RESEARCH PAPER

WILEY   JOURNAL OF Evolutionary Biology ⊰eseb⊱

Sexually dimorphic gene expression and transcriptome evolution provide mixed evidence for a fast-Z effect in *Heliconius*

Ana Pinharanda[1,2] (ID) | Marjolaine Rousselle[3] | Simon H. Martin[1] | Joe J. Hanly[1] | John W. Davey[1,4] | Sujai Kumar[5] | Nicolas Galtier[2] | Chris D. Jiggins[1] (ID)

[1]Department of Zoology, University of Cambridge, Cambridge, UK

[2]Department of Ecology and Evolutionary Biology and Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey

[3]UMR 5554 Institut des Sciences de l'Evolution, CNRS, IRD, EPHE, Université de Montpellier, Montpellier, France

[4]Department of Biology, University of York, York, UK

[5]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

Correspondence
Ana Pinharanda, or Professor CD Jiggins, Department of Zoology, University of Cambridge, Cambridge, UK.
Emails: alpp@princeton.edu; c.jiggins@zoo.cam.ac.uk

Abstract

Sex chromosomes have different evolutionary properties compared to autosomes due to their hemizygous nature. In particular, recessive mutations are more readily exposed to selection, which can lead to faster rates of molecular evolution. Here, we report patterns of gene expression and molecular evolution for a group of butterflies. First, we improve the completeness of the *Heliconius melpomene* reference annotation, a neotropical butterfly with a ZW sex determination system. Then, we analyse RNA from male and female whole abdomens and sequence female ovary and gut tissue to identify sex- and tissue-specific gene expression profiles in *H. melpomene*. Using these expression profiles, we compare (a) sequence divergence and polymorphism; (b) the strength of positive and negative selection; and (c) rates of adaptive evolution, for Z and autosomal genes between two species of *Heliconius* butterflies, *H. melpomene* and *H. erato*. We show that the rate of adaptive substitutions is higher for Z than autosomal genes, but contrary to expectation, it is also higher for male-biased than female-biased genes. Additionally, we find no significant increase in the rate of adaptive evolution or purifying selection on genes expressed in ovary tissue, a heterogametic-specific tissue. Our results contribute to a growing body of literature from other ZW systems that also provide mixed evidence for a fast-Z effect where hemizygosity influences the rate of adaptive substitutions.

KEYWORDS
gene expression, genomics, insects, molecular evolution, sex chromosomes

1 | INTRODUCTION

Heteromorphic sex chromosomes have different evolutionary properties compared to autosomes (Rice, 1984). Specifically, because recessive mutations are exposed to selection more readily on the sex chromosomes, positive selection and purifying selection—as well as the strength of genetic drift—are expected to result in different rates of molecular evolution between sex chromosomes and autosomes. An increased evolutionary rate of sex chromosomes relative to autosomes, known as the fast-X effect (Charlesworth, Coyne, & Barton, 1987), has been observed in *Drosophila* (e.g., Avila et al., 2014). X genes are expected to diverge faster between species than autosomal genes mainly due to the higher substitution rate of recessive, advantageous mutations. However, this process is also influenced by

(a) patterns of selection in males versus females; (b) mutation; (c) recombination; and (d) demography (Connallon, Singh, & Clark, 2012; Kirkpatrick & Hall, 2004; Orr, 2010; Orr & Betancourt, 2001; Pool & Nielsen, 2007; Vicoso & Charlesworth, 2006, 2009).

Patterns of molecular evolution on sex chromosomes are particularly influenced by gene expression patterns. Sexually dimorphic expression is often caused by natural and/or sexual selection favouring phenotypes that influence the fitness of one of the sexes (Grath & Parsch, 2016). In species with genetic sex determination, the majority of sexually dimorphic traits results from the differential expression of genes present in both male and female genomes (Ellegren & Parsch, 2007). Sex-biased expression is common across taxa from mammals (Rinn & Snyder, 2005) to Diptera (Assis, Zhou, & Bachtrog, 2012), reptiles (Cox et al., 2017), birds (Mank, Nam, & Ellegren, 2010; Mank, Vicoso, Berlin, & Charlesworth, 2010) and Lepidoptera (Rousselle, Faivre, Ballenghien, Galtier, & Nabholz, 2016). For example, in *Drosophila melanogaster*, 57% of genes have been categorized as sex biased (Assis et al., 2012), and, in *Heliconius melpomene*, analysis of two different tissues identified up to 29% of expressed genes as sex biased (Walters, Hardcastle, & Jiggins, 2015). The vast majority of genes that exhibit sexually dimorphic expression are active in reproductive tissues and tend to also have distinctive rates of molecular evolution compared to genes without dimorphic expression (Avila, Campos, & Charlesworth, 2015; Parisi et al., 2003, 2004). Ultimately, the identification of sex-biased genes, and subsequent analysis of patterns of molecular evolution, will contribute to a better understanding of the evolutionary forces shaping sex chromosome and autosome evolution (Assis et al., 2012; Kirkpatrick & Hall, 2004; Zhang, Hambuch, & Parsch, 2004).

Empirical studies of the fast-X effect typically measure two different metrics: (a) the ratio of nonsynonymous to synonymous substitution rates (dN/dS) and (b) the amount of adaptive evolution (α) using the McDonald–Kreitman (MK) test (McDonald & Kreitman, 1991). Studies measuring dN/dS usually test for "faster-X divergence." Although this approach may be useful for comparing sex chromosome and autosomal divergence, measuring the relative rate of nonsynonymous substitutions captures the effects of both adaptive and neutral (or slightly deleterious) mutations. Estimates of α can better test for an excess of adaptive substitution in the sex chromosome ("faster-X adaptation") by combining measures of within-species polymorphism and between-species divergence, but α is still sensitive to the rate of accumulation of slightly deleterious mutations and demography (Fay, 2011). For instance, an increase in N_e is expected to result in decreased dN/dS and increased α even when the rate of adaptive substitutions remains unchanged. To overcome this problem, extensions of the MK test such as ω_a were developed to estimate the rate of adaptation by calculating the frequency distribution of polymorphism after correcting for demographic history and distribution of deleterious effects at functional sites (Galtier, 2016).

The analysis of evolutionary rates between sex and autosomal genes, however, has produced mixed evidence in support of fast-X evolution (Meisel & Connallon, 2013). In some taxa, there is strong evidence for faster-X divergence but not faster-X adaptation, or vice versa

(Meisel, Malone, & Clark, 2012). For example, the first calculations of faster-X divergence were carried out in *Drosophila* where support for elevated dN/dS in X genes has been mixed. Studies that used autosome-to-X translocations to control for gene content effect did not reach a consensus on the existence of faster-X divergence (Counterman, Ortiz-Barrientos, & Noor, 2004; Thornton, Bachtrog, & Andolfatto, 2006; Zhou & Bachtrog, 2012) but X-linked duplicate genes have elevated dN/dS compared to autosomal duplicates (Thornton & Long, 2002). Signals of faster-X sequence divergence in *Drosophila* have been shown to affect noncoding regulatory regions as well, and might be at least partly explained by differences in gene composition on the X versus the autosomes (Hu, Eisen, Thornton, & Andolfatto, 2013). However, faster-X divergence in other taxa has received stronger support. For example, in humans, chimpanzees and rodents, dN/dS is higher for X genes (Mank, Vicoso et al., 2010; Nielsen et al., 2005).

In contrast, whole-genome analyses of adaptive substitutions have resulted in stronger evidence for faster-X adaptation in *Drosophila* (Mackay et al., 2012), whereas support for faster-X adaptation in vertebrates is less clear. McDonald–Kreitman tests support faster-X adaptation in wild mouse populations (Baines & Harr, 2007) but, for the European rabbit (*Oryctolagus cuniculus*), a clear faster-X adaptation signal is only present in populations with large effective population sizes (Carneiro et al., 2012).

Taxa with ZW sex determination provide an interesting contrast. For female heterogametic taxa such as birds, females only have one copy of the Z chromosome. A fast-Z effect may be expected to result from the expression of recessive mutations on the Z chromosome as Z genes are immediately exposed to selection in females (Charlesworth, 2012). In birds, fast-Z divergence has been reported, but Z male-biased genes were not less accelerated than unbiased genes or female-biased genes (Wright, Zimmer, Harrison, & Mank, 2015). This would not be expected if the fast-Z effect was driven by recessive beneficial mutations, and so, it was suggested that fast-Z in birds does not reflect positive selection (Mank, Nam, et al., 2010; Wright et al., 2015).

