Patterns of Z chromosome divergence among *Heliconius* species highlight the importance of historical demography

Steven M. Van Belleghem\(^1,2,3,4\) | Margarita Baquero\(^2\) | Riccardo Papa\(^3\) | Camilo Salazar\(^5\) | W. Owen McMillan\(^4\) | Brian A. Counterman\(^2\) | Chris D. Jiggins\(^1\) | Simon H. Martin\(^1\)

\(^1\)Department of Zoology, University of Cambridge, Cambridge, UK
\(^2\)Department of Biological Sciences, Mississippi State University, Mississippi State, MS, USA
\(^3\)Department of Biology, Center for Applied Tropical Ecology and Conservation, University of Puerto Rico, Rio Piedras, Puerto Rico
\(^4\)Smithsonian Tropical Research Institute, Apartado, Panamá, Panamá
\(^5\)Biology Program, Faculty of Natural Sciences and Mathematics, Universidad del Rosario, Carrera, Bogota, Colombia

**Correspondence**
Steven M. Van Belleghem, Department of Zoology, University of Cambridge, Cambridge, UK.
Email: vanbelleghemsteven@hotmail.com

**Funding information**
COLCIENCIAS, Grant/Award Number: FP44842-5-2017; H2020 European Research Council, Grant/Award Number: 339873; National Science Foundation, Grant/Award Number: DEB 1257689; St John's College, Cambridge; NSF EPSCoR RII Track-2 FEC (OIA 1736026)

**Abstract**
Sex chromosomes are disproportionately involved in reproductive isolation and adaptation. In support of such a “large-X” effect, genome scans between recently diverged populations and species pairs often identify distinct patterns of divergence on the sex chromosome compared to autosomes. When measures of divergence between populations are higher on the sex chromosome compared to autosomes, such patterns could be interpreted as evidence for faster divergence on the sex chromosome, that is “faster-X”, barriers to gene flow on the sex chromosome. However, demographic changes can strongly skew divergence estimates and are not always taken into consideration. We used 224 whole-genome sequences representing 36 populations from two *Heliconius* butterfly clades (*H. erato* and *H. melpomene*) to explore patterns of Z chromosome divergence. We show that increased divergence compared to equilibrium expectations can in many cases be explained by demographic change. Among *Heliconius erato* populations, for instance, population size increase in the ancestral population can explain increased absolute divergence measures on the Z chromosome compared to the autosomes, as a result of increased ancestral Z chromosome genetic diversity. Nonetheless, we do identify increased divergence on the Z chromosome relative to the autosomes in parapatric or sympatric species comparisons that imply postzygotic reproductive barriers. Using simulations, we show that this is consistent with reduced gene flow on the Z chromosome, perhaps due to greater accumulation of incompatibilities. Our work demonstrates the importance of taking demography into account to interpret patterns of divergence on the Z chromosome, but nonetheless provides evidence to support the Z chromosome as a strong barrier to gene flow in incipient *Heliconius* butterfly species.

**KEYWORDS**
absolute divergence measures, demography, *Heliconius*, large-X effect, relative divergence measures, speciation

---

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2018 The Authors. Molecular Ecology Published by John Wiley & Sons Ltd

1 INTRODUCTION

Comparisons between genomes of diverging populations or species have revealed elevated differentiation on the sex chromosomes in several animals, such as flycatchers (Ellegren et al., 2012), crows (Poelstra et al., 2014), Darwin’s finches (Lamichhaney et al., 2015), ducks (Lavretsky et al., 2015) and Heliconius butterflies (Kronforst et al., 2013; Martín et al., 2013; Van Belleghem et al., 2017). These patterns of elevated sex chromosome divergence are sometimes readily interpreted as the result of increased reproductive isolation and reduced admixture on the sex chromosomes and, thus, ascribed to a large-X effect (Box 1). However, it remains unresolved whether such elevated sex-linked divergence actually results from more rapid accumulation of isolating barriers on the sex chromosome or could be explained by differences in effective population size between the sex chromosomes and the autosomes (Meisel & Connallon, 2013; Pool & Nielsen, 2007; Wolf & Ellegren, 2017).

When comparing divergence between genomic regions, such as sex chromosomes versus autosomes, measures of population divergence are influenced by within-population diversity (Charlesworth, 1998; Cruickshank & Hahn, 2014) (Box 2). This is explicitly the case for relative measures such as $F_{ST}$, but also less directly for absolute measures of divergence such as $d_{XY}$. For absolute divergence measures, this is because the genetic divergence between two alleles sampled from two species includes both divergence accumulated postspeciation, but also diversity already present in the ancestral population before the split. The latter is strongly dependent on effective population size. In a population under equilibrium conditions where the two sexes have an identical distribution of offspring number, the X chromosome-effective population size and genetic diversity is expected to be three-quarters that of the autosomes. Deviations from this ratio can result from multiple unique features of the sex chromosomes (Box 1), and population size changes in particular can have strong differential influence on sex chromosome compared to autosome diversity (Pool & Nielsen, 2007). Previous studies attempted to control for differences in effective population size on the sex chromosome, for instance among recently diverged duck species from Mexico, but such studies generally do not account for population size changes (Lavretsky et al., 2015). In order to interpret both relative and absolute measures of divergence on the sex chromosomes as evidence of a disproportionate contribution to species divergence and/or reduced admixture, we need to also account for demographic changes that can influence diversity of the sex chromosomes.

Here, we explore diversity and divergence on the Z chromosome relative to the autosomes among populations of the Heliconius erato and Heliconius melpomene butterfly clades, using these different measures. The H. erato and H. melpomene clades diverged about 13 million years ago (Kozak et al., 2015) and represent unpalatable and warningly coloured butterflies that have independently radiated into many divergent geographic races and reproductively isolated species. Within both clades, speciation has been accompanied by shifts in Müllerian mimicry (Mallet, McMillan, & Jiggins, 1998), and where populations come into contact, hybrid phenotypes usually have reduced survival rates due to strong frequency-dependent selection against intermediate colour pattern phenotypes (Jiggins, McMillan, Neukirchen, Mallet, & Nw, 1996; Mallet & Barton, 1989; Merrill et al., 2012; Naisbit, Jiggins, & Mallet, 2001). Two species, H. himera and H. e. chestertonii, are geographic replacements of H. erato in dry Andean valleys. They are partially reproductively isolated, but individuals of hybrid ancestry make up about 10% of the population in narrow transition zones between forms (McMillan, Jiggins, & Mallet, 1997; Merrill, Chia, & Nadeau, 2014; Muñoz, Salazar, Castaño, Jiggins, & Linares, 2010). Similarly, H. cydno and H. timareta are geographic replacements of each other and both are broadly sympatric with H. melpomene. Here, both species are reproductively isolated from H. melpomene by a combination of pre- and postmating isolation (Merot, Salazar, Jiggins, & Joron, 2017). Species integrity does not seem to involve structural variation such as chromosomal inversions (Davey et al., 2017). Instead, reproductive barriers include strong selection against hybrids, mate choice and postzygotic incompatibilities (Davey et al., 2017). Assortative mating has evolved in both the H. erato and H. melpomene clades (Jiggins, Naisbit, Coe, & Mallet, 2001; McMillan et al., 1997; Merrill et al., 2014; Muñoz et al., 2010). In the H. erato clade, sterility and reciprocal-cross asymmetry of hybrid sterility have been reported in crosses between H. erato and H. e. chestertonii (Muñoz et al., 2010), but hybrid sterility is absent between H. erato and H. himera (McMillan et al., 1997). In the H. melpomene clade, female sterility (Haldane’s rule; Box 1) and reciprocal-cross asymmetry of hybrid sterility occur in crosses between H. melpomene and H. cydno (Naisbit, Jiggins, Linares, Salazar, & Mallet, 2002), H. melpomene and H. heurippa (Salazar et al., 2005) and H. melpomene and H. timareta (Sánchez et al., 2015), as well as between allopatric H. melpomene populations from French Guiana and those from Panama and Colombia (Jiggins, Linares et al., 2001). In support of a large-X effect, sterility in these crosses (H. melpomene × H. cydno, H. melpomene × H. heurippa and H. melpomene × H. timareta) was found to be Z-linked.

