detect differences. Moreover, the effective population size of
464 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 explains 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 adaption 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 α, 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 indicate 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 vs 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 deleterious 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. While the ancestral Z is down-regulated in males such that expression is balanced between the sexes (ZZ↓ = Z), the neo-Z is upregulated in females to create balance (Z↑=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 silkmoth 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 α 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 towards alleles benefitting the homogametic sex (e.g. male-benefitting, 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: Toups 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.

References

Ahola, V., R. Lehtonen, P. Somervuo, L. Salmela, P. Koskinen, P. Rastas, N. Välimäki, L. Paulin, J. Kvist, N. Wahlberg, J. Tanskanen, E. A. Hornett, L. C. Ferguson, S. Luo, Z. Cao, M. A. de Jong, A. Duplouy, O.-P. Smolander, H. Vogel, R. C. McCoy, K. Qian, W. S. Chong, Q. Zhang, F. Ahmad, J. K. Haukka, A. Joshi, J. Salojärvi, C. W. Wheat, E. Grosse-Wilde, D. Hughes, R. Katainen, E. Pitkänen, J. Ylinen, R. M. Waterhouse, M. Turunen, A. Vähärautio, S. P. Ojanen, A. H. Schulman, M. Taipale, D. Lawson, E. Ukkonen, V. Mäkinen, M. R. Goldsmith, L. Holm, P. Auvinen, M. J. Frilander, and I. Hanski. 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 biology and evolution 4:1189–1200. Oxford University Press.

Bachtrog, D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nature reviews. Genetics 14:113–24.

603 Baines, J. F., S. A. Sawyer, D. L. Hartl, and J. Parsch. 2008. Effects of X-Linkage and Sex-Biased Gene
604 Expression on the Rate of Adaptive Protein Evolution in Drosophila. Molecular Biology and
605 Evolution 25:1639–1650.
606 Bateman, A. J. 1948. Intra-sexual selection in Drosophila. Heredity 2:349–368. Essex.
607 Cao, X., and H. Jiang. 2017. An analysis of 67 RNA-seq datasets from various tissues at different stages of
608 a model insect, *Manduca sexta*. BMC Genomics 18.
609 Chapman, T., G. Arnqvist, J. Bangham, and L. Rowe. 2003. Sexual conflict. Trends in Ecology & Evolution
610 18:41–47.
611 Charlesworth, B., and D. Charlesworth. 2000. The degeneration of Y chromosomes. Philosophical
612 Transactions of the Royal Society B-Biological Sciences 355:1563–1572.
613 Charlesworth, B., J. A. Coyne, and N. H. Barton. 1987. The relative rates of evolution of sex
614 chromosomes and autosomes. The American Naturalist 130:113–146.
615 Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. 2012.
616 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff:
617 SNPs in the genome of *Drosophila melanogaster* strain w 1118; iso-2; iso-3. Fly 6:80–92.
618 Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T.
619 Marth, and S. T. Sherry. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–
620 2158. Oxford University Press.
621 Dapper, A. L., and M. J. Wade. 2016. The evolution of sperm competition genes: The effect of mating
622 system on levels of genetic variation within and between species. Evolution 70:502–511.
623 Davey, J. W., S. L. Barker, P. M. Rastas, A. Pinharanda, S. H. Martin, R. Durbin, W. O. McMillan, R. M.
624 Merrill, and C. D. Jiggins. 2017. No evidence for maintenance of a sympatric Heliconius species
625 barrier by chromosomal inversions. Evolution letters 1:138–154. Wiley Online Library.