For Lepidoptera, results have also been mixed. Sackton et al. (2014) reported that faster-Z evolution was driven by positive selection in silkworms. But, in satyrine butterflies, there were no significant differences in adaptive evolutionary rates between the Z and the autosomes (i.e., no fast-Z adaptation). However, the comparison of male-biased, female-biased and unbiased Z genes in satyrine butterflies revealed increased purifying selection against recessive deleterious mutations in female-biased Z genes (Rousselle et al., 2016). Therefore, considerable uncertainty remains regarding the prevalence and magnitude of the fast-X/Z effect on divergence and adaptation.

Here, we investigate the effects of hemizygosity on the rates of adaptive substitution in the neotropical butterfly genus *Heliconius*, a ZW sex determination system, by analysing polymorphism, divergence and gene expression genomewide. We test whether there is a fast-Z effect in *Heliconius* using two species from the *H. melpomene* and *H. erato* clades which diverged 12 million years ago (synonymous divergence = 0.15) (Kozak et al., 2015; Martin et al., 2016). Previous analyses of *Heliconius* transcriptome data have focused on the evolution of

dosage compensation and the impact of sex-specific dosage on the levels of gene expression (Walters et al., 2015). In this study, using the same transcriptome data, we first compute sex-biased expression. Then, accounting for sex-biased gene expression, we (a) calculate coding sequence divergence and polymorphism in *H. melpomene* and (b) assess the strength of positive and negative selection, and rates of adaptive evolution between *H. melpomene* and *H. erato*. We then analyse newly generated female transcriptome data from *H. melpomene* ovary and gut tissue in order to investigate whether genes expressed in the reproductive tissue of the heterogametic sex have higher rates of adaptive evolution than those expressed in somatic tissues.

2 | MATERIALS AND METHODS

2.1 | Updated *H. melpomene* annotation

The Hmel2 annotation of the *H. melpomene* genome has 13,178 predicted transcripts spanning 16,897,139 bp (Davey et al., 2016 The Heliconius Genome Consortium, 2012). The Hmel2 annotation is incomplete, as there are 20,118 high-quality predicted transcripts in *H. erato* spanning 33,669,374 bp (Van Belleghem et al., 2018). To improve the completeness of the annotation for *H. melpomene*, we downloaded RNA-seq reads from NCBI repositories ArrayExpress ID: E-TAB-1500 (Briscoe et al., 2013) and BioProject PRJNA283415 (Walters et al., 2015), published since Hmel1 release. We also used data from 10 wing RNA-seq libraries (Hanly, 2017). We used the BRAKER1 pipeline to perform unsupervised RNA-seq-based genome annotation (Hoff, Lange, Lomsadze, Borodovsky, & Stanke, 2016). GeneMark-ET was used to perform iterative training, generating initial gene structures, and AUGUSTUS was used for training and subsequent integration of RNA-seq read information into the final gene predictions (Hoff et al., 2016 ; Lomsadze, Burns, & Borodovsky, 2014; Stanke, Diekhans, Baertsch, & Haussler, 2008). This resulted in 26,017 predicted transcripts spanning 32,222,367 bp. 6,532 of these transcripts had a 90% single hit match to the Repbase repeat database. We considered the 6,532 transcripts to be repeat proteins and removed them (Bao, Kojima, & Kohany, 2015). We transferred 428 manually annotated genes (441 transcripts/protein) from the original Hmel2 annotation and removed any BRAKER1 predictions that overlapped. We also transferred 189 genes (189 transcripts/proteins) that have been manually annotated and published since Hmel2 release. Specifically, we transferred 73 gustatory receptors, 31 immune response and 85 glutathione-S-transferases and glucuronosyltransferases (Briscoe et al., 2013; van Schooten, Jiggins, Briscoe, & Papa, 2016; Yu, Fang, Zhang, & Jiggins, 2016) and removed any BRAKER1 predictions that were overlapping. Moreover, BRAKER1 predictions that had 1-to-1 overlaps with Hmel2 names were replaced by their original Hmel2 name. For many-to-1 mapping between the BRAKER1 predictions and Hmel2, Hmel2 names were reused and a suffix of g1/g2/g3/etc. was added. The rest were renamed from HMEL030000 onwards.

2.2 | Samples for gene expression analysis

Gene expression data were calculated using (a) Illumina 100-bp paired-end RNA-seq data from five Panamanian *H. m. rosina* whole-male abdomens, and five Panamanian *H. m. rosina* whole-female abdomens, downloaded from GenBank (BioProject PRJNA283415) (Walters et al., 2015) and (b) newly sequenced Illumina HiSeq 2500 150-bp paired-end directional (stranded) RNA-seq data from ovary tissue of seven young (1 hr) and six old (20 days) *H. m. rosina* females and from gut tissue of six young (1 hr) and six old (20 days) *H. m. rosina* females (25 samples from 13 different individuals, Supporting Information Table S1).

For these 25 samples, *H. m. rosina* females were reared in insectaries in Gamboa, Panama. *Passiflora platyloba* potted plants were monitored daily and 5th-instar caterpillars were removed and taken to the laboratory in large individual containers where they were allowed to pupate and emerge at a constant temperature (24–25°C). The pupating containers in the laboratory were monitored several times a day. Sexually antagonistic pressures are expected to be greater in mature adults, and genes exhibiting extreme sexually dimorphic patterns should be expressed in mature individuals (Gibson, Chippindale, & Rice, 2002). In addition, mating induces behavioural and physiological changes and has been shown to trigger regulatory changes in sex-biased genes (e.g., Immonen, Sayadi, Bayram, & Arnqvist, 2017). By collecting gene expression data from ovaries at two different life-history time points, we aimed to increase the number of sex-specific genes identified. When a female emerged, it was either (a) returned to the insectaries to be mated to a *H. m. rosina* male (Treatment: old, Supporting Information Table S1) or (b) dissected 1 hr after eclosion under controlled laboratory conditions (Treatment: young, Supporting Information Table S1). Mated females were kept in individual 1 m × 1 m × 2 m cages for 20 days until dissection.

Guts and ovaries were dissected in RNAlater (ThermoFisher, Waltham, MA) at 24–25°C, and tissue was stored in RNAlater at 4°C for 24 hr and −20°C thereafter. Total RNA was extracted with a combined guanidium thiocyanate–phenol–chloroform and silica matrix protocol using TRIzol (Invitrogen, Carlsbad, CA), RNeasy columns (Qiagen, Valencia, CA) and DNaseI (Ambion, Naugatuck, CT). mRNA was isolated from total RNA via poly-A pull-down, and directional cDNA library preparation and sequencing (Illumina HiSeq 2500, 150-bp paired end) were performed by Novogene Bioinformatics Technologies (Hong Kong, China) (Supporting Information Table S1).

2.3 | Read mapping, counting and identification of sex- and ovary- and gut-biased genes

FASTQ reads were aligned to gene sequences from *H. melpomene* v2.5 annotation using HISAT2 (Kim, Langmead, & Salzberg, 2015) with default mapping parameters. Mapping statistics were calculated using SAMTOOLS FLAGSTAT (v1.2) (Li et al., 2009). We used htseq-count to determine the number of aligned sequencing reads mapped to each genic feature (HTSEQ v0.6.1; python v2.7.10; option: -m union) (Anders, Pyl, & Huber, 2015).