The presence of incipient species pairs with different levels of reproductive isolation allows us to examine the relative rate of autosomal and Z chromosomal evolution and the factors that are likely influencing patterns of divergence. We take advantage of a large genomic data set composed of 224 whole genomes representing 20 populations of the H. erato clade and 16 populations of the H. melpomene clade. We also use simulations to evaluate the effect that demographic changes have on the estimate of relative rates of divergence on the Z versus the autosomes and demonstrate that in many comparisons, demography can explain much of the observed elevated divergence on the Z relative to the autosomes. However, by taking into account geographic distance or autosomal divergence as a proxy for gene flow, we show that there is evidence for increased divergence on the Z chromosome for species pairs with known postzygotic reproductive barriers. These rates of increased divergence likely reflect reduced admixture on the Z chromosome and provide support for the Z chromosome being a greater barrier to gene flow in some incipient Heliconius butterfly species.
genes expressed in male reproductive tissues (reviewed in Meisel & Connallon, 2013). In Lepidoptera (butterflies and moths), Haldane by three widespread intrinsic postmating effects (Johnson & Lachance, 2012; Turelli & Moyle, 2007): (i) Haldane Sex chromosomes have been repeatedly shown to have a disproportionate role during speciation (Coyne & Orr, 2004), demonstrated due to their uniparental inheritance and, thus, also explain asymmetric reproductive isolation (Turelli & Moyle, 2007). Hence, hemizygous expression of recessive alleles on the sex chromosome has been put forward as a cause for the disproportionate role of the sex chromosomes during speciation (dominance theory) (Turelli & Orr, 1995).

In contrast to the dominance theory, there is, however, a large body of observations and theory that propose alternative or additional explanations that can cause a large-X effect and/or explain Haldane's rule (Presgraves, 2008; Wu & Davis, 1993). These factors include faster male evolution resulting from intense sexual selection among males (Wu & Davis, 1993), meiotic drive (Frank, 1991), dosage compensation (Jablonka & Lamb, 1991) and faster-X evolution (Charlesworth, Coyne, & Barton, 1987; Vicoso & Charlesworth, 2006). Faster male evolution targets genes with male-biased effects, which may be sex-linked or autosomal, in which case it can cause Haldane's rule only when males are the heterogametic sex. By contrast, faster-X evolution does not necessarily depend on selection related to sex and is predicted if adaptive new mutations are on average partially recessive and thus are presented more readily to selection on the hemizygous sex chromosomes (Charlesworth et al., 1987). In Drosophila, faster-X evolution has been studied extensively. Although it is not ubiquitous, there is clear evidence for faster-X divergence and adaptation (Counterman, Ortiz-Barrientos, & Noor, 2004), particularly for X-linked genes expressed in male reproductive tissues (reviewed in Meisel & Connallon, 2013). In Lepidoptera (butterflies and moths), Haldane's rule and the large-X (or Z) effect have been reported for numerous species (Presgraves, 2002; Prowell, 1998; Sperling, 1994). As lepidopteran females are heterogametic ZW, while males are ZZ, the Z is equivalent to the X. However, as females are heterogametic, faster male evolution is insufficient to explain Haldane’s, but faster-X evolution remains a viable explanation (Sackton et al., 2014). Moreover, in Lepidoptera, the large-X effect extends beyond intrinsic isolating barriers and there are differences in many traits and behaviours that map disproportionately to the Z chromosome (Prowell, 1998; Sperling, 1994). These observations are consistent with the faster accumulation of differences on the Z chromosome (faster-X evolution).

Factors Affecting Sex/Autosome Diversity Ratios
Apart from population size changes, factors that can result in deviations from the expected three-quarter X/autosome (X/A) diversity ratio, and could thus potentially affect divergence measures, include (i) sex-biased demographic events leading to different effective population sizes of males and females (Charlesworth, 2001), (ii) selective sweeps and background selection differently affecting the sex chromosomes (Charlesworth, 2012) and (iii) differences in mutation rates between sexes or between the sex chromosomes and the autosomes (Johnson & Lachance, 2012; Sayres & Makova, 2011).

First, different population sizes of males and females can influence the X/A diversity ratio because two-thirds of the X chromosome population is transmitted through females. A male-biased population would thus decrease the X/A diversity ratio below three-quarters, whereas a female-biased population would increase the ratio. This effect would be opposite in female heterogametic sex systems (ZW).

Second, the hemizygous expression of the sex chromosome could result in both higher purifying selection and more efficient selection of beneficial recessive mutations (selective sweeps) and result in a decrease in the expected X/A diversity ratio (Charlesworth et al., 1987). Additionally, differences in recombination rates can lead to different extent of loss of variation through linked selection and thus background selection (Charlesworth, 2012). In Lepidoptera, meiosis is commonly achiastic (no recombination) in the heterogametic sex (females) (Suomalainen, Cook, & Turner, 1973; Turner & Sheppard, 1975). A reduction in recombination rate on the

(Continues)
sex chromosomes compared to autosomes, which is commonly found in Drosophila (Vicoso & Charlesworth, 2009), should thus not be expected to decrease Z/A diversity ratios through increased background selection in Heliconius. On the other hand, it has been suggested that effective recombination should be higher, and thus background selection lower, for the Z chromosome when recombination is absent in females (Charlesworth, 2012). This is because the Z chromosomes spend two-thirds of their time in recombining males, whereas autosomes only spend half of their time in recombining males.

Third, because the male germ line generally involves more cell divisions and thus opportunities for replication errors, sex-linked genes may have different mutation rates. Because X-linked genes spend only one-third of their time in males and two-thirds of their time in females, the X chromosome may be subjected to a lower mutation rate. Conversely, the Z chromosome spends two-thirds of its time in males and may therefore become enriched in genetic variation compared to the autosomes (Johnson & Lachance, 2012; Sayres & Makova, 2011; Vicoso & Charlesworth, 2006). Such increased mutation rates on the Z chromosome could also increase the rate of divergence between Z chromosomes (Kirkpatrick & Hall, 2004).

**FIGURE B1** The effect of population size on the coalescent and measures of diversity and divergence. The branches represent two populations, 1 and 2, that have split at a certain time (grey dashed line). This branching event occurs on two chromosomes that have a different population size, such as the autosomes (grey) and X chromosome (green). The black lines show the coalescent of two alleles in each population. The branches show that the coalescence process is influenced by the split time as well as the population size. Population size affects the nucleotide diversity within each population (π), the total nucleotide diversity (πT) and absolute divergence dXY, but not dA as indicated by the vertical coloured lines. For dA, the average of the within-population nucleotide diversity (πS) is used as the estimate of the ancestral nucleotide diversity (πanc). The influence of population size on FST can be seen as resulting from a decrease in the denominator (πT), but not in the numerator (πT and π change proportionally).
BOX 2 Measures of Divergence Depend on Population Size

The mutational diversity in present-day samples is directly related to population size, structure and age. This is because population size determines the rate of coalescence within and between populations (Figure B1). This relationship can be seen with \( F_{ST} \), which was developed to measure the normalized difference in allele frequencies between populations (Wright, 1931). The dependence of \( F_{ST} \) on population size can be understood by interpreting \( F_{ST} \) as the rate of coalescence within populations compared to the overall coalescence rate (Slatkin & Voelm, 1991). Comparing pairs of populations with different effective population sizes will therefore show distinct \( F_{ST} \) estimates even when the split time is the same (Charlesworth, 1998). Absolute divergence \( d_{XY} \) is the average number of pairwise differences between sequences sampled from two populations (Nei & Li, 1979). The measure \( d_{XY} \) is not influenced by changes to within-population diversity that occur after the split, but does depend on diversity that was present at the time the populations split (Gillespie & Langley, 1979). Therefore, population pairs that had a smaller population size at the time they split will, consequently, have smaller \( d_{XY} \) estimates.