626 Davey, J. W., M. Chouteau, S. L. Barker, L. Maroja, S. W. Baxter, F. Simpson, M. Joron, J. Mallet, K. K.
627 Dasmahapatra, and C. D. Jiggins. 2016. Major improvements to the *Heliconius melpomene*
628 genome assembly used to confirm 10 chromosome fusion events in 6 million years of butterfly
629 evolution. G3: Genes| Genomes| Genetics g3-115. Genetics Society of America.
630 Dean, R., P. W. Harrison, A. E. Wright, F. Zimmer, and J. E. Mank. 2015. Positive selection underlies
631 faster-Z evolution of gene expression in birds. Molecular biology and evolution 32:2646–2656.
632 Ellegren, H. 2009. The different levels of genetic diversity in sex chromosomes and autosomes. Trends in
633 Genetics 25:278–284. Elsevier.
634 Ellegren, H., L. Hultin-Rosenberg, B. Brunström, L. Dencker, K. Kultima, and B. Scholz. 2007. Faced with
635 inequality: chicken do not have a general dosage compensation of sex-linked genes. BMC
636 biology 5:40–52.
637 Fraïsse, C., M. A. Picard, and B. Vicoso. 2017. The deep conservation of the Lepidoptera Z chromosome
638 suggests a non-canonical origin of the W. Nature communications 8:1486.
639 Gershoni, M., and S. Pietrokovski. 2014. Reduced selection and accumulation of deleterious mutations in
640 genes exclusively expressed in men. Nature Communications 5.
641 Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R.
642 Raychowdhury, and Q. Zeng. 2011. Trinity: reconstructing a full-length transcriptome without a
643 genome from RNA-Seq data. Nature biotechnology 29:644. NIH Public Access.
644 Griffiths, R., M. C. Double, K. Orr, and R. J. G. Dawson. 1998. A DNA test to sex most birds. Molecular
645 Ecology 7:1071–1075.
646 Gu, L., P. F. Reilly, J. J. Lewis, R. D. Reed, P. Andolfatto, and J. R. Walters. 2019. Dichotomy of Dosage
647 Compensation along the Neo Z Chromosome of the Monarch Butterfly. Current Biology 29:4071-
648 4077.e3.

649 Gu, L., and J. R. Walters. 2017. Evolution of Sex Chromosome Dosage Compensation in Animals: A
650 Beautiful Theory, Undermined by Facts and Bedeviled by Details. Genome Biology and Evolution
651 9:2461–2476.
652 Hayes, K., H. J. Barton, and K. Zeng. 2020. A Study of Faster-Z Evolution in the Great Tit (Parus major).
653 Genome Biology and Evolution 12:210–222.
654 Hedrick, P. W. 2007. Sex: differences in mutation, recombination, selection, gene flow, and genetic drift.
655 Evolution: International Journal of Organic Evolution 61:2750–2771. Wiley Online Library.
656 Kanost, M. R., E. L. Arrese, X. Cao, Y.-R. Chen, S. Chellapilla, M. Goldsmith, E. Grosse-Wilde, D. G. Heckel,
657 N. Herndon, and H. Jiang. 2016. Multifaceted biological insights from a draft genome sequence
658 of the tobacco hornworm moth, *Manduca sexta*. Insect Biochemistry and Molecular Biology
659 76:118–147. Pergamon.
660 Kawahara, A. Y., and J. W. Breinholt. 2014. Phylogenomics provides strong evidence for relationships of
661 butterflies and moths. Proceedings of the Royal Society B: Biological Sciences 281:20140970.
662 Korneliussen, T. S., A. Albrechtsen, and R. Nielsen. 2014. ANGSD: Analysis of Next Generation
663 Sequencing Data. BMC Bioinformatics 15.
664 Kousathanas, A., D. L. Halligan, and P. D. Keightley. 2014. Faster-X adaptive protein evolution in house
665 mice. Genetics 196:1131–1143.
666 Kryuchkova-Mostacci, N., and M. Robinson-Rechavi. 2017. A benchmark of gene expression tissue-
667 specificity metrics. Briefings in Bioinformatics 18:205–214.
668 Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–
669 359.
670 Li, B., and C. N. Dewey. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or
671 without a reference genome. BMC bioinformatics 12:323. Springer.