Estimation of variance–mean dependence from the count data was performed with DESEQ2 (v1.14.1) (Love, Huber, & Anders, 2014) using BIOCONDUCTOR v3.4 and R v3.2.5, using the constructor function DESeqDataSetFromHTSeqCount(design = ~batch + tissue) for ovary- and gut-biased genes. All the result tables were built using the DESeq2 results() function (options: betaPrior = false, test = Wald). We filtered the results as in Walters et al. (2015) with FDR < 0.05 (alpha = 0.05) (Walters et al., 2015). Ovary- and gut-biased genes have a $\log_2$-fold change significance threshold >1.5 (option: lcf-Threshold < 1.5, altHypothesis = "greaterAbs"). We defined male, female and unbiased genes as in Rousselle et al. (2016). First, we calculated reads per kilobase per million (RPKM) as:

$$RPKM = \frac{N_c \times 10^9}{N_{tot} \times L_c}$$

where N_c is the number of reads mapped to the genic feature, N_{tot} is the total number of reads mapped in the sample, and L_c is the length of the genic sequence in base pairs (Mortazavi, Williams, McCue, Schaeffer, & Wold, 2008). $RPKM_i$ is the mean RPKM of gene i across the 10 individuals. Genes for which $RPKM_f/RPKM_m > 1.5$ were classified as female biased, genes for which $RPKM_f/RPKM_m < 0.66$ were classified as male biased, and the others were classified as unbiased (Rousselle et al., 2016).

2.4 | Extraction of orthologous genes, coding sequence alignment and SNP calling

OrthoFinder was used to identify orthologous groups of genes in the *H. melpomene* and the *H. erato* transcriptomes (options: -t 48 -a 6). 1-1 orthologous gene sequences were selected for use in subsequent analysis (Supporting Information Table S2). Using Gff-Ex, a genome feature extraction package (Rastogi & Gupta, 2014), we extracted coding sequences from (a) 10 whole-genome short-read resequenced wild *H. m. rosina* from Panama (Supporting Information Table S3; Van Belleghem et al., 2018) mapped to Hmel2 (Davey et al., 2016) with bwa-mem (Li & Durbin, 2009) and (b) the reference *H. erato* genome (Van Belleghem et al., 2018).

For the 10 whole-genome resequence *H. m. rosina* samples (Van Belleghem et al., 2018), genotypes were called using HaplotypeCaller (GATK v3.4-0-g7e26428) (DePristo et al., 2011), and genotypes were designated as missing if the read depth for a given individual at a given site was <8. Coding sequences for 1-1 orthologous genes were extracted in fasta format from (a) and (b) and aligned using MACSE, accounting for frameshifts and stop codons (Ranwez, Harispe, Delsuc, & Douzery, 2011).

2.5 | Calculation of diversity and selection statistics for 1-1 ortholog alignments between *H. melpomene* and *H. erato*: *Classic approach*

The adaptive substitution rate was estimated by comparing synonymous and nonsynonymous variation in the polymorphism and divergence compartments, as first proposed by McDonald & Kreitman,

1991; see also Bustamante et al., 2005, and Macpherson, Sella, Davis, & Petrov, 2007). We first used the original MK test (referred to as *Classic approach* hereafter) to estimate the rate of adaptive substitution for all genes found to be orthologous between *H. melpomene* and *H. erato*. We calculated (a) synonymous polymorphism (P_s) and (b) nonsynonymous polymorphism (P_n) in *H. melpomene*, as well as (c) synonymous fixed divergence (dS), and (d) nonsynonymous fixed divergence (dN) between *H. melpomene* and *H. erato*. We estimated the rate of adaptive molecular evolution (α) between the two species as:

$$\alpha = 1 - \frac{dS \times P_n}{dN \times P_s}$$

α assumes that nonsynonymous mutations are either adaptive, neutral or strongly deleterious (McDonald & Kreitman, 1991), with $-\infty > \alpha \geq 1$, where α = 0 represents the null hypothesis that nonsynonymous mutations are neutral (dN/dS = P_n/P_S). α > 0 corresponds to dN/dS > P_n/P_S and indicates positive selection, whereas α < 0 corresponds to dN/dS < P_n/P_S and indicates negative selection. These values were calculated using the EggLib C++ function polymorphismBPP (v2.1.11) (De Mita & Siol, 2012) and Bio++ (v2.2.0) (Dutheil & Boussau, 2008) in python (v2.7.5) using scripts adapted from https://github.com/tatumdmortimer (last accessed 09/04/2018) (O'Neill, Mortimer, & Pepperell, 2015).

2.6 | Calculation of diversity and selection statistics for 1-1 ortholog alignments between *H. melpomene* and *H. erato*: *Modelling approach*

The *Classic approach* to the MK test is robust to differences in mutation rates and variation in coalescent histories across genomic locations (McDonald & Kreitman, 1991). Inference of positive selection using the *Classic approach* of the MK test is not robust, however, to the occurrence of slightly deleterious mutations and demographic change. To account for these confounders, we used a *Modelling* approach to estimate the strength of positive and purifying selection in addition to the *Classic approach* described above, using the method of Eyre-Walker and Keightley (2009) as implemented in Galtier (2016) and Rousselle et al. (2016).

The *Modelling approach* uses the frequency distribution of polymorphism to assess the distribution of deleterious mutations at functional sites. This elaborates on the *Classic approach* of the MK test by modelling the distribution of fitness effects (DFE) of deleterious nonsynonymous mutations as a negative Gamma distribution. The model is fitted to the synonymous and nonsynonymous site frequency spectra (SFS) and the expected dN/dS under near-neutrality is inferred. The difference between the observed and expected dN/dS provides an estimate of the proportion of adaptive nonsynonymous substitutions (α). The per-mutation rate of adaptive substitutions is calculated as:

$$\omega_a = \alpha \times \frac{dN}{dS}$$

and the per-mutation rate of nonadaptive substitutions is calculated as:

$$\omega_{na} = (1 - \alpha) \times \frac{dN}{dS}$$

2.7 | Gene expression level and π_n/π_s ratios

To test whether gene expression level and chromosome type have a significant effect on π_n/π_s ratios, we used a multiple regression analysis. We established the linear model:

$$\log(\pi_{nij}) \sim \log(\pi_{nsj}) + chromosome\_type_j + \log(RPKM)$$

using R (v3.2.5). Chromosome type is either autosome or sex chromosome as assigned in Hmel2 reference genome (Davey et al., 2016). 477 genes with no polymorphism were removed from the analysis. We plotted diagnostic plots of residuals versus fitted values.

3 | RESULTS

3.1 | Hmel2.5 annotation and 1-1 ortholog prediction with *H. erato*

There are 20,118 transcripts predicted in *H. erato* (Van Belleghem et al., 2018) and 20,097 genes (21,565 transcripts/proteins) in *H. melpomene* Hmel2.5. OrthoFinder returned 11,062 clusters of genes, 8085 of which included exactly one sequence per species. 14,841 (73.8%) of the total number of genes in *H. erato* were assigned to an orthogroup, and 14,857 (68.6%) of the total number of genes in *H. melpomene* version Hmel2.5 were assigned to an orthogroup (Supporting Information Table S2). Conversely, the *H. melpomene* Hmel2 annotation has 13,019 predicted gene models (Davey et al., 2016). Using the Hmel2 annotation, OrthoFinder returned 9,320 clusters of genes, 6,846 of which included exactly one sequence per species (i.e., single-copy orthogroups). 13,744 (68.3%) genes were assigned to an orthogroup in *H. erato* and 10,530 (80.9%) were assigned to an orthogroup in *H. melpomene*. The Hmel2.5 annotation set for *H. melpomene* is therefore more comparable to the published *H. erato* gene annotation and is more appropriate for future gene-based analysis in *H. melpomene*. The Hmel2.5 annotation has 1,093 genes mapping to the Z chromosome and 18,835 mapping to the autosomes. This new version of *H. melpomene* genome annotation was numbered Hmel2.5 (available at LepBase http://ensembl.lepbase.org/Heliconius_melpomene_melpomene_hmel25/Info/Index, last accessed 20 June 2018, Challis et al., BioRxiv preprint).

3.2 | RNA-sequencing and read mapping

Analysis of gene expression profiles in the data retrieved from Walters et al. (2015) by principal component, the first principal component separates gene expression in whole abdomen by sex and explains 97% of variance (Supporting Information Figure S1). The 25 *H. melpomene* samples sequenced for this project have a median total number of reads of 34.86 M (min. 27.81 M; max. 46.12 M), similar to previously published gene expression studies in *Heliconius* (Briscoe et al., 2013; Walters et al., 2015. Mapping success is high compared to other published studies (e.g., Walters et al., 2015; Yu et al., 2016) (Supporting Information Table S1). We analysed data from two different time points and from non-sex- and sex-specific tissue separately (Treatment: Young and Old). There is a clear separation of the 25 samples by tissue when we compare gene expression profiles between them. In total, 51% of the total variance is explained by the two-first principal components. PC1 separates the samples by tissue and explains 40% of variance. PC2 explains 11% of total variance and separates samples by age (Supporting Information Figure S3). Ovarian tissue clusters by age more tightly than non-sex-specific tissue (Gut) (Supporting Information Figure S4A,B).