To compare pairs of populations that had different ancestral population sizes, \( d_a \) has been proposed, which subtracts an estimate of the diversity in the ancestral population from the absolute divergence measure \( d_{XY} \) (Nei & Li, 1979). An approximation of ancestral diversity can be obtained by taking the average of the nucleotide diversity observed in the two present-day populations. Such a correction should result in the "net" nucleotide differences that have accumulated since the time of split.

To show how these different divergence measures perform, we simulated a simplified population split with varying degrees of migration (m) (Figure B2). As expected, the values \( F_{ST} \), \( d_{XY} \) and \( d_a \) all increase when migration between populations decreases. \( F_{ST} \) and \( d_{XY} \) are clearly influenced by population size. While for \( d_{XY} \), this results from the variation present at the time of the split, \( F_{ST} \) does not show a simple linear relationship with population size. Only \( d_a \) represents the net accumulation of differences that can be compared between populations of different sizes, such as the sex chromosomes versus autosomes (but see section 3.3 in Results and discussion).

FIGURE B2 Simulated effect of population size differences on divergence measures \( F_{ST} \), \( d_{XY} \) and \( d_a \). Simulations were performed for two populations that split 4 million generations ago and vary in their degree of migration (m). A lower effective population size, such as for the X chromosome (green) compared to autosomes (grey), results in higher \( F_{ST} \) and lower \( d_{XY} \) estimates, but has no effect on \( d_a \) under these assumptions.

(Panama, \( n = 3 \)), \( H. m. melpomene \) (French Guiana, \( n = 10 \)), \( H. m. melpomene \) (Colombia, \( n = 5 \)), \( H. m. rosinia \) (Panama, \( n = 10 \)), \( H. m. malleti \) (Colombia, \( n = 10 \)), \( H. m. vulcanus \) (Colombia, \( n = 10 \)), \( H. m. plesseni \) (Ecuador, \( n = 3 \)), \( H. m. aglaope \) (Peru, \( n = 4 \)), \( H. m. amaryllis \) (Peru, \( n = 10 \)), and \( H. m. nanna \) (Brazil, \( n = 4 \)). Three populations were sampled from the \( H. timareta \) clade: \( H. heurippa \) (Colombia, \( n = 3 \)), \( H. t. thekinoe \) (Peru, \( n = 10 \)) and \( H. t. florencia \) (Colombia, \( n = 10 \)). Three populations were sampled from the \( H. cydno \) clade: \( H. c. chionus \) (Panama, \( n = 10 \)), \( H. c. cordula \) (Venezuela, \( n = 3 \)) and \( H. c. zelinde \) (Colombia, \( n = 10 \)).

2.2 Sequencing and genotyping

Whole-genome 100-bp paired-end Illumina resequencing data from \( H. erato \) and \( H. melpomene \) clade samples were aligned to the \( H. erato \) v1 (Van Belleghem et al., 2017) and \( H. melpomene \) v2 (Davey et al., 2016) reference genomes, respectively, using \( bwa \) v0.7 (Li, 2013). PCR duplicated reads were removed using \( picard \) v1.138 (http://picard.sourceforge.net) and sorted using \( samtools \) (Li et al., 2009). Genotypes were called using the genome analysis tool kit (GATK) Haplotypecaller (Van der Auwera et al., 2013). Individual genomic VCF records (gVCF) were jointly genotyped using GATK’s genotype GVCFs. Genotype calls were only considered in downstream analysis if they had a minimum depth (DP) \( \geq 10 \), maximum depth (DP) \( \leq 100 \) (to avoid false SNPs due to mapping in repetitive regions), and for variant calls, a minimum genotype quality (GQ) \( \geq 30 \). The W chromosome has not been identified in Heliconius, but read depth comparisons between Z and autosomes in males and females (e.g., see supplemental material Martin et al., 2013) support the hypothesis...
FIGURE 1  Diversity and sampling of the *Heliconius erato* and *Heliconius melpomene* clade. (a) *Heliconius erato chestertonii* (green) is reproductively isolated from *H. erato* (pink) by spatial separation (parapatry), mate choice and (asymmetric) reduced hybrid fertility of both sexes (i.e., no Haldane’s rule). *Heliconius himera* (blue) is reproductively isolated from *H. erato* by spatial separation and mate choice, but hybrids show no reduced fertility. *Heliconius cydno* (green) and *H. timareta* (blue) occur sympatrically with *H. melpomene* (pink) populations, but are both reproductively isolated by strong mate choice and (asymmetric) reduced fertility of F1 hybrids (i.e., Haldane’s rule). (b) Localities of sampled populations included in this study. Within *H. erato*, *H. melpomene*, *H. timareta* and *H. cydno* names represent different races that display distinct colour patterns. Shapes represent geographic regions: Mexico and Panama (diamond), west of the Andes (triangles) and east of the Andes (circles). (c) PCA plots of autosomal SNP variation. Note that *H. m. nanna* has not been included in the PCA as the signal of geographic isolation between *H. m. nanna* and the other populations dominates the signal (see Fig. S1) [Colour figure can be viewed at wileyonlinelibrary.com]
that there is no significant mapping of W-linked reads to the Z and the W is, thus, unlikely to interfere with genotyping. The absence of mapping of W-linked reads to the Z is likely due to the degenerate sequence and highly repetitive nature of the W chromosome. The data set contained 31 and 11 female samples (ZW) randomly distributed among the H. erato and H. melpomene clade populations, respectively (Tables S1 and S2). These samples had lower read and, consequentially, lower genotyping coverage for the Z chromosome, but using the stringent filter thresholds, this does not affect variant and nonvariant sites differently and does not affect measures of nucleotide diversity and divergence.

2.3 Population structure and historical demography

To discern population structure among the sampled H. erato and H. melpomene clade individuals, we performed principal component analysis (PCA) using EIGENSTRAT SmartPCA (Price et al., 2006). For this analysis, we only considered autosomal biallelic sites that had coverage in all individuals.

We inferred changes in the historical population size from individual consensus genome sequences using pairwise sequentially Markovian coalescent (PSMC) analyses as implemented in MSMC (Schiffels & Durbin, 2014). This method fits a model of changing population size by estimating the distribution of times to the most recent common ancestor along diploid genomes. When used on single diploid genomes, this method is similar to pairwise sequentially Markovian coalescent (PSMC) analyses (Li & Durbin, 2011). Genotypes were inferred from bwa v0.7 (Li, 2013) mapped reads separately from previous genotyping analysis using samtools v0.1.19 (Li et al., 2009) according to authors’ suggestions. This involved a minimum mapping (-q) and base (-Q) quality of 20 and adjustment of mapping quality (-C) 50. A mask file was generated for regions of the genome with a minimum coverage depth of 30 and was provided together with heterozygosity calls to the MSMC tool. MSMC was run on heterozygosity calls from all contiguous scaffolds longer than 500 kb, excluding scaffolds on the Z chromosome. We scaled the PSMC estimates using a generation time of 0.25 years and a mutation rate of 2e-9 as estimated for H. melpomene (Keightley et al., 2014).