672 Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for
673 RNA-seq data with DESeq2. Genome biology 15:1–21. BioMed Central.
674 Lunter, G., and M. Goodson. 2011. Stampy: A statistical algorithm for sensitive and fast mapping of
675 Illumina sequence reads. Genome Research 21:936–939.
676 Lyons, J. I., A. A. Pierce, S. M. Barribeau, E. D. Sternberg, A. J. Mongue, and J. C. de Roode. 2012. Lack of
677 genetic differentiation between monarch butterflies with divergent migration destinations.
678 Molecular Ecology 21:3433–3444.
679 Mank, J. E., D. J. Hosken, and N. Wedell. 2014. Conflict on the sex chromosomes: cause, effect, and
680 complexity. Cold Spring Harbor perspectives in biology 6:a017715. Cold Spring Harbor Lab.
681 Mank, J. E., K. Nam, and H. Ellegren. 2009. Faster-Z evolution is predominantly due to genetic drift.
682 Molecular biology and evolution 27:661–670.
683 McKenna, A. H., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler,
684 S. Gabriel, M. Daly, and M. Depristo. 2010. The Genome Analysis Toolkit: A MapReduce
685 framework for analyzing next-generation DNA sequencing data. Genome Research 20:1297–
686 1303.
687 Meisel, R. P., and T. Connallon. 2013. The faster-X effect: integrating theory and data. Trends in Genetics
688 29:537–544.
689 Meisel, R. P., J. H. Malone, and A. G. Clark. 2012. Disentangling the relationship between sex-biased
690 gene expression and X-linkage. Genome Research 22:1255–1265.
691 Mongue, A. J., M. E. Hansen, L. Gu, C. E. Sorenson, and J. R. Walters. 2019. Nonfertilizing sperm in
692 Lepidoptera show little evidence for recurrent positive selection. Molecular Ecology 28:2517–
693 2530.
694 Mongue, A. J., and A. Y. Kawahara. 2020. Population differentiation and structural variation in the
695 genome of Manduca sexta across the United States. bioRxiv. Cold Spring Harbor Laboratory.

696 Mongue, A. J., P. Nguyen, A. Volenikova, and J. R. Walters. 2017. Neo-sex chromosomes in the Monarch
697 butterfly, Danaus plexippus. G3: Genes|Genomes|Genetics 7:3281–3294.
698 Mongue, A. J., and J. Walters. 2017. The Z chromosome is enriched for sperm proteins in two divergent
699 species of Lepidoptera. Genome 61:248–253. NRC Research Press.
700 Nemenyi, P. 1962. Distribution-free multiple comparisons. P. 263 *in*. International Biometric Soc 1441 I
701 ST, NW, SUITE 700, WASHINGTON, DC 20005-2210.
702 Nguyen, P., and L. C. Paladino. 2016. On the neo-sex chromosomes of Lepidoptera. Pp. 171–185 *in*
703 Evolutionary Biology. Springer.
704 Nguyen, P., M. Sýkorová, J. Šíchová, V. Kůta, M. Dalíková, R. Čapková Frydrychová, L. G. Neven, K.
705 Sahara, and F. Marec. 2013. Neo-sex chromosomes and adaptive potential in tortricid pests.
706 Proceedings of the National Academy of Sciences of the United States of America 110:6931–6.
707 Oberhauser, K. S., and M. J. Solensky. 2004. The Monarch butterfly: biology & conservation. Cornell
708 university press.
709 Ohta, T. 1992. The Nearly Neutral Theory of Molecular Evolution. Annual Review of Ecology and
710 Systematics 23:263–286.
711 Parsch, J., and H. Ellegren. 2013. The evolutionary causes and consequences of sex-biased gene
712 expression. Nature reviews. Genetics 14:83–7.
713 Pinharanda, A., M. Rousselle, S. H. Martin, J. J. Hanly, J. W. Davey, S. Kumar, N. Galtier, and C. D. Jiggins.
714 2019. Sexually dimorphic gene expression and transcriptome evolution provide mixed evidence
715 for a fast-Z effect in Heliconius. Journal of evolutionary biology 32:194–204.
716 Pliske, T. E. 1973. Factors determining mating frequencies in some New World butterflies and skippers.
717 Annals of the Entomological Society of America 66:164–169. Oxford University Press Oxford, UK.
718 Pohlert, T. 2014. The pairwise multiple comparison of mean ranks package (PMCMR). R package 27:10.