In order to test whether there was a fast-Z effect in *Heliconius*, we calculated the following statistics for autosomal and Z genes: dN/dS, α, ω_a and ω_{na}. We also calculated π_{sZ}/π_{sA} and investigated the relationship between π_n and gene expression. dN/dS tests for "faster-Z divergence." If there is a faster-Z effect in a female heterogametic system, genes with female-biased expression will have higher dN/dS than male-biased or unbiased genes. α, ω_a and ω_{na} test for "faster-Z adaptation." In a female heterogametic system, faster-Z adaptation would predict that the proportion of adaptive nonsynonymous substitutions (α) in female-biased genes will be significantly higher in the Z chromosome compared to the autosomes. ω_a measures the per-mutation rate of adaptive substitutions and, if there is "fast-Z adaptation," ω_a is also predicted to be significantly higher for female-biased Z genes. However, in a female heterogametic system with reduced purifying selection on the Z, the per-mutation rate of nonadaptive substitution (ω_{na}) would also be higher for female-biased Z genes. Another metric that can be used as an indicator of the strength of purifying selection is the π_{sZ}/π_{sA} ratio. In a female heterogametic system, stronger purifying selection on the Z chromosome would lead to a π_{sZ}/π_{sA} ratio <0.75 due to background selection. Overall, patterns of diversity tend to be associated with gene expression levels. π_n is expected to be negatively correlated with expression levels if there is increased purifying selection on highly expressed genes.

3.3 | Coding sequence divergence does not support a significant fast-Z effect

We first compared rates of Z and autosomal sequence divergence using dN/dS comparisons of 1-1 orthologous genes between *H. melpomene* and *H. erato*. The dN/dS ratio for the Z chromosome genes is not significantly higher than dNdS for autosomal genes (dN/dS$_{Auto}$ = 0.110; 95% CI = [0.106–0.113]; dN/dS$_Z$ = 0.120; 95% CI = [0.098–0.145]), indicating no obvious faster-Z divergence of coding sequence.

More highly expressed genes are more exposed to selection, so in a female heterogametic system with a fast-Z effect, genes with female-biased expression are expected to have higher rates of amino acid substitution if dN/dS is driven by positive selection. However, the dN/dS ratio of Z female-biased genes (dN/dS$_Z$ = 0.120; 95% CI = [0.069–0.183]) was not significantly different to that for male-biased genes (dN/dS$_Z$ = 0.148 [0.122–0.172]) or unbiased genes (dNdS$_Z$ = 0.107; 95% CI = [0.078–0.143]). By contrast, among autosomal genes, those that are unbiased have a significantly lower coding sequence divergence compared to both male-biased and female-biased autosomal genes (dNdS$_{Auto}$ = 0.0978; 95% CI = [0.093–0.102]) (Table 1).

Finally, dS on the Z chromosome (dS$_Z$ = 0.189; 95% CI = [0.18–0.2]) is higher than dS on the autosomes (dS$_{Auto}$ = 0.162; 95% CI = [0.16–0.17]), consistent with either a: (a) male-biased mutation rate or (b) difference in coalescence time for autosomes and Z, but does not support a fast-Z effect (Table 1).

3.4 | π_{sZ}/π_{sA} diversity ratio is lower than 0.75

Next, we explored patterns of within-species diversity as an indicator of the strength of purifying selection. In a population at equilibrium with a 1:1 sex ratio, the π_{sZ}/π_{sA} diversity ratio is expected to be 0.75, but stronger purifying selection on the Z chromosome would lead to a reduction in this ratio due to background selection. The π_{sZ}/π_{sA} ratio for *H. melpomene* is approximately 0.44 (Table 2), which might indicate purifying selection on the Z. However, this ratio can also be influenced by a biased sex ratio (Vicoso & Charlesworth, 2006), differences in recombination rates (Charlesworth, 2012), sex-biased

mutation rates (Vicoso & Charlesworth, 2009) or a historical reduction in population size. Recent calculations for *H. melpomene* from Panama using whole-genome short-read sequencing data estimated the π_{sZ}/π_{sA} diversity ratio value to be 0.611; CI = [0.570–0.653] with only weak evidence for a population bottleneck (Van Belleghem et al., 2018). The more pronounced reduction in diversity at synonymous sites observed might, therefore, indicate enhanced background selection in genic regions of the Z chromosome.

3.5 | Increased strength of purifying selection on highly expressed genes

Patterns of diversity were, however, strongly associated with expression levels. Using a multiple regression approach, we found that functional genetic diversity, π_n, was significantly negatively correlated with expression level for both autosomal and Z genes ($p < 0.01$) consistent with increased purifying selection on highly expressed genes (Supporting Information Figure S2) (Figure 1).

3.6 | Z and autosomal rates of adaptive substitution: testing fast-Z adaptation

We next explored patterns of adaptive evolution using (a) the *Classic* MK test and (b) the *Modelling* approach which accounts for the effect of mildly deleterious mutations. We computed (a) the proportion of adaptive nonsynonymous substitutions (α) for both the *Classic* and the *Modelling* approaches and (b) ω_a and ω_{na} for the *Modelling approach*. ω_a is the per-mutation rate of adaptive substitutions and ω_{na} is the per-mutation rate of nonadaptive substitutions.

TABLE 1 Ratios of π_n/π_s, dN/dS in *H. melpomene*; and calculations of α, ω_a and ω_{na} for autosomal and Z male-biased, female-biased and unbiased genes between *H. melpomene* and *H. erato*

	Linkage	All	Female biased	Male biased	Unbiased
π_n/π_s	Autosomal	0.103 [0.100–0.107]	0.106 [0.10–0.113]	**0.127**[A] [0.118–0.138]	**0.094**[A] [0.091–0.098]
	Z	0.111 [0.098–0.126]	0.094 [0.059–0.136]	0.136 [0.112–0.162]	0.104 [0.09–0.125]
dN/dS	Autosomal	0.110 [0.106; 0.113]	0.113 [0.105; 0.121]	0.113 [0.145; 0.167]	**0.098**[A] [0.093; 0.102]
	Z	0.120 [0.098–0.145]	0.120 [0.069–0.183]	0.148 [0.122–0.172]	0.107 [0.078–0.143]
α Classic	Autosomal	0.24 [−0.486 to 0.697]	0.12 [−0.54 to 0.687]	0.305 [−0.309 to 0.669]	0.225 [−0.544 to 0.719]
	Z	0.434 [−0.526 to 0.866]	0.463 [−0.276 to 0.837]	0.279 [−0.728 to 0.794]	0.535 [−0.475 to 1.0]
α Modelling	Autosomal	0.629 [0.622–0.636]	0.635 [0.620–0.650]	0.630 [0.616–0.646]	**0.538**[B] [0.529–0.547]
	Z	**0.675**[A] [0.647–0.704]	0.699 [0.595–0.811]	0.646 [0.596–0.697]	**0.537**[B] [0.500–0.576]
ω_a	Autosomal	0.062 [0.061–0.063]	0.066 [0.065–0.068]	**0.087**[B] [0.085–0.089]	**0.047**[B] [0.046–0.048]
	Z	**0.069**[A] [0.066–0.072]	0.069 [0.058–0.080]	**0.090**[B] [0.083–0.097]	**0.048**[B] [0.044–0.051]
ω_{na}	Autosomal	0.036 [0.036–0.037]	0.038 [0.037–0.040]	**0.051**[A] [0.049–0.053]	0.040 [0.039–0.041]
	Z	0.033 [0.030–0.036]	0.029 [0.019–0.040]	0.049 [0.042–0.056]	0.041 [0.038–0.044]
#Genes	Autosomal	7464	1231	1238	4739
	Z	200	28	96	193

Notes. π_n/π_s, dN/dS ratios, α, ω_a and ω_{na} were calculated for autosomal and Z male-biased, female-biased and unbiased genes. π_n/π_s, dN/dS ratios were calculated for *H. melpomene* samples, and α, ω_a and ω_{na} were calculated between *H. melpomene* and *H. erato*. Intervals represent 95% confidence intervals obtained by bootstrapping 1,000 times. Bold[A] denotes significant values within either Z or autosomal categories. Bold[B] denotes significant values within both Z and autosomal categories. Significance indicated separately for *All* and for sex-biased expression (*female, male* and *unbiased*).