2.4 Population genomic diversity and divergence statistics

We first estimated diversity within populations as well as divergence between parapatric and sympatric populations in nonoverlapping 50-kb windows along the autosomes and Z chromosome using python scripts and egglib (data presented in Figure 3, 7 and S4–6) (De Mita & Siol, 2012). We only considered windows for which at least 10% of the positions were genotyped for at least 75% of the individuals within each population. For females, haploid was enforced when calculating divergence and diversity statistics. Sex of individuals was inferred from heterozygosity on the Z. FST was estimated as in Hudson, Slatkin, and Maddison (1992), as

\[ F_{ST} = \frac{\pi_T - \pi_S}{\pi_T} \]

with nucleotide diversity in a population (\(\pi\)) calculated as

\[ \pi = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} d_{ij,k}}{\binom{n}{2}} \]

and average within-population nucleotide diversity (\(\bar{\pi}\)) calculated as the weighted (\(w\)) average of the nucleotide diversity (\(\pi\)) within each population in and k, as

\[ \bar{\pi} = w\pi_1 + (1 - w)\pi_2. \]

Total nucleotide diversity (\(\pi_T\)) was calculated as the average number of nucleotide differences per site between two DNA sequences in all possible pairs in the sampled population (Hudson et al., 1992), as

\[ \pi_T = \frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} d_{ij,k} + \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} d_{ij,2}}{\binom{n_1 + n_2}{2}}. \]

Between-population sequence divergence \(d_{XY}\) was estimated as the average pairwise difference between sequences sampled from two different populations (Nei & Li, 1979), as

\[ d_{XY} = \frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} d_{ij,2}}{n_1 + n_2}. \]

The relative measure of divergence \(d_s\) was calculated by subtracting \(d_{XY}\) with an estimate of the nucleotide diversity (\(\bar{\pi}\)) in the ancestral populations (Nei & Li, 1979),

\[ d_s = d_{XY} - \bar{\pi}. \]

Tajima’s D was calculated as a measure of deviation from a population evolving neutrally with a constant size, with negative values indicating an excess of rare alleles (selective sweep or population expansion) and positive values indicating a lack of rare alleles (balancing selection or population contraction) (Tajima, 1989). To overcome the problem of nonindependence between loci, estimates of the variance in nucleotide diversity (\(\pi\)) and Tajima’s D within populations along the genomes were obtained using block-jackknife deletion over 1-Mb intervals along the genome (chosen to be much longer than linkage disequilibrium in Heliconius (Martin et al., 2013)) (Künsch, 1989).

To calculate pairwise \(d_{XY}\) values between each individual, we subsampled the genomes by only considering genomic sites that were at least 500 bp apart and had coverage for at least one individual in each population (data presented in Figures 8 and 9). For the H. erato clade data set, this resulted in a high coverage data set with 322,082 and 15,382 sites on the autosomes and Z chromosome, respectively. For the H. melpomene clade data set, this resulted in 335,636 and 18,623 sites on the autosomes and Z chromosome,
respectively. Pairwise $d_{XY}$ values between each individual were used to evaluate the relationship between absolute genetic divergence ($d_{XY}$) and geographic distance using Mantel’s tests (Mantel, 1967). Mantel’s tests are commonly used to test for correlations between pairwise distance matrices and were performed using the R package VEGAN (Oksanen et al., 2016). Pairwise distances between populations were calculated from the average of the sample coordinates obtained for each population (Table S3, S4).

2.5 | Simulations

To compare patterns in our data to expectations, we simulated genealogies in 50-kb sequence windows under certain evolutionary scenarios. The simulations were performed with a population recombination rate $(4Ne_{R})$ of 0.01 using the coalescent simulator msms (Ewing & Hermisson, 2010). Subsequently, from the simulated genealogies, we simulated 50-kb sequences with a mutation rate of 2e-9 a Hasegawa–Kishino–Yano substitution model using seq-gen (Rambaut & Grass, 1997).

In a first set of simulations, we considered one population that underwent a single population size change of a magnitude ($\lambda$) ranging from 0.01 to 100 and at a certain moment backwards in time ($t$). In a second set of simulations, we considered pairs of populations that were connected through migration ($m$) ranging from 0 to 1e-6 and for which migration was reduced with a factor $d$ on the Z chromosome. To compare changes in the nucleotide diversity on autosomes and the Z chromosome, we simulated the Z chromosome as a separate population for which the effective population size was set to three-quarters of that of the autosomal population.

To compare populations with a different effective population size ($N_e$), such as the autosomes and the Z chromosome, we expressed time in generations and migration rates as a proportion of the effective population size. Comparable to the Heliconius sampling, we sampled five individuals from each population and ran 300 replicates for each parameter combination. Pseudocode to run the msms command lines are provided in Tables S5. Tajima's $D$, nucleotide diversity ($\pi$) and $d_{XY}$ were calculated from the simulated sequences using python scripts and egglib (De Mita & Siol, 2012).

3 | RESULTS AND DISCUSSION

3.1 | Population structure and historical demography in Heliconius erato and Heliconius melpomene

We mapped a total of 109 Heliconius erato clade resequenced genomes to the Heliconius erato v1 reference genome (Van Belleghem et al., 2017) and 115 Heliconius melpomene clade genomes to the Heliconius melpomene v2 reference genome (Davey et al., 2016). These samples represent 20 H. erato clade and 16 H. melpomene clade populations covering nearly the entire geographic distribution of these species groups (Figure 1a and b).

Phylogenies from whole-genome data for the H. erato clade and H. melpomene clade have been presented previously in Van Belleghem et al. (2017) and Martin et al. (2016), respectively. Given the difficulty in presenting phylogenies for hybridizing populations and species, we here instead summarized relationships using principal component analysis (PCA) of the autosomal SNP variation using EIGENSTRAT SmartPCA (Price et al., 2006). PCA grouped the H. erato clade samples mainly according to geography, apart from H. himera individuals from Ecuador and northern Peru and H. e. chestertonii from Colombia (Figure 1c). Four main geographic groups were apparent: populations from Mexico, Panama, west of the Andes and east of the Andes. Heliconius erato populations east of the Andes as far as 3000 km apart were closely clustered in the PCA. The separate grouping of H. himera and H. e. chestertonii individuals supports these populations as representing incipient species that maintain their integrity, despite ample opportunity for hybridization and gene flow (Arias et al., 2008; Jiggins et al., 1996; McMillan et al., 1997). In the PCA, H. himera was more closely related to the H. erato populations east of the Andes, whereas H. e. chestertonii was more closely related to the West Andean populations.

PCA of the H. melpomene clade grouped individuals from west of the Andes and Panama closely together, with H. melpomene from Colombia being most similar to this population pair (Figure 1c). Heliconius melpomene populations from east of the Andes further clustered in three distinct groups, largely in agreement with geographic distance: populations from the eastern slopes of the Andes, the French Guiana population and H. m. nanna from Brazil (Figures 1c and S1). While phylogenetic reconstructions have suggested that H. melpomene and the Heliconius cydno/timareta clades are reciprocally monophyletic (Dasmahapatra et al., 2012; Martin et al., 2013; 2016; Nadeau et al., 2013), such patterns are hard to interpret from the PCA and patterns of relatedness may be influenced by more recent admixture. Nevertheless, H. cydno and Heliconius timareta clustered distinctly. Heliconius cydno formed a distinct cluster with little difference between samples from Panama, west or east of the Andes. Heliconius timareta grouped most closely with Heliconius heurippa, consistent with previous analysis (Arias et al., 2014; Nadeau et al., 2013).