719 Pool, J. E., and R. Nielsen. 2007. POPULATION SIZE CHANGES RESHAPE GENOMIC PATTERNS OF
720 DIVERSITY. Evolution 61:3001–3006. John Wiley & Sons, Ltd.
721 Quinlan, A. R., and I. M. Hall. 2010. BEDTools: A flexible suite of utilities for comparing genomic features.
722 Bioinformatics 26:841–842.
723 R Core Team. 2017. R Core Team (2017). R: A language and environment for statistical computing. R
724 Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. R
725 Foundation for Statistical Computing.
726 Rice, W. R. 1984. Sex chromosomes and the evolution of sexual dimorphism. Evolution 38:735–742.
727 Rousselle, M., N. Faivre, M. Ballenghien, N. Galtier, and B. Nabholz. 2016. Hemizygosity enhances
728 purifying selection: lack of fast-Z evolution in two satyrine butterflies. Genome biology and
729 evolution 8:3108–3119.
730 Sackton, T. B., R. B. Corbett-Detig, J. Nagaraju, L. Vaishna, K. P. Arunkumar, and D. L. Hartl. 2014. Positive
731 selection drives faster-Z evolution in silkmoths. Evolution 68:2331–2342.
732 Smith, D. A. S., I. J. Gordon, W. Traut, J. Herren, S. Collins, D. J. Martins, K. Saitoti, and P. Ireri. 2016. A
733 neo-W chromosome in a tropical butterfly links colour pattern, male-killing, and speciation. P.
734 20160821 *in* Proc. R. Soc. B. The Royal Society.
735 Smith, G., Y. R. Chen, G. W. Blissard, and A. D. Briscoe. 2014. Complete dosage compensation and sex-
736 biased gene expression in the moth *Manduca sexta*. Genome Biology and Evolution 6:526–537.
737 Stoletzki, N., and A. Eyre-Walker. 2011. Estimation of the neutrality index. Molecular Biology and
738 Evolution 28:63–70.
739 Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism.
740 Genetics 123:585–595.
741 Toups, M. A., J. B. Pease, and M. W. Hahn. 2011. No Excess Gene Movement Is Detected off the Avian or
742 Lepidopteran Z Chromosome. Genome Biology and Evolution 3:1381–1390.

743 Turner, J. R. G., and P. M. Sheppard. 1975. Absence of Crossover in Female Butterflies (Heliconius).
744 Heredity 34:265–269.
745 Vicoso, B., and B. Charlesworth. 2009. Effective population size and the faster-X effect: an extended
746 model. Evolution: International Journal of Organic Evolution 63:2413–2426.
747 Walters, J. R., and T. J. Hardcastle. 2011. Getting a full dose? reconsidering sex chromosome dosage
748 compensation in the silkworm, Bombyx mori. Genome Biology and Evolution 3:491–504.
749 Wang, J., M. Long, and M. D. Vibranovski. 2012. Retrogenes Moved Out of the Z Chromosome in the
750 Silkworm. Journal of Molecular Evolution 74:113–126.
751 Wang, Z., J. Zhang, W. Yang, N. An, P. Zhang, G. Zhang, and Q. Zhou. 2014. Temporal genomic evolution
752 of bird sex chromosomes. BMC Evolutionary Biology 14:250.
753 Whittle, C. A., A. Kulkarni, and C. G. Extavour. 2020. Absence of a faster-X effect in beetles (Tribolium,
754 Coleoptera). G3: Genes, Genomes, Genetics 10:1125–1136. G3: Genes, Genomes, Genetics.
755 Wilson Sayres, M. A. 2018. Genetic Diversity on the Sex Chromosomes. Genome Biology and Evolution
756 10:1064–1078.
757 Wright, A. E., P. W. Harrison, F. Zimmer, S. H. Montgomery, M. A. Pointer, and J. E. Mank. 2015.
758 Variation in promiscuity and sexual selection drives avian rate of Faster-Z evolution. Molecular
759 ecology 24:1218–1235.
760 Wright, A. E., H. K. Moghadam, and J. E. Mank. 2012. Trade-off between selection for dosage
761 compensation and masculinization on the avian Z chromosome. Genetics 192:1433–1445.
762 Genetics Soc America.
763 Xu, L., S. Y. Wa Sin, P. Grayson, S. V. Edwards, and T. B. Sackton. 2019. Evolutionary Dynamics of Sex
764 Chromosomes of Paleognathous Birds. Genome Biology and Evolution 11:2376–2390.