TABLE 2 *H. melpomene* π_s and π_{sZ}/π_{sA} ratio from pairwise alignments for Z and autosomal genes

	Linkage	All	Female biased	Male biased	Unbiased
π_s	Autosomal	**0.027**[(A)] [0.026–0.027]	0.025 [0.024–0.027]	**0.035**[(A)] [0.033–0.036]	0.025 [0.024–0.026]
	Z	0.012 [0.011–0.013]	0.016 [0.013–0.020]	0.015 [0.01–0.02]	0.0106 [0.009–0.012]
π_{sZ}/π_{sA}	NA	0.444	NA	NA	NA

Notes. π_s calculated from pairwise alignments for Z and autosomal genes. π_{sZ}/π_{sA} ratio used to estimate N_{eZ}/N_{eA}. Intervals represent 95% confidence intervals obtained by bootstrapping genes (1,000 replicates). Bold [(A)] denotes significant values within either Z or autosomal categories. Significance indicated separately for *All* and for sex biased (*female, male* and *unbiased*).

There are no significant differences in α values between Z and autosomal genes under the *Classic approach* (Table 1). However, using the *Modelling approach*, when all genes are considered, Z genes have a marginally but significantly higher α (α_Z = 0.675; 95% CI = [0.647–0.704]) than those that are autosomal (α_{Auto} = 0.629; 95% CI = [0.622–0.636]). Nonetheless, α is not significantly different between the Z chromosome and autosomes for female-biased (α_{Auto} = 0.635; 95% CI = [0.62–0.65]; α_Z = 0.699; 95% CI = [0.595–0.811]) or male-biased genes (α_{Auto} = 0.63; 95% CI = [0.616–0.646]; α_Z = 0.646; 95% CI = [0.596–0.697]). Unbiased genes have significantly lower α values than female- or male-biased genes for both Z and autosomes, but within the unbiased genes there is no significant difference in α between Z (α_Z = 0.537; 95% CI = [0.5–0.567]) and autosomes (α_{Auto} = 0.538; 95% CI = [0.529–0.547]) (Table 1). This means that the significant effect we find in the Z (α_Z = 0.675; 95% CI = [0.647–0.704]) when analysing all the genes together is due to the sex-biased genes and not due to the unbiased gene category. This observation is in accordance with the predictions from Rice (1984), where the accumulation of alleles under sex-specific and antagonistic selection on the Z was expected if the alleles were recessive in females and dominant in males (Rice, 1984). The lack of significant differences in α between sex-biased genes is not consistent with the expectations of fast-Z adaptation, which would predict faster evolution of female-biased genes due to hemizygosity compared to autosomes, but this could also reflect a lack of power to detect the signal when the total number of genes is reduced.
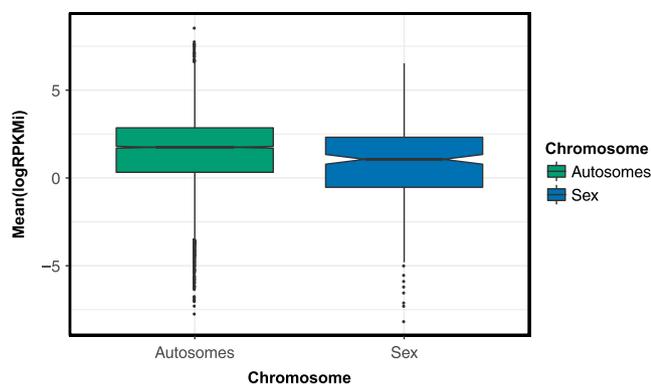
FIGURE 1 Expression level of Z and autosomal genes. Median expression level of Z genes is significantly lower than autosomal genes ($p < 0.05$). Notches on boxplot display the confidence intervals around the median

3.7 | Hemizygosity and the rate of adaptive substitutions

There was no evidence for reduced purifying selection on the Z chromosome, as the per-mutation rate of nonadaptive substitution (ω_{na}) is lower for Z genes (ω_{naZ} = 0.033; 95% CI = [0.030–0.036] and $\omega_{na\,Auto}$ = 0.036; 95% CI = [0.036–0.037]). Female-biased genes have the lowest ω_{na} ($\omega_{na\,Auto}$ = 0.038; 95% CI = [0.037–0.040]; $\omega_{na\,Z}$ = 0.029; 95% CI = [0.019–0.040]) compared to male-biased ($\omega_{na\,Auto}$ = 0.051; 95% CI = [0.049–0.053]; $\omega_{na\,Z}$ = 0.049; 95% CI = [0.042–0.056]) and unbiased ($\omega_{na\,Auto}$ = 0.04; 95% CI = [0.039–0.041]; $\omega_{na\,Z}$ = 0.041; 95% CI = [0.038–0.044]) genes, which confirms the low π_n/π_s already reported and would suggest that purifying selection is stronger in female-biased genes.

3.8 | Female ovary-biased and gut-biased genes

Next, we explored the expression of genes in female reproductive tissue. Overall there were a greater number of genes with gut-biased expression (#Gut$_{Auto}$ = 153) than ovary-biased expression (#Ovary$_{Auto}$ = 40) in the autosomes. However, there was an over-representation of Z ovary-expressed genes than expected by chance (#Gut$_Z$ = 6; #Ovary$_Z$ = 6; chi-square test; $p < 0.05$). However, the number of genes in each category is relatively small so these tests should be treated with caution.

Of the 205 differentially expressed genes between the two tissues, only 9 in the ovaries and 26 in the gut could be used to calculate dN/dS, π_n/π_s and α. The other genes either do not have a 1-1 ortholog with *H. erato* or there were too many undetermined characters (gaps or Ns) to be able to estimate the parameters. Of the 35 genes for which molecular evolution statistics could be calculated, all 9 ovary-biased and 25 of 26 gut-biased genes are autosomal; and 1 gut-biased gene maps to the Z (Gut dN/dS $_Z$ = 0.365; Gut $\pi_n/\pi_{s\,Z}$ = 0.093; Gut α_Z = 0.295). We did not detect any significant differences in π_n/π_s; dN/dS; or α for autosomal ovary- and gut-biased genes.

4 | DISCUSSION

Elevated rates of coding sequence evolution on the sex chromosome relative to autosomes have been reported for several species, consistent with the theoretical prediction of fast-X evolution. Here, we find evidence for enhanced rates of adaptation on the *Heliconius* Z chromosome: Z genes have a significantly higher rate of adaptive

evolution when all expressed genes are considered. However, fast-X theory predicts that genes highly expressed in the hemizygous sex should be especially prone to fast-X evolution, and this prediction was not satisfied in our data. Female-biased genes did not evolve faster when located on the Z chromosome. The evidence for fast-Z evolution in *Heliconius* is, therefore, mixed.

In other taxa, there is strongest support for fast-X evolution in groups with complete dosage compensation (Mank, Vicoso, et al., 2010; Meisel & Connallon, 2013). Theory predicts that opportunities for fast-X evolution should increase in species with somatic X-inactivation such as eutherian mammals, as there is effectively haploid expression of the sex chromosome in cells, increasing the chances of recessive beneficial mutations being fixed (Charlesworth, 2012). Groups such as Lepidoptera have been reported to have more complex patterns of sex chromosome dosage compensation. In *Heliconius* males, expression of Z genes is reduced below autosomal levels, but this dosage compensation mechanism is imperfect, with males showing increased expression relative to females on Z chromosome genes (Walters et al., 2015). However, the apparent incomplete dosage compensation could be a consequence of an uneven distribution of sex-biased genes on sex chromosomes (Gu & Walters, 2017, Huylmans, Macon, & Vicoso, 2017). Regardless, when we compare rates of divergence and adaptation for genes with sex-biased expression, the expectations of fast-Z evolution are not clearly met.