We inferred changes in the historical population size from individual consensus genome sequences using pairwise sequentially Markovian coalescent (PSMC) (Schiffels & Durbin, 2014) (Figures 2 and S2–3). Heliconius erato populations east of the Andes were inferred to have the strongest increase in population size, starting about 1 MYA. Similarly, but to a lesser extent, H. erato populations from west of the Andes and Panama are inferred to have had a continuous population size increase during the past million years. In contrast, after a period of population size increase, H. himera has undergone continuous population size decrease since about 300 KYA. The H. e. chestertonii population seems to have decreased in size since about 300 KYA and increased in size after 30 KYA. Note that all the population size estimates for H. himera and H. e. chestertonii, as well as H. erato from Panama and east and west of the Andes start to deviate about 300 to 400 KYA. In Mexico, the
sampled *H. erato* population has undergone a more recent steep population size increase after a period of population decrease. The absence of convergence of the population size of the Mexican population with the other *H. erato* populations agrees with an old divergence of this population that likely falls out of the detection limit of the PSMC method (Van Belleghem et al., 2017).

In the *H. melpomene* clade, we find that *H. cydno* populations and *H. melpomene* from east of the Andes have undergone a continuous population size increase since about 1 MYA (Figure 2). In contrast, *H. melpomene* from French Guiana only shows this population size increase up to 100 KYA, after which it has experienced slight population size decrease. *Heliconius melpomene* populations from west of the Andes, Panama and Colombia show a more complex demographic history with population increase up to 300 KYA, followed by decrease and increase again about 30 to 40 KYA. *Heliconius melpomene nanna* from East Brazil is characterized by a steep population size decrease since 500 KYA and a steep population size increase starting 40–50 KYA. Next, *H. timareta* (i.e., *H. t. florencia* and *H. t. thelxinoe*) is characterized by population size increase until 200 to 300 KYA, followed by population size decrease and increase again about 40 KYA. Finally, *H. heurippa* has a similar demographic history as the other *H. timareta* species, but with a continuous decrease in population size since about 300 KYA.

The PSMC estimates give inference up to about 1 MYA (Figure 2). In contrast, the split time between *H. melpomene* and *H. cydno* has been estimated between 0.9 and 1.4 MYA ago (~3.6–5.6 million *Heliconius* generations) (Kronforst et al., 2013; Lohse, Chmelik, Martin, & Barton, 2016). A similar split time can be expected between *H. melpomene* and *H. timareta*, as *H. timareta* and *H. cydno* likely diverged after the split from *H. melpomene* (Martin et al., 2016). While such divergence time estimates that account for migration are unavailable for *H. e. chestertonii* and *H. himera*, their split times from *H. erato* are likely also older than most of the time interval in which historical demography is inferred by PSMC (Flanagan et al., 2004). Therefore, the PSMC results likely reflect only demographic changes that have occurred after these populations split. We also note that PSMC does not account for hybridization, which might impact the inferred histories. However, almost all sympatric and parapatric species pairs showed very different population histories to one another (e.g., *H. himera* and *H. e. cyrbia*; *H. t. thelxinoe* and *H. m. amaryllis*), suggesting that the differences observed between populations are largely driven by real demographic change rather than artefacts of hybridization.

### 3.2 Z chromosome divergence in *Heliconius erato* and *Heliconius melpomene*

To compare rates of divergence between the Z chromosome and autosomes, we calculated three commonly used measures of divergence, $F_{ST}$, $d_{XY}$ and $d_a$, between incipient species and population pairs of *H. erato* and *H. melpomene* (Figures 3 and S4–6). All three measures of sequence divergence are calculated from mutational diversity in the data, but are each dependent on population size in different ways (Box 2). In *Heliconius*, $F_{ST}$ has been frequently used to identify regions in the genome under strong divergent selection (Martin et al., 2013; Nadeau et al., 2012; Van Belleghem et al., 2017). In

![FIGURE 2](https://wileyonlinelibrary.com)
comparisons between parapatric colour pattern races of both *H. erato* and *H. melpomene*, sharp *F*<sub>ST</sub> peaks are present near the major colour pattern loci, suggesting both strong divergent selection and reduced gene flow (Figure S4). Additionally, increased *F*<sub>ST</sub> values can be observed on the Z chromosome in comparisons between populations with assortative mating and hybrid inviability and sterility (Figures 3 and S4). However, *F*<sub>ST</sub> is influenced by effective population sizes (Box 2). It is therefore problematic to obtain insights about selection or migration when comparing genomic regions with different effective population sizes, such as the Z chromosome and autosomes. Given equal numbers of breeding males and females, the Z chromosome is expected to have an effective population size three-quarters that of the autosomes. Smaller population size and the resulting lower nucleotide diversity on the Z chromosome may therefore partly explain inflated *F*<sub>ST</sub> estimates on the Z chromosome.

In contrast, *d*<sub>XY</sub> values on the Z chromosome tend to be similar to or slightly lower than the average values on the autosomes in most species comparisons of the *H. erato* and *H. melpomene* clade (Figures 3 and S5). Under equilibrium conditions, *d*<sub>XY</sub> on the Z chromosome is expected to be three-quarters that of the autosomes at the time of the split. As time progresses and differences between populations accumulate, the proportion of the coalescent that is affected by the ancestral population size will become smaller and the ratio of *d*<sub>XY</sub> on the Z to *d*<sub>XY</sub> on the autosomes is expected to move towards one (Box 2). Estimating the exact split time is difficult, and finding the expected absolute divergence for the Z chromosome compared to the autosomes is complicated (Patterson, Richter, Gnerre, Lander, & Reich, 2006). However, in contrast to the expectation that the ratio of *d*<sub>XY</sub> will move towards one, *d*<sub>XY</sub> on the Z chromosome is higher than on the autosomes for *H. e. cyria–H. himera* and *H. e. venus–H. e. chestertonii* comparisons (Figure 3).

Finally, by subtracting an estimate of diversity in the ancestral population from the absolute divergence measure *d*<sub>XY</sub>, known as *d*<sub>a</sub> (Nei & Li, 1979), we obtain an estimate of nucleotide differences that have accumulated since the time of split (Box 2). The *d*<sub>a</sub> estimates show significantly higher divergence on the Z chromosome in the comparisons *H. himera–H. e. cyria*, *H. e. venus–H. e. chestertonii*, and in *H. melpomene–H. cydno* and *H. melpomene–H. timareta* (Figures 3 and S6). Overall, the increased *d*<sub>XY</sub> in the *H. e. cyria* and *H. himera* and the *H. e. venus–H. e. chestertonii* comparisons and the higher *d*<sub>a</sub> values on the Z chromosome relative to the autosomes appear to support a faster rate of divergence between *Heliconius* species pairs on the Z chromosome.
3.3 | Population size changes affect the Z chromosome differently

Apart from the overall difference in effective population size between the Z chromosome and autosomes, there are additional demographic factors that can contribute to differences in $F_{ST}$, $d_{XY}$ and $d_a$ values between the Z chromosome and autosomes. Population size changes can alter the equilibrium expectation that Z-linked diversity should be three-quarters of autosomal diversity (Pool & Nielsen, 2007). To explore this, we performed coalescent simulations of sequences from populations that underwent a single size change in the past, varying the time and magnitude of this event (Figure 4). In these simulated populations, the decrease in nucleotide diversity that follows population contraction occurs much faster than the increase in diversity that follows an expansion of the same magnitude (Figure 4a). This is because an increase in diversity requires mutation accumulation, whereas drift can more rapidly remove variation to reach a new equilibrium. Additionally, population size changes have proportionately stronger effects on diversity on the Z chromosome compared to the autosomes (Figure 4b). This results from populations with a smaller effective population size, such as the Z chromosome, converging faster to their new equilibrium after a population size change (Pool & Nielsen, 2007).