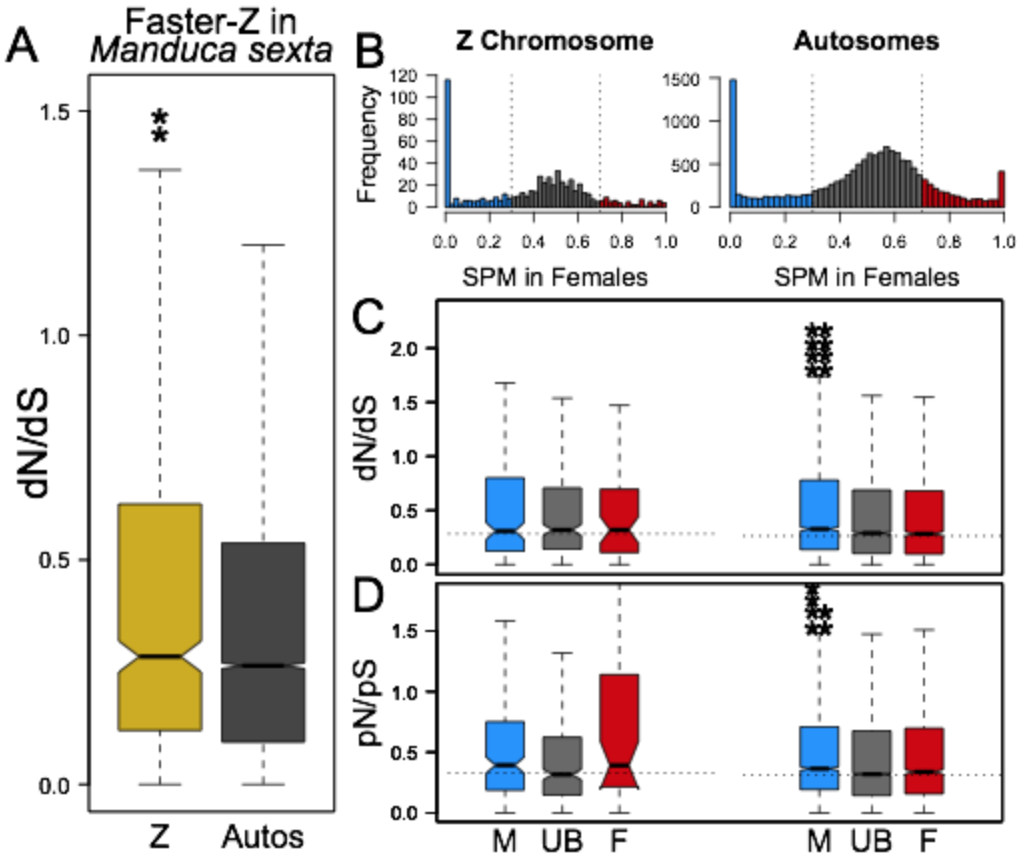
765 Yamamoto, K., J. Nohata, K. Kadono-Okuda, J. Narukawa, M. Sasanuma, S. Sasanuma, H. Minami, M.
766 Shimomura, Y. Suetsugu, and Y. Banno. 2008. A BAC-based integrated linkage map of the
767 silkworm Bombyx mori. Genome biology 9:1–14. BioMed Central.
768 Zhan, S., and S. M. Reppert. 2013. MonarchBase: The monarch butterfly genome database. Nucleic Acids
769 Research 41:D758–D763.
770 Zhan, S., W. Zhang, K. Niitepõld, J. Hsu, J. F. Haeger, M. P. Zalucki, S. Altizer, J. C. de Roode, S. M.
771 Reppert, and M. R. Kronforst. 2014. The genetics of monarch butterfly migration and warning
772 colouration. Nature 514:317–321. Nature Publishing Group.
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791