Although we might expect faster rates of adaptive evolution for female-biased genes, we observe a weak tendency for faster rates of evolution in male-biased genes. This might mean that the Z could in fact be a hotspot for dominant alleles that benefit males. If this is the case, faster rates of evolution in male Z genes could contribute to a fast-Z effect that is unrelated to hemizygous nature of the Z in females (Rice, 1984). It is important to note that sex-specific genes can evolve rapidly both due to (a) neutral processes (as they experience relaxed selection in the other sex) and (b) the accumulation moderately deleterious mutations (Dapper & Wade, 2016; Gershoni & Pietrokovski, 2014, 2017; Mank, 2017). Nevertheless, our results of ω_{α} for male-biased Z-linked genes seem to support faster adaptive evolution of such genes: as the Z spends less time in females as compared to autosomes, relaxed selection should be greater for autosomal male-biased genes (Rice, 1984).

Although a fast-Z effect has been observed in *Bombyx mori* (Sackton et al., 2014), no such pattern was reported in two satyrine butterflies where the dN/dS ratio of Z genes was slightly lower than for autosomal genes (Rousselle et al., 2016). In *Heliconius*, although dN/dS was not significantly different between autosomal and Z genes, we did find evidence for a faster rate of adaptive substitution. Interestingly, our data also show that dS on the Z chromosome is higher than dS on the autosomes perhaps indicating a male-biased mutation rate, as Z chromosomes spend more time in males than in females (Miyata, Hayashida, Kuma, Mitsuyasu, & Yasunaga, 1987). Although hemizygosity is expected to expose beneficial mutations to selection and increase rates of adaptive evolution on the Z chromosome, it is also expected to increase the efficacy of purifying selection, which would act to reduce evolutionary rates. It may be that

the balance between these two forces differs across lepidopteran species, leading to the mixed pattern of fast-Z evolution in some taxa but not others. It is important to add, however, that the α values estimated in this study are substantially higher than α reported in Martin et al. (2016). Martin et al. (2016) estimated α using the approach developed by Messer and Petrov (2013), and, in simulations, it has been shown that it is possible that there is an overestimation of DFE-α (the method used in this study) in scenarios with strong sweeps or population expansion (Messer & Petrov, 2013).

Wright et al. (2015) interpreted the high dN/dS in Z genes of birds as a consequence of reduced effective population size rather than positive selection. The difference in effective population size between sex chromosomes and autosomes in female heterogametic systems is predicted to be larger than in male heterogametic systems due to higher variance of male reproductive success (Mank, Nam et al., 2010). Indeed, we estimate that coding regions on the Z chromosome have an N_e 0.44 times that of autosomes. We might therefore predict a considerable reduction in the efficacy of purifying selection on butterfly Z chromosomes. This should lead to higher ω_{na} and π_n/π_s ratios on the Z due to stronger genetic drift. However, as in satyrine butterflies (Rousselle et al., 2016), in *Heliconius*, ω_{na} is not higher on the Z relative to autosomes. In *Heliconius*, dN/dS and π_n/π_s are higher in the Z compared to autosomes, but this is not significant. This means that, in contrast to birds, the difference in the effective population size of the Z relative to autosomes is not sufficient to reduce the efficacy of purifying selection at a detectable level.

One possible explanation for this difference is the generally much higher effective population sizes of Lepidoptera, which could allow for efficient selection even on sex chromosome (Rousselle et al., 2016). Another is that by not using all genomic sites to estimate the π_{sZ}/π_{sA} diversity ratio, we might be underestimating its true value due to a stronger effect of background selection. The latter is supported by the observation that, in a recently published paper using all genomic sites to estimate π_{sZ}/π_{sA}, Van Belleghem et al. (2018) calculated it to be 0.661 CI = [0.570–0.653]. Regardless, both ours and Van Belleghem et al. (2018) estimates of the π_{sZ}/π_{sA} diversity ratio are significantly lower than the expected 0.75, and there is no observable reduction in the efficacy of purifying selection in *H. melpomene*.

Another factor that might counteract the fast-Z effect is adaptation from standing variation. Larger populations are more polymorphic and, therefore, have an increased probability of adaption from standing genetic variation. Adaptation from standing genetic variation is expected to result in faster autosome evolution, independent of the dominance of beneficial alleles (Orr & Betancourt, 2001), which would counteract the fast-Z effect. This may be especially relevant when overall population sizes are large, as in many *Heliconius* species, such that standing variation becomes a comparatively important source of adaptive variation compared to de novo mutations.

As sex genes tend to be expressed in sex-specific tissue such as the testis and the ovaries, we aimed to investigate patterns of molecular

evolution in ovary-biased genes. Unfortunately, there are no ovary-biased genes with 1-1 orthologs between *H. melpomene* and *H. erato* that are in the Z. This means we could not test the effect of hemizygosity on non-sex-specific and sex-specific female expression directly. The lack of 1-1 orthology may mean that these genes are rapidly evolving. On the one hand, adaptive evolution rates could be higher. And, indeed, autosomal ovary-expressed genes have higher rates of adaptive evolution than gut expressed genes. On the other hand, the lack of 1-1 orthology could also be due to rapid nonadaptive evolution (Gershoni & Pietrokovski, 2014).

It has been recently shown that, like in other taxa, sex-biased genes can experience rapid turnovers in butterflies (e.g., *Papilio*) (Huylmans et al., 2017). Consequently, it is possible for *H. melpomene* and *H. erato* to have different sex-biased genes. However, if there are differences in sex-biased genes between *H. melpomene* and *H. erato*, these are likely to only occur in small number of genes. *H. melpomene* and *H. erato* diverged 12MYA (Kozak et al., 2015) compared to 35MYA for *P. xuthus* and *P. machaon* (Zakharov, Caterino, & Sperling, 2004). Regardless, any gene that is sex biased in *H. erato* but not in *H. melpomene* is categorized as unbiased in our data. So, for example, if a fast-evolving Z gene is female biased in *H. erato*, but unbiased in *H. melpomene*, the divergence and adaptation estimates could be inflated for the *Unbiased* category. In the future, analysing gene expression data from *H. erato* to understand whether sex-biased genes are different for the two species would contribute to understand the actual gene turnover.

Together these results illustrate the need to study substitution rates in other ZW systems considering sex-biased expression. This genomewide analysis of polymorphism, divergence and gene expression data contributes to a growing body of literature on sex chromosome evolution in ZW systems, and reveals the complexity of the different evolutionary forces shaping transcriptome evolution in *Heliconius* and, consistent with previous work, shows limited evidence of fast-Z evolution in this taxon.

ORCID

Ana Pinharanda (iD) https://orcid.org/0000-0002-4696-9218

Chris D. Jiggins (iD) https://orcid.org/0000-0002-7809-062X

REFERENCES

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*, 166–169. https://doi.org/10.1093/bioinformatics/btu638

Assis, R., Zhou, Q., & Bachtrog, D. (2012). Sex biased transcriptome evolution in Drosophila. *Genome Biology and Evolution*, *4*, 1189–1200. https://doi.org/10.1093/gbe/evs093

Avila, V., Campos, J. L., & Charlesworth, B. (2015). The effects of sex biased gene expression and X-linkage on rates of adaptive protein sequence evolution in Drosophila. *Biology Letters*, *11*, 20150117. https://doi.org/10.1098/rsbl.2015.0117

Avila, V., Procé, S. M., Campos, J. L., Borthwick, H., Charlesworth, B., & Betancourt, A. J. (2014). Faster-X effects in two drosophila lineages. *Genome Biology and Evolution*, *6*(10), 2968–2982. https://doi.org/10.1093/gbe/evu229

Baines, J. F., & Harr, B. (2007). Reduced X-linked diversity in derived populations of house mice. *Genetics*, *175*, 1911–1921. https://doi.org/10.1534/genetics.106.069419

Bao, W., Kojima, K. K., & Kohany, O. (2015). Repbase update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA*, *6*, 11. https://doi.org/10.1186/s13100-015-0041-9