The result of population size change differently affecting the Z chromosome is that divergence measures are also differentially affected by population size change on the Z chromosome compared to the autosomes. The Z chromosome to autosome (Z/A) diversity ratio will be larger than expected in populations that experienced a recent expansion and smaller than expected in those that experienced a recent contraction (Figure 4b). Therefore, in pairwise comparisons, if population size change occurred in the ancestral population before the two populations split, it would alter the ancestral Z/A diversity ratio and therefore confound comparisons of divergence between Z and autosomes using either relative or absolute measures of divergence, as all are influenced by ancestral diversity (Figure 5). By contrast, if population size change occurred in one or both daughter populations after the split, it would affect the relative measures of divergence $F_{ST}$ and $d_a$ but not absolute divergence ($d_{XY}$), which is only dependent on ancestral diversity and not on diversity within each population. The effect size will depend on the timing and magnitude of the population size change. In our simulations, Z/A diversity ratios ranged from 0.40 to 0.86 under the most extreme simulated population size changes, compared to the expected diversity ratio of 0.75 under equilibrium expectations (Figure 4b). All the simulations were run for timescales relevant to Heliconius divergence and, therefore, demonstrate that a return to equilibrium values is unlikely after a population increase during the history of these species. In particular, a return to equilibrium Z/A diversity ratios after population size increase can be slow and long-lasting during the evolutionary history of a population.

3.4 | Demography and its influence on Z/A diversity ratios in Heliconius

To explore how population size changes might have affected Z/A diversity ratios and thus Z/A divergence comparisons within Heliconius clades, we used the behaviour of Tajima’s $D$ as a way to assess likely population size changes within species. Tajima’s $D$ is a population genetic measure commonly used to detect whether a locus is evolving neutrally in a population (Tajima, 1989). At a genomewide scale, negative values reflect population size expansion, whereas positive values can reflect population size decrease or population subdivision. Due to the different response of Tajima’s $D$ to
population size increase and decrease, Tajima's $D$ can give an indication of population size changes and their effect on nucleotide diversity. As the simulations show, negative Tajima's $D$ values (population size increase) are correlated with increased nucleotide diversity, whereas positive Tajima's $D$ values (population size decrease) are correlated with reduced nucleotide diversity (Figure 6). As with the Z/A diversity ratio, the timescale of the influence of population size change on Tajima's $D$ values is different for population expansion versus population contraction, as a population that has contracted returns to equilibrium faster than one that has expanded. Moreover, because smaller populations respond faster to such population size changes, the Tajima's $D$ values are also expected to be correlated with Z/A diversity ratios. Although this results in a complex relationship (Figure 7), *H. erato* and *H. melpomene* clade populations that showed more negative Tajima's $D$ values (population size increase) all had higher nucleotide diversity ($\pi$) values, as well as higher Z/A diversity ratios (Figure 7). Z/A diversity ratios ranged from 0.38 to 0.93 in the *H. erato* clade and from 0.47 to 0.85 in the *H. melpomene* clade, similar to the range of values obtained in the simulations (Figure 4b and 7). It should be noted that multiple population size change events (e.g., population size expansion followed by a bottleneck), continuous increase and different durations of population size changes would further complicate the relation between Tajima's $D$ estimates and the expected nucleotide diversity as well as the Z/A diversity ratio. Potentially, this also explains the more extreme negative Tajima's $D$ and Z/A diversity ratios in several of the *H. erato* and *H. melpomene* clade populations compared to our simulated scenarios (Figure 7). Nevertheless, Tajima's $D$ estimates for the *H. erato*
and H. melpomene clade do capture the average demographic history as inferred using the PSMC method. We find more negative Tajima’s D values for populations that have undergone continuous population size increase and higher Tajima’s D values for populations that have undergone steep population declines (Figures 2 and 7). Therefore, the patterns among these Heliconius populations suggest that differences in nucleotide diversity as well as differences in the Z/A diversity ratios are likely driven at least in part by population size changes. Given that samples assigned to a population were collected in close proximity, it is unlikely that estimated Tajima’s D values are influenced by hidden population structure in the data (which could result in positive Tajima’s D values).

Broad patterns of nucleotide diversity can also give insights into the general demography of the studied species. Higher nucleotide diversity in the H. erato clade as compared to the H. melpomene clade is consistent with the generally greater abundance of H. erato observed in nature (Mallet, Jiggins, & McMillan, 1998) (Figure 7). These population size differences likely also explain the large differences in absolute divergence levels among the H. erato clade populations compared to the H. melpomene clade populations (Figure 8). Absolute divergence between H. melpomene and H. timareta or H. cydno, for instance, is smaller than within-population diversity of most H. erato populations (Figure 7), despite the former pairs being clearly distinct species.

While changes in population size can have strong effects on measures of sequence divergence, jointly considering patterns of variation on the Z chromosome and autosome can give further insights into the evolutionary history. Among H. erato populations from east of the Andes that show little differentiation, $d_{XY(Z)}/d_{XY(A)}$ ratios are above the 0.75 ratio that would be expected immediately after the populations split (0.91 ± 0.11) (Figure 8) and there is also increased Z/A nucleotide diversity (Figure 7). This likely resulted from population size increase in the ancestral population. Similarly, if this population size increase occurred before the divergence of H. himera and H. e. chestertonii from H. erato, this could have contributed to elevated $d_{XY(Z)}/d_{XY(A)}$ in these comparisons (Figure 7). In contrast, the $d_{XY(Z)}/d_{XY(A)}$ ratios among H. melpomene from east of the Andes are closer to the 0.75 (Figure 8). While comparisons between H. melpomene populations east and west of the Andes, Panama and Colombia show deeper divergence, their $d_{XY(Z)}/d_{XY(A)}$ ratios are much lower, consistent with a population size decrease deeper in the ancestry of H. melpomene. Finally, the lower $d_{XY(Z)}/d_{XY(A)}$ ratios in H. cydno–H. timareta comparisons relative to the H. melpomene–H. cydno and H. melpomene–H. timareta comparisons potentially suggest a population contraction of the ancestral population of H. cydno and H. timareta, but after they split from H. melpomene.

Within the H. erato clade, nucleotide diversity as well as Z/A diversity ratios were distinctly higher in populations from east of the Andes and Panama and lower in the H. e. chestertonii and H. himera populations (Figure 7). These populations shared a common ancestor, so differences in nucleotide diversity likely result from population size changes that occurred after divergence and thus confound the relative $F_{ST}$ and $d_a$ divergence measures. Although absolute divergence $d_{XY}$ is clearly higher between H. erato and H. himera or H. e. chestertonii than among H. erato populations east of the Andes (Figure 8), a population size decrease in H. himera and H. e. chestertonii may inflate the $F_{ST}$ and $d_a$ estimates when comparing these populations to geographically abutting H. erato populations (Figure 3). Additionally, any population size changes that occurred before the split of H. himera from H. erato and H. e. chestertonii from H. erato may have affected current $d_{XY}$ estimates. Importantly, if such demographic changes differently affected the ancestor of H. himera as compared to the ancestor of H. e. chestertonii, the $d_{XY}$ values may not necessarily reflect different degrees or stages of the speciation process. This difficulty may also apply when comparing divergence between H. melpomene and H. timareta and between H. melpomene and H. cydno.