Table 1. Sex bias of the Z chromosomes in the two species studied with gene counts and proportions in parentheses. 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 X^2 p-values < $1.0*10^{-6}$; note that this significant result holds in *Danaus plexippus* whether the Z is considered as one category or two (i.e. neo and ancestral)). The *Manduca sexta* Z is depleted for female-biased genes, while the monarch (ancestral-)Z is depleted for unbiased genes.

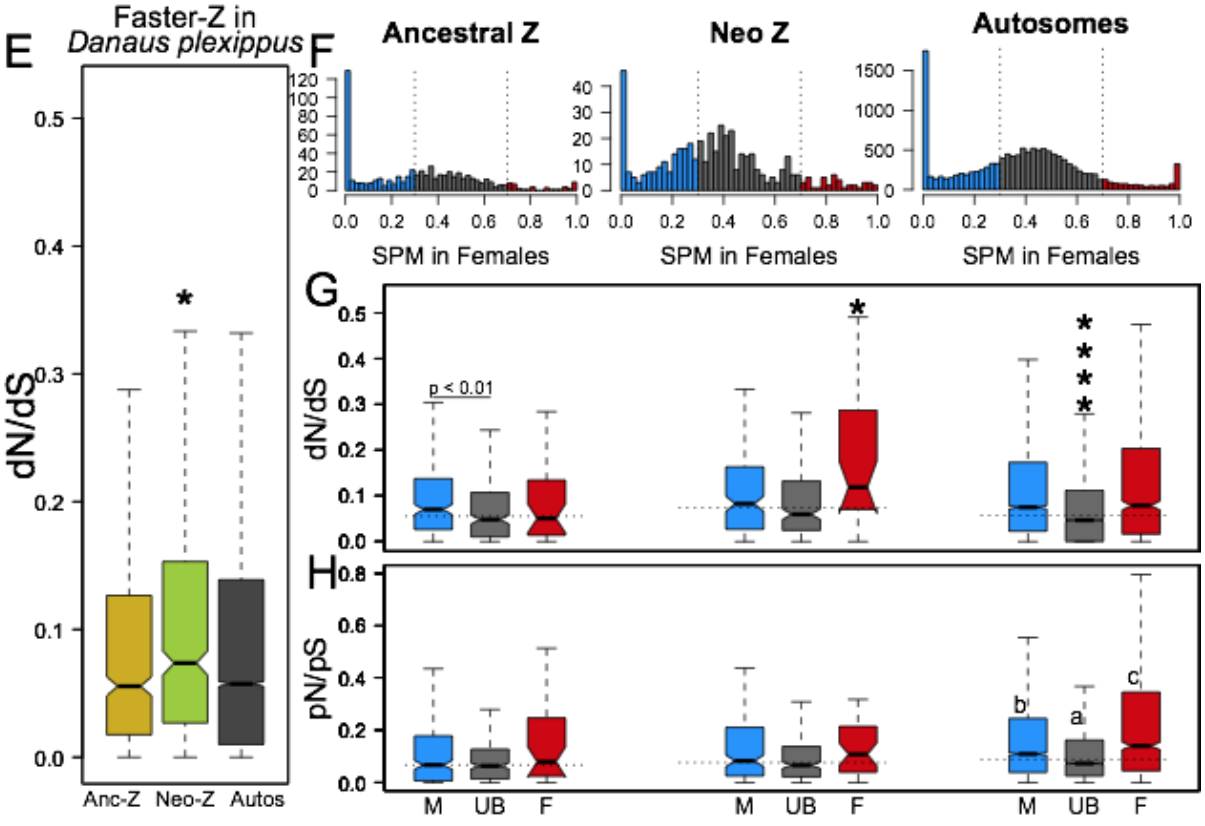
	Carolina sphinx moth (*Manduca sexta*)		Monarch butterfly (*Danaus plexippus*)		
	Autosomes	Z	Autosomes	Ancestral Z	Neo-Z
Male-biased	2477 (0.21)	177 (0.34)	4721 (0.35)	279 (0.47)	184 (0.39)
Unbiased	7219 (0.63)	295 (0.56)	7529 (0.56)	278 (0.46)	243 (0.52)
Female-biased	1856 (0.16)	55 (0.10)	1248 (0.09)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