Briscoe, A. D., Macias-Muñoz, A., Kozak, K. M., Walters, J. R., Yuan, F., Jamie, G. A., … Jiggins, C. D. (2013). Female behaviour drives expression and evolution of gustatory receptors in butterflies. *PLoS Genetics*, *9*, e1003620. https://doi.org/10.1371/journal.pgen.1003620

Bustamante, C. D., Fledel-Alon, A., Williamson, S., Nielsen, R., Hubisz, M. T., Glanowski, S., … Clark, A. G. (2005). Natural selection on protein-coding genes in the human genome. *Nature*, *437*(7062), 1153–1157. https://doi.org/10.1038/nature04240

Carneiro, M., Albert, F. W., Melo-Ferreira, J., Galtier, N., Gayral, P., Blanco-Aguiar, J. A., … Ferrand, N. (2012). Evidence for widespread positive and purifying selection across the European rabbit (*Oryctolagus cuniculus*) genome. *Molecular Biology and Evolution*, *29*, 1837–1849. https://doi.org/10.1093/molbev/mss025

Charlesworth, B. (2012). The role of background selection in shaping patterns of molecular evolution and variation: Evidence from variability on the Drosophila X chromosome. *Genetics*, *191*(1), 233–246. https://doi.org/10.1534/genetics.111.138073

Charlesworth, B., Coyne, J. A., & Barton, N. H. (1987). The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist*, *130*(1), 113–146.

Connallon, T., Singh, N. D., & Clark, A. G. (2012). Impact of genetic architecture on the relative rates of X versus autosomal adaptive

substitution. *Molecular Biology and Evolution*, 29, 1933–1942. https://doi.org/10.1093/molbev/mss057

Counterman, B. A., Ortiz-Barrientos, D., & Noor, M. A. F. (2004). Using comparative genomic data to test for fast-X evolution. *Evolution*, 58(3), 656–660.

Cox, R. M., Cox, C. L., McGlothlin, J. W., Card, D. C., Andrew, A. L., & Castoe, T. A. (2017). Hormonally mediated increases in sex biased gene expression accompany the breakdown of between-sex genetic correlations in a sexually dimorphic lizard. *American Naturalist*, 189, 315–332. https://doi.org/10.1086/690105

Counterman, B. A., Ortiz-Barrientos, D., & Noor, M. A. F. (2004). Using comparative genomic data to test for fast-X evolution. *Evolution*, 58(3), 656–660.

Dapper, A., & Wade, M. (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. https://doi.org/10.1111/evo.12848

Davey, J. W., Chouteau, M., Barker, S. L., Maroja, L., Baxter, S. W., Simpson, F., … Jiggins, C. D. (2016). Major improvements to the *Heliconius melpomene* genome assembly used to confirm 10 chromosome fusion events in 6 million years of butterfly evolution. *G3: Genes, Genomes, Genetics*, 6, 695–708. https://doi.org/10.1534/g3.115.023655

De Mita, S., & Siol, M. (2012). EggLib: Processing, analysis and simulation tools for population genetics and genomics. *BMC Genetics*, 13, 27. https://doi.org/10.1186/1471-2156-13-27

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., … Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43, 491–498. https://doi.org/10.1038/ng.806

Dutheil, J., & Boussau, B. (2008). Non-homogeneous models of sequence evolution in the Bio++ suite of libraries and programs. *BMC Evolutionary Biology*, 8, 255. https://doi.org/10.1186/1471-2148-8-255

Ellegren, H., & Parsch, J. (2007). The evolution of sex biased genes and sex biased gene expression. *Nature Reviews Genetics*, 8, 689–698. https://doi.org/10.1038/nrg2167

Eyre-Walker, A., & Keightley, P. D. (2009). Estimating the rate of adaptive molecular evolution in the presence of slightly deleterious mutations and population size change. *Molecular Biology and Evolution*, 26(9), 2097–2108. https://doi.org/10.1093/molbev/msp119

Fay, J. C. (2011). Weighing the evidence for adaptation at the molecular level. *Trends in Genetics*, 27(9), 343–349. https://doi.org/10.1016/j.tig.2011.06.003

Galtier, N. (2016). Adaptive protein evolution in animals and the effective population size hypothesis. *PLoS Genetics*, 12, e1005774. https://doi.org/10.1371/journal.pgen.1005774

Gershoni, M., & Pietrokovski, S. (2014). Reduced selection and accumulation of deleterious mutations in genes exclusively expressed in men. *Nature Communications*, 5, 4438. https://doi.org/10.1038/ncomms5438

Gershoni, M., & Pietrokovski, S. (2017). The landscape of sex-differential transcriptome and its consequent selection in human adults. *BMC Biology*, 15, 7. https://doi.org/10.1186/s12915-017-0352-z

Gibson, J. R., Chippindale, A. K., & Rice, W. R. (2002). The X chromosome is a hot spot for sexually antagonistic fitness variation. *Proceedings of the Royal Society of London Series B: Biological Sciences*, 269, 499–505. https://doi.org/10.1098/rspb.2001.1863

Grath, S., & Parsch, J. (2016). Sex biased gene expression. *Annual Review of Genetics*, 50, 29–44. https://doi.org/10.1146/annurev-genet-120215-035429

Gu, L., & Walters, J. R. (2017). Evolution of sex chromosome dosage compensation in animals: A beautiful theory, undermined by facts and bedeviled by details. *Genome Biology and Evolution*, 9(9), 2461–2476. https://doi.org/10.1093/gbe/evx154

Hanly, J. (2017). *Developmental basis of wing pattern diversity in heliconius butterflies* (Doctoral thesis). https://doi.org/10.17863/CAM.13701

Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M., & Stanke, M. (2016). BRAKER1: Unsupervised RNA-seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*, 32(5), 767–769. https://doi.org/10.1093/bioinformatics/btv661

Hu, T. T., Eisen, M. B., Thornton, K. R., & Andolfatto, P. (2013). A second-generation assembly of the *Drosophila simulans* genome provides new insights into patterns of lineage-specific divergence. *Genome Research*, 23, 89–98. https://doi.org/10.1101/gr.141689.112

Huylmans, A. K., Macon, A., & Vicoso, B. (2017). Global dosage compensation is ubiquitous in Lepidoptera, but counteracted by the masculinization of the Z chromosome. *Molecular Biology and Evolution*, 34(10), 2637–2649. https://doi.org/10.1093/molbev/msx190

Immonen, E., Sayadi, A., Bayram, H., & Arnqvist, G. (2017). Mating changes sexually dimorphic gene expression in the seed beetle *Callosobruchus maculatus*. *Genome Biology and Evolution*, 9, 677–699. https://doi.org/10.1093/gbe/evx029

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, 12, 357–360. https://doi.org/10.1038/nmeth.3317

Kirkpatrick, M., & Hall, D. W. (2004). Male biased mutation, sex linkage, and the rate of adaptive evolution. *Evolution*, 58, 437–440. https://doi.org/10.1111/j.0014-3820.2004.tb01659.x

Kozak, K. M., Wahlberg, N., Neild, A. F. E., Dasmahapatra, K. K., Mallet, J., & Jiggins, C. D. (2015). Multilocus species trees show the recent adaptive radiation of the mimetic Heliconius butterflies. *Systematic Biology*, 64(3), 505–524. https://doi.org/10.1093/sysbio/syv007

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lomsadze, A., Burns, P. D., & Borodovsky, M. (2014). Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Research*, 42(15), e119–e119. https://doi.org/10.1093/nar/gku557

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Mackay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., … Gibbs, R. A. (2012). The *Drosophila melanogaster* genetic reference panel. *Nature*, 482, 173–178. https://doi.org/10.1038/nature10811

Macpherson, M., Sella, G., Davis, J. C., & Petrov, D. A. (2007). Genome-wide spatial correspondence between non-synonymous divergence and neutral polymorphism reveals extensive adaptation in Drosophila. *Genetics*, 177, 2083–2089. https://doi.org/10.1534/genetics.107.080226

Mank, J. (2017). The transcriptional architecture of phenotypic dimorphism. *Nature Ecology and Evolution*, 1, 0006.