### 3.5 Sex-linked incompatibilities increase Z/A absolute divergence ratio

Despite the difficulties in directly comparing divergence on sex chromosomes and autosomes, it may be possible to detect enhanced barriers to migration on sex chromosomes (i.e., reduced effective migration) by comparing population pairs with different levels of absolute migration due to physical isolation, but that otherwise share the same common history. This can be achieved by comparing pairs of populations from the same two species that differ in their extent of geographic isolation. Indeed, previous analyses of sympatric and
allopatric populations of *H. melpomene*, *H. cydno* and *H. timareta*, based on shared derived alleles (i.e., the ABBA-BABA test), found evidence of extensive gene flow between the species in sympatry, but with a strong reduction on the Z chromosome (Martin et al., 2013). Here, we instead use our broad sampling scheme to investigate how patterns of sequence divergence differ with differing levels of geographic separation, and ask whether this signal can detect reduced effective migration on the Z chromosome.

Among *H. erato* and *H. melpomene* clade populations, absolute divergence generally increases with increased distance between population pairs (Figure 8). This trend is strongest for population comparisons that are less obstructed by geographic barriers, such as among *H. erato* (Mantel’s test: \( R^2 = .18; p = .012 \)) and *H. melpomene* (excluding Colombia; Mantel’s test: \( R^2 = .95; p = .001 \)) populations from east of the Andes. As expected, the correlation between distance and absolute divergence is reduced by geographical barriers, such as when comparing *H. erato* (Mantel’s test: \( R^2 = .15; p = .019 \)) and *H. melpomene* (Mantel’s test: \( R^2 = .55; p = .001 \)) populations from Panama, east of the Andes and west of the Andes. We also observed a significant trend of increased absolute divergence with distance between populations of *H. erato–H. e. chestertonii* (Mantel’s test: \( R^2 = .44; p = .001 \)), *H. melpomene–H. cydno* (Mantel’s test: \( R^2 = .55; p = .001 \)) and *H. melpomene–H. timareta* (Mantel’s test: \( R^2 = .75; p = .001 \)).

**FIGURE 7** Relation between Tajima’s D, average nucleotide diversity on the autosomes \( (\pi_A) \) (upper panels) and the ratio of nucleotide diversity between the Z chromosome and autosomes \( (\pi_Z/\pi_A) \) (lower panels) for populations of the *Heliconius erato* and *Heliconius melpomene* clade and simulated data. Points represent average nucleotide diversity measures obtained from autosomes \( (\pi_A) \) and the Z chromosome \( (\pi_Z) \) in 50-kb windows. Grey vertical bars represent 95% confidence intervals estimated from block-jackknifing (note that these are too small to see in the Tajima’s D versus \( \pi_A \) plots). Schematics in the upper right panel represent the simulated population size changes. We simulated a single population that underwent a population size change of magnitude \( x \), ranging from 0.01 to 100 (right panels), with population size change occurring between 1000 and 4,000,000 generations \( (t) \) ago (colours). Triangles indicate population size contractions, and circles indicate population size increase. Population size \( (N_e) \) was 3e6 and 2.25e6 for the autosomes and Z chromosome, respectively. The dashed lines indicate expectations under neutrality [Colour figure can be viewed at wileyonlinelibrary.com]
$R^2 = .69; p = .001$) and $H. melpomene$–$H. timareta$ (Mantel’s test: $R^2 = .56; p = .001$). This is consistent with gene flow among these species pairs where they are in contact. In contrast, no significant trend between absolute divergence and distance was observed between $H. erato$–$H. himera$ and $H. cydno$–$H. timareta$, suggesting that these species pairs may be more strongly isolated.

Simulations show that if distance is considered a proxy for migration, reduced rates of admixture on the Z chromosome may become apparent as increased Z/A absolute divergence ($d_{XY}(Z)/d_{XY}(A)$) ratios over short distances, with the ratio decreasing between pairs that are geographically more isolated (Figure 8). As the effective rate of migration is reduced on the Z chromosome relative to autosomes, the $d_{XY}(Z)/d_{XY}(A)$ ratio increases, and this increase is most pronounced when overall migration rates are high. This relationship can be explained by the absolute difference in effective migration on the Z chromosome compared to the autosomes becoming smaller as overall migration decreases. While overall $d_{XY}(Z)/d_{XY}(A)$ ratios may be influenced by ancestral population size changes, the trend should be independent from population size changes that occurred after the populations split. Our widespread sampling of both Heliconius clades therefore allowed us to test for reduced effective migration on the Z chromosome.

First, we examined the $d_{XY}(Z)/d_{XY}(A)$ ratios and its relationship to geographic distance. If rates of admixture between Heliconius populations are similar on the Z chromosome compared to the
autosomes, we would not expect any relation between distance and \( d_{XY}(Z)/d_{XY}(A) \) ratios. In contrast, we observed increased \( d_{XY}(Z)/d_{XY}(A) \) ratios among geographically more closely located population pairs for *H. melpomene*–*H. timareta* (Manzel’s test: \( R^2 = .25; p = .004 \)) and *H. melpomene*–*H. cydno* (Manzel’s test: \( R^2 = .35; p = .001 \)) comparisons (Figure 8). Similarly, a tendency for increased \( d_{XY}(Z)/d_{XY}(A) \) ratios between *H. erato*–*H. e. chestertoni* was observed among the geographically closest comparisons, although this was not significant (Manzel’s test: \( R^2 = .26; p = .06 \)). Finding this trend, however, can be obscured by geographic barriers that would reduce the relation between distance and admixture. For instance, *H. e. chestertoni* comes into close contact with *H. e. venus* west of the Andes, but is geographically isolated from relatively closely located *H. erato* populations east of the Andes. Similarly, PCA of the *H. melpomene* populations indicates splits between populations east of the Andes, which may reflect additional geographic barriers that do not correlate linearly with distance (Figure 1).

Next, to account for geographic barriers, we also carried out a similar comparison using absolute divergence on the autosomes instead of geographic distance, which might reflect a more direct relationship with migration. Using the absolute divergence on the autosomes as a proxy for gene flow, we find a pattern of increased \( R \) with migration. Using the absolute divergence on the sex chromosomes, resulting from increased mutation rate and that this could explain the observed variation in Z/A diversity ratios among the *Heliconius* populations. Alternatively, in *Heliconius*, male-biased sex ratios have been reported in the field, which could result in increased Z/A diversity ratios. However, it has been argued that these male-biased sex ratios are most likely explained by differences in behaviour, resulting in male-biased captures rather than effective sex ratio differences (Jiggins, 2017). Nonetheless, a *Heliconius* characteristic that could potentially amplify sex ratio biases is that *H. erato* and *H. melpomene* clade populations are characterized by contrasting pupal-mating and adult-mating strategies, respectively (Beltrán, Jiggins, Brower, Bermingham, & Mallet, 2007; Gilbert, 1976). Pupal maters are largely monandrous (females mate only once), whereas adult maters are polyandrous (Walters, Stafford, Hardcastle, & Jiggins, 2012). Such differences in mating system could potentially result in increased variance of male reproductive success and decreased Z/A diversity ratios for monandrous mating systems (Charlesworth, 2001). However, the frequency of remating in polyandrous *Heliconius* species is estimated to be only 25–30% higher than in monandrous species (Walters et al., 2012) and adult mating is likely still prevalent in presumed pupal-mating species (Thurman, Brodie, Evans, & McMillan, 2018). Correspondingly, we did not find any clear difference in Z/A diversity ratios between the pupal-mating *H. erato* and adult-mating *H. melpomene* clade populations (Figure 7). Finally, it has been suggested that effective recombination may be higher for the Z chromosome when recombination is completely absent in females (Charlesworth, 2012). This is because the Z chromosomes spend two-thirds of their time in recombining males, whereas autosomes only spend half of their time in recombining males. This could lead to less of a reduction in diversity on the Z compared to autosomes than the 25% null expectation. However, this cannot explain the correspondence of Tajima’s D and the PSMC inferences and the observed Z/A diversity ratios. Similarly, the pattern of increased Z/A divergence that results from reduced admixture on the Z in population comparisons of geographically closely located *Heliconius* species should not be affected by sex ratio or recombination and mutation biases. Overall, in *Heliconius*, the largest influence on variation in Z/A diversity ratios is likely to be demographic changes.