Table 2. Population genetic parameters across the genomes of both Lepidoptera. Median values are given for divergence and polymorphism estimates (to avoid skew from outliers), while means are reported for Tajima's D (as in every case, the median value is centered on zero). Mean linkage disequilibrium (ρ^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. **<u>Bolded and underlined numbers</u>** 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.

	M. sexta		*D. plexippus*		
	Autosomes	Z	Autosomes	Ancestral Z	Neo-Z
dN	0.0037 (±0.016)	**0.0049** (±0.008)	0.0016 (±0.006)	**0.0032** (±0.006)	**<u>0.0044</u>** (±0.009)
dS	0.0158 (±0.028)	**0.0181** (±0.034)	0.0347 (±0.045)	**0.0667** (±0.048)	**0.0750** (±0.054)
dN/dS	0.2589 (±0.665)	**0.2805** (±0.817)	0.0583 (±0.366)	0.0570 (±0.255)	**0.0757** (±0.269)
pN	**0.0068** (±0.022)	0.0034 (±0.019)	**0.0043** (±0.012)	0.0020 (±0.006)	**0.0037** (±0.007)
pS	**0.0232** (±0.050)	0.0104 (±0.069)	**0.0545** (±0.053)	0.0346 (±0.039)	**0. 0522** (±0.040)
pN/pS	0.3056 (±0.282)	0.3188 (±0.325)	**0.0908** (±0.245)	0.0678 (±0.203)	0.0776 (±0.115)
π_N	**$8.82*10^{-9}$** (±0.030)	$4.68*10^{-9}$ (±0.022)	**<u>$1.83*10^{-5}$</u>** (±0.029)	$7.80*10^{-6}$ (±0.022)	**$1.03*10^{-5}$** (±0.024)
π_S	**$5.96*10^{-8}$** (±0.082)	$2.23*10^{-8}$ (±0.055)	**<u>$1.36*10^{-4}$</u>** (±0.080)	$5.19*10^{-5}$ (±0.061)	**$1.16*10^{-4}$** (±0.080)
Tajima's D_0	-0.0795 (±0.433)	**-0.0419** (±0.319)	**-0.3341** (±0.673)	**<u>-0.2749</u>** (±0.613)	-0.3692 (±0.696)
Tajima's D_4	-0.0259 (±0.510)	**-0.0171** (±0.392)	**-0.2702** (±0.631)	-0.2734 (±0.643)	**-0.2584** (±0.629)
ρ^2	0.3546 (±0.396)	**0.3577** (±0.398)	0.1364 (±0.308)	0.1437 (±0.317)	0.1430 (±0.316)