Mank, J., Nam, K., & Ellegren, H. (2010). Faster-Z evolution is predominantly due to genetic drift. *Molecular Biology and Evolution*, 27, 661–670. https://doi.org/10.1093/molbev/msp282

Mank, J., Vicoso, B., Berlin, S., & Charlesworth, B. (2010). Effective population size and the Faster-X effect: Empirical results and their interpretation. *Evolution*, 64, 663–674. https://doi.org/10.1111/j.1558-5646.2009.00853.x

Martin, S. H., Möst, M., Palmer, W. J., Salazar, C., McMillan, W. O., Jiggins, F. M., & Jiggins, C. D. (2016). Natural selection and genetic diversity in the butterfly *Heliconius melpomene*. *Genetics*, 203(1), 525–541. https://doi.org/10.1534/genetics.115.183285

McDonald, J. H., & Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in Drosophila. *Nature*, *351*, 652–654. https://doi.org/10.1038/351652a0

Meisel, R. P., & Connallon, T. (2013). The faster-X effect: Integrating theory and data. *Trends in Genetics*, *29*, 537–544. https://doi.org/10.1016/j.tig.2013.05.009

Meisel, R. P., Malone, J. H., & Clark, A. G. (2012). Faster-X evolution of gene expression in Drosophila. *PLoS Genetics*, *8*, e1003013. https://doi.org/10.1371/journal.pgen.1003013

Messer, P. W., & Petrov, D. A. (2013). Frequent adaptation and the McDonald-Kreitman test. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(21), 8615–8620. https://doi.org/10.1073/pnas.1220835110

Miyata, T., Hayashida, H., Kuma, K., Mitsuyasu, K., & Yasunaga, T. (1987). Male-driven molecular evolution: A model and nucleotide sequence analysis. *Cold Spring Harbor Symposia on Quantitative Biology*, *52*, 863–867. https://doi.org/10.1101/SQB.1987.052.01.094

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, *5*(7), 621–628. https://doi.org/10.1038/nmeth.1226

Nielsen, R., Bustamante, C., Clark, A. G., Glanowski, S., Sackton, T. B., Hubisz, M. J., ... Cargill, M. (2005). A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biology*, *3*, e170. https://doi.org/10.1371/journal.pbio.0030170

O'Neill, M. B., Mortimer, T. D., & Pepperell, C. S. (2015). Diversity of *Mycobacterium tuberculosis* across evolutionary scales. *PLoS Pathogens*, *11*, e1005257. https://doi.org/10.1371/journal.ppat.1005257

Orr, H. A. (2010). The population genetics of beneficial mutations. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, *365*, 1195–1201. https://doi.org/10.1098/rstb.2009.0282

Orr, H. A., & Betancourt, A. J. (2001). Haldane's sieve and adaptation from the standing genetic variation. *Genetics*, *157*, 875–884.

Parisi, M., Nuttall, R., Edwards, P., Minor, J., Naiman, D., Lü, J., ... Oliver, B. (2004). A survey of ovary, testis-, and soma biased gene expression in *Drosophila melanogaster* adults. *Genome Biology*, *5*, R40. https://doi.org/10.1186/gb-2004-5-6-r40

Parisi, M., Nuttall, R., Naiman, D., Bouffard, G., Malley, J., Andrews, J., ... Oliver, B. (2003). Paucity of genes on the Drosophila X chromosome showing male biased expression. *Science*, *299*, 697–700. https://doi.org/10.1126/science.1079190

Pool, J. E., & Nielsen, R. (2007). Population size changes reshape genomic patterns of diversity. *Evolution*, *61*(12), 3001–3006. https://doi.org/10.1111/j.1558-5646.2007.00238.x

Ranwez, V., Harispe, S., Delsuc, F., & Douzery, E. J. P. (2011). MACSE: Multiple alignment of coding sequences accounting for frameshifts and stop codons. *PLoS ONE*, *6*, e22594. https://doi.org/10.1371/journal.pone.0022594

Rastogi, A., & Gupta, D. (2014). GFF-Ex: A genome feature extraction package. *BMC Research Notes*, *7*, 315. https://doi.org/10.1186/1756-0500-7-315

Rice, W. R. (1984). Sex chromosomes and the evolution of sexual dimorphism. *Evolution*, *38*, 735–742. https://doi.org/10.1111/j.1558-5646.1984.tb00346.x

Rinn, J. L., & Snyder, M. (2005). Sexual dimorphism in mammalian gene expression. *Trends in Genetics*, *21*(5), 298–305. https://doi.org/10.1016/j.tig.2005.03.005

Rousselle, M., Faivre, N., Ballenghien, M., Galtier, N., & Nabholz, B. (2016). Hemizygosity enhances purifying selection: Lack of fast-Z evolution in two satyrine butterflies. *Genome Biology and Evolution*, *8*, 3108–3119. https://doi.org/10.1093/gbe/evw214

Sackton, T. B., Corbett-Detig, R. B., Nagaraju, J., Vaishna, L., Arunkumar, K. P., & Hartl, D. L. (2014). Positive selection drives faster-Z evolution in silkmoths. *Evolution*, *68*, 2331–2342.

Stanke, M., Diekhans, M., Baertsch, R., & Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics*, *24*(5), 637–644. https://doi.org/10.1093/bioinformatics/btn013

The Heliconius Genome Consortium (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, *487*, 94–98.

Thornton, K., Bachtrog, D., & Andolfatto, P. (2006). X chromosomes and autosomes evolve at similar rates in Drosophila: No evidence for faster-X protein evolution. *Genome Research*, *16*, 498–504. https://doi.org/10.1101/gr.4447906

Thornton, K., & Long, M. (2002). Rapid divergence of gene duplicates on the *Drosophila melanogaster* X chromosome. *Molecular Biology and Evolution*, *19*, 918–925. https://doi.org/10.1093/oxfordjournals.molbev.a004149

Van Belleghem, S. M., Baquero, M., Papa, R., Salazar, C., McMillan, W. O., Counterman, B. A., ... Martin, S. H. (2018). Patterns of Z chromosome divergence among *Heliconius* species highlight the importance of historical demography. *Molecular Ecology*, *27*(19), 3852–3872. https://doi.org/10.1111/mec.14560

van Schooten, B., Jiggins, C. D., Briscoe, A. D., & Papa, R. (2016). Genome-wide analysis of ionotropic receptors provides insight into their evolution in *Heliconius* butterflies. *BMC Genomics*, *17*(1), 254. https://doi.org/10.1186/s12864-016-2572-y

Vicoso, B., & Charlesworth, B. (2006). Evolution on the X chromosome: Unusual patterns and processes. *Nature Reviews Genetics*, *7*, 645–653. https://doi.org/10.1038/nrg1914

Vicoso, B., & Charlesworth, B. (2009). Effective population size and the faster-X effect: An extended model. *Evolution*, *63*, 2413–2426. https://doi.org/10.1111/j.1558-5646.2009.00719.x

Walters, J. R., Hardcastle, T. J., & Jiggins, C. D. (2015). Sex chromosome dosage compensation in Heliconius butterflies: Global yet still incomplete? *Genome Biology and Evolution*, *7*, 2545–2559. https://doi.org/10.1093/gbe/evv156

Wright, A. E., Zimmer, F., Harrison, P. W., & Mank, J. E. (2015). Conservation of regional variation in sex-specific sex chromosome regulation. *Genetics*, *201*(2), 587–598. https://doi.org/10.1534/genetics.115.179234

Yu, Q.-Y., Fang, S.-M., Zhang, Z., & Jiggins, C. D. (2016). The transcriptome response of Heliconius melpomene larvae to a novel host plant. *Molecular Ecology*, *25*(19), 4850–4865. https://doi.org/10.1111/mec.13826

Zakharov, E., Caterino, M., & Sperling, F. (2004). Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera: Papilionidae). *Systematic Biology*, *53*, 193–215. https://doi.org/10.1080/10635150490423403

Zhang, Z., Hambuch, T., & Parsch, J. (2004). Molecular evolution of sex biased genes in Drosophila. *Molecular Biology and Evolution*, *21*, 2130–2139. https://doi.org/10.1093/molbev/msh223

Zhou, Q., & Bachtrog, D. (2012). Sex-specific adaptation drives early sex chromosome evolution in Drosophila. *Science*, *337*, 341–345. https://doi.org/10.1126/science.1225385

SUPPORTING INFORMATION

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