### 3.7 Consequences for other study systems

Extensive genomic sampling is available for a number of other natural systems that have recently diverged, particularly for birds that also have ZW sex chromosomes, such as flycatchers (Ellegren et al., 2012), crows (Poelstra et al., 2014) and Darwin’s finches (Lamichhaney et al., 2015). In these systems, increased coalescence rates (~lineage sorting) on the Z and/or W chromosome have been accredited to the smaller effective population sizes of the sex chromosome. However, it remains unclear whether elevated measures of divergence could indicate elevated rates of between species divergence on the sex chromosomes, resulting from increased mutation or reduced admixture.

In the adaptive radiation of Darwin’s finches, there is no evidence for Haldane’s rule nor for reduced viability of hybrids due to postmating incompatibilities (Grant & Grant, 1992) and the
maintenance of isolating barriers is best explained as resulting from ecological selection and assortative mating (Grant & Grant, 2008). In crows, the divergence between hooded and carrion crows seems to be solely associated with colour-mediated assortative mating even in the apparent absence of ecological selection (Poelstra et al., 2014; Randler, 2007). The populations of both Darwin’s finches and crows can be characterized by distinct demographic histories (Lamichhaney et al., 2015; Vijay et al., 2016). Therefore, in these species, deviations in divergence measures from neutral expectation on the Z chromosome are potentially also explained by demography. In the divergence of pied and collared flycatchers, species recognition and species-specific male plumage traits are Z-linked (Saether et al., 2007) and female hybrids are completely sterile compared to only low levels of reduced fertility in males (Veen et al., 2001). In agreement with the large-X effect and disjunct rates of admixture between the sex chromosomes and autosomes, genome scans have found signals of increased relative divergence on the Z and W chromosomes (Ellegren et al., 2012; Smeds et al., 2015). The demographic history of these populations is, however, characterized by a severe decrease in population size since their divergence (Nadachowska-Brzyska et al., 2013). In particular, for the W chromosome, the reported excessive decrease in diversity and the high values of relative divergence can thus likely be partly explained by demography (Smeds et al., 2015). However, the excess of rare alleles (negative Tajima’s D) on the W chromosome does contrast with these inferred demographic histories and provides support that the reduced diversity and increased FST measures result from selection (Smeds et al., 2015).

4 | CONCLUSION

The disproportionate role of sex chromosomes during speciation has been well documented based on genetic analysis. However, it is less clear how this influences patterns of divergence in natural populations. In Heliconius, we find much of the observed increased absolute divergence on the Z chromosome relative to neutral expectation can be explained by population size changes. This caution against highlighting increased sex chromosome divergence alone as evidence for a disproportionate role in species incompatibilities or as evidence for faster-X evolution. Although relative measures of divergence are most prone to demographic changes, absolute divergence measures can also be strongly influenced by population size changes. Absolute measures do not therefore provide a solution to the problems inherent in using relative measures to compare patterns of divergence across genomes (Cruickshank & Hahn, 2014). Despite these difficulties, we do find patterns consistent with decreased effective migration on the Z for species pairs with known postzygotic reproductive barriers, in agreement with hybrid sterility and inviability being linked to the Z chromosome in
these cases (Jiggins, Linares, et al., 2001; Naisbit et al., 2002; Salazar et al., 2005; Sánchez et al., 2015). Successfully disentangling the influence of a large-X effect and faster-X evolution on relative rates of divergence will require modelling of the demographic history of each population, including changes that may have occurred before the split of the populations. Such modelling would allow us to better contrast (i) expected within-population Z/A diversity ratios with hypotheses of increased mutation rates, selective sweeps, background selection and mating system and (ii) expected between population Z/A divergence ratios with hypotheses of increased mutation rates or adaptive divergence on the Z chromosome. Additionally, our strategy of contrasting $d_{CV}(Z)/d_{CV}(A)$ ratios with geographic distance provides opportunities for testing reduced admixture between sex chromosomes in systems for which tree-based approaches and/or crossing experiments are unfeasible.

ACKNOWLEDGEMENTS

We thank the Ecuadorian Ministerio del Ambiente (No. 005-13 IC-FAU-DNB/MA), Peruvian Ministerio de Agricultura and Instituto Nacional de Recursos Naturales (201-2013-MINAGRI-DGFFS/ DGEFFSS) and Autoridad Nacional De Licencias Ambientales-ANLA in Colombia (Permiso Marco 0530) for permission to collect butterflies. This work was funded by ERC grant SpeciationGenetics (339873) to CJ, NSF grant (DEB 1257689) to BAC and WOM and NSF EPSCoR RII Track-2 FEC (OIA 1736026) to RP and BAC. SHM was funded by a research fellowship from St John’s College, Cambridge. CS was funded by COLCIENCIAS (Grant FP44842-5-2017). For computational resources, we thank the University of Puerto Rico, the Puerto Rico INBRE Grant P20 GM103475 from the National Institute for General Medical Sciences (NIGMS), a component of the National Institutes of Health (NIH); and awards 1010094 and 1002410 from the EPSCoR program of the NSF. This work was also performed using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service (http://www.hpc.crac.ac.uk/), provided by Dell Inc., using Strategic Research Infrastructure Funding from the Higher Education Funding Council for England and funding from the Science and Technology Facilities Council.

DATA ACCESSIBILITY

Genome assemblies are available on lepbase.org. Sequencing reads are deposited in the Sequence Read Archive (SRA). See Tables S1 and S2 for accession numbers.

AUTHOR CONTRIBUTIONS

S.V.B., M.B., B.A.C., C.D.J. and S.H.M. conceived of the study. S.V.B., M.B. and S.H.M. analyzed the data. S.V.B., M.B., W.O.M., B.A.C., C.D.J. and S.H.M. wrote the paper. R.P. and C.S. provided samples. All authors discussed the results and contributed to the final manuscript.

ORCID

Steven M. Van Belleghem http://orcid.org/0000-0001-9399-1007
Camilo Salazar https://orcid.org/0000-0001-9217-6588
W. Owen McMillan https://orcid.org/0000-0003-2805-2745
Brian A. Counterman https://orcid.org/0000-0003-2724-071X
Chris D. Jiggins https://orcid.org/0000-0002-7809-062X
Simon H. Martin https://orcid.org/0000-0002-0747-7456

REFERENCES

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, 6, 695–708. https://doi.org/10.1534/g3.115.023655


**SUPPORTING INFORMATION**

Additional Supporting Information may be found online in the supporting information tab for this article.