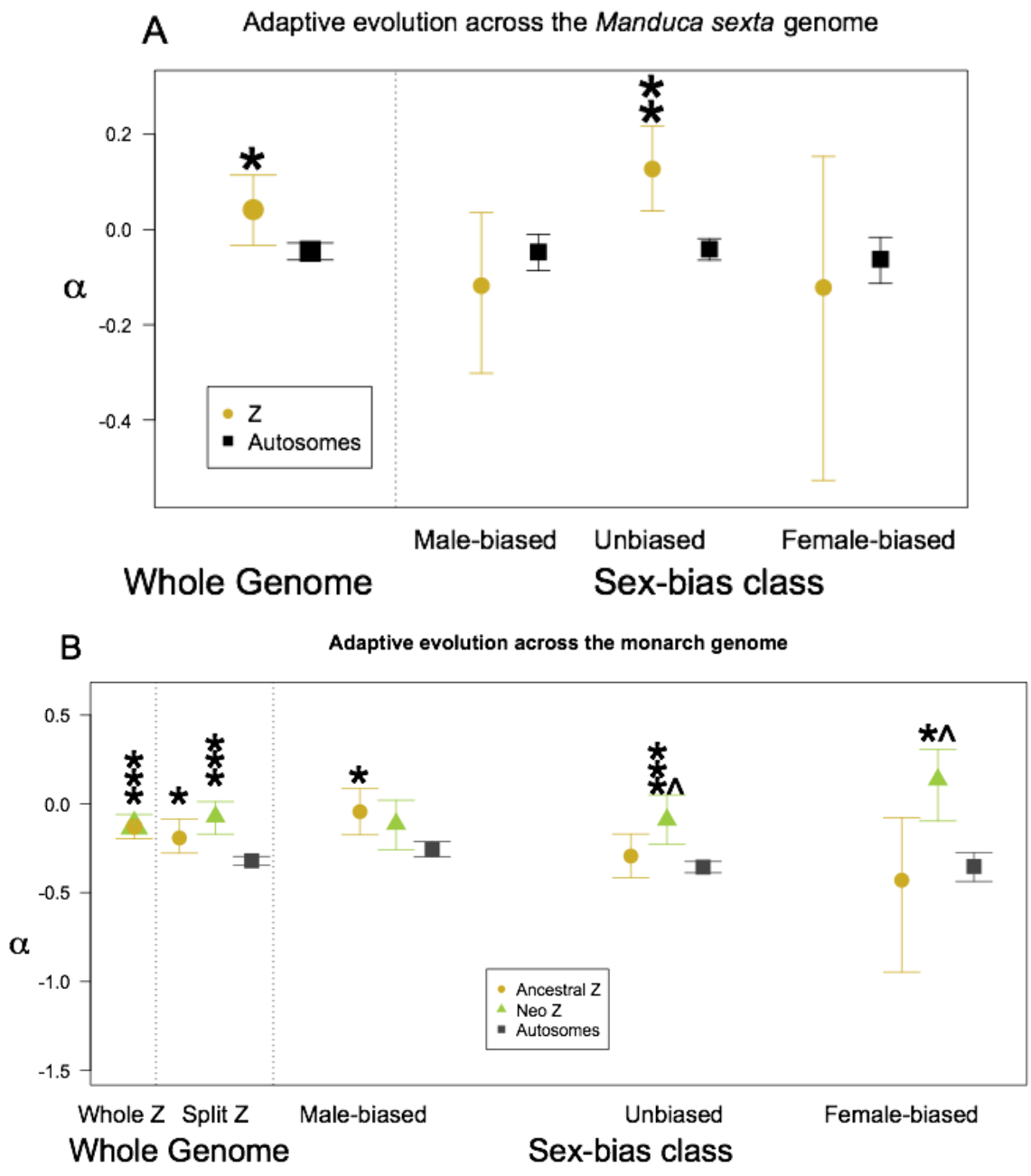

A
Faster-Z in
Manduca sexta
dN/dS
Z
Autos
B
Z Chromosome
Autosomes
Frequency
SPM in Females
SPM in Females
C
dN/dS
D
pN/pS
M UB F
M UB F

836


E
Faster-Z in
Danaus plexippus
dN/dS
Anc-Z Neo-Z Autos
F
Ancestral Z
Neo Z
Autosomes
SPM in Females
SPM in Females
SPM in Females
G
dN/dS
p < 0.01
H
pN/pS
M UB F
M UB F
M UB F

837

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 (* < 0.05, ** <0.01, *** <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 *Manduca 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, while 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).



867

868

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