Parallel clines in a quantitative trait in butterfly co-mimics species despite different levels of genomic divergence and selection

Emma V. Curran[1*], Sean Stankowski[1], Carolina Pardo-Diaz[2], Camilo Salazar[2], Mauricio Linares[2], Nicola J. Nadeau[1, 3*]

[1]Department of Animal and Plant Sciences, University of Sheffield, UK;

[2]Biology Program, Faculty of Natural Sciences and Mathematics, Universidad del Rosario, Bogota, Colombia;

[3]The Smithsonian Tropical Research Institute, Panama City, Republic of Panama

*Corresponding authors: e.v.curran@sheffield.ac.uk, n.nadeau@sheffield.ac.uk

Running title: Parallel divergence in a quantitative trait

Abstract word count: 225

Total word count: 7998

Abstract

Hybrid zones, where phenotypically distinct populations meet and interbreed, give insight into how differences between populations are maintained despite gene flow. Divergence in quantitative traits, controlled by multiple loci, may require stronger barriers to gene flow than traits controlled by few, major effect loci. The butterflies *Heliconius erato* and *Heliconius melpomene* are distantly related Müllerian mimics that show parallel divergence in wing colour patterns between geographical races across South America. We investigated intraspecific hybrid zones in which the colour patterns of both species show discrete variation in pigmentation, and quantitative variation in iridescent blue. We tested whether differentiation across hybrid zones persisted due to genome-wide barriers to gene flow maintained by indirect selection, or whether it resulted from direct selection on colour patterns in the face of gene flow. Analysis of phenotypic clines revealed parallel clines in iridescence between species, consistent with direct selection for mimicry, but cline widths were different between species, indicating differences in the strength of selection on iridescence. Genotyping-by-sequencing revealed less defined population structure and weaker genomic differentiation in *H. melpomene*, suggesting the hybrid zone has evolved differently in the two species. In both species, iridescence clines were not concordant with genome-wide ancestry clines, suggesting that they are targets of direct selection. This is the first direct comparison of divergence in quantitative and Mendelian traits in this classic Müllerian mimicry system.

Key words: Hybrid zones, quantitative trait variation, gene flow, selection, cline analysis, Müllerian mimicry

Introduction

Hybrid zones, where genetically differentiated populations are in contact and interbreed, have long been a valuable resource in understanding the evolutionary processes shaping taxonomic boundaries. Hybrid zones can form in continuously distributed populations, where different alleles are favoured at either end of an ecological gradient or transition, a process called primary intergradation. Alternatively, they can form when previously isolated populations, which have become genetically differentiated in allopatry, come into secondary contact (Endler, 1977). Both scenarios can result in clinal variation in quantitative traits and allele frequencies across the hybrid zone, where such clines are maintained by the balance between gene flow following dispersal, and selection (Barton & Hewitt, 1985).

Total selection acting on a divergent locus is composed of direct selection on that locus, plus the indirect selection it experiences from other loci also under divergent selection, depending on the strength of linkage disequilibrium (LD) between it and those other diverged loci (Barton, 1983; Barton & Gale, 1993). When the strength of direct selection outweighs indirect selection, the width of a cline informs about the strength of selection acting on that trait or genetic locus, with cline width being proportional to the strength of selection for a given dispersal parameter (Barton & Hewitt, 1985). Under this scenario the cline centre indicates the point at which the direction of divergent selection changes, so cline coincidence can indicate a common selective agent (Barton & Hewitt, 1985). The shape and position of clines can also be determined by the number of loci involved and the strength of LD between them. If the strength of direct selection on each locus is outweighed by indirect selection from many loci in LD, the total selection affecting such loci will be approximately equal, resulting in them having similar centres and widths, and being steeper at the centre (or “stepped”) where LD is strongest (Kruuk, Baird, Gale, & Barton, 1999; Szymura & Barton, 1991).

Differentiation in phenotypes showing discrete variation across hybrid zones can be maintained by strong divergent selection, despite high levels of gene flow across the rest of the genome (Gompert, Mandeville, & Buerkle, 2017). This has been demonstrated in crows (Poelstra et al., 2014), warblers

(Toews et al., 2016), and monkeyflowers (Stankowski, Sobel, & Streisfeld, 2015). In contrast, studies which examine clines in quantitative traits under divergent selection across hybrid zones, often find that they are accompanied by stepped clines in genetic markers, indicating indirect selection on many different loci across the genome and genetic structure across the hybrid zone, despite some gene flow [e.g. *Larus* gulls (Gay, Crochet, Bell, & Lenormand, 2008) and *Bombina* toads (Szymura & Barton, 1986)], although divergent selection can also maintain differentiation in quantitative traits across hybrid zones without population structure [e.g. floral traits in *Ipomopsis* (Milano, Kenney, & Juenger, 2016)]. Quantitative traits tend to have a more complex, polygenic genetic basis [e.g. human height (Allen et al., 2010)]. Polygenic local adaptation may only require small allele frequency changes, but can also involve greater levels of covariance between loci (Le Corre & Kremer, 2012). The combined action of divergent selection and the build-up of statistical associations between loci can reduce effective migration rates across the genome (Flaxman, Wacholder, Feder, & Nosil, 2014; Kruuk et al., 1999). Therefore, an increased level of overall genome-wide differentiation, and population level genetic structure may be expected across hybrid zones over which quantitative variation is maintained.

Here, we studied the clinal variation of two colour pattern traits in South and Central American *Heliconius* butterflies: the yellow hindwing bar (Mendelian) and iridescence (quantitative). *Heliconius erato* and *Heliconius melpomene* are Müllerian co-mimics with bright aposematic wing colour patterns. Where they co-occur, they converge on almost identical wing colour patterns to share the cost of educating predators of their distastefulness (Brown, 1981). Both species comprise many parapatric colour pattern races, or subspecies, connected by hybrid zones (Mallet, 1993; Rosser, Dasmahapatra, & Mallet, 2014). When members of different subspecies hybridise, their offspring can display novel or heterozygous phenotypes (Arias et al., 2008; Mallet, 1989; Mallet, 1986). Predators are less likely to learn to avoid rare phenotypes, causing frequency-dependent selection on colour patterns (Langham, 2004; Mallet & Barton, 1989). This maintains stable hybrid zones (Mallet, 1986; Rosser et al., 2014). The diverse colouration seen in the *Heliconius* genus has been extensively studied, the vast majority of which is determined by a genetic ‘tool kit’ of around five major-effect

loci (Baxter, Johnston, & Jiggins, 2009; Kronforst et al., 2006; Martin et al., 2012; Nadeau, 2016; Nadeau et al., 2016; Reed et al., 2011; Westerman et al., 2018). Previous studies have found low levels of genetic differentiation between parapatric colour races, with a few diverged loci which mainly control colour pattern differences (Martin et al., 2013; Nadeau et al., 2014, 2012; Supple et al., 2013).

Near the Panamanian-Colombian border, there are co-occurring hybrid zones between subspecies of *H. erato* and *H. melpomene*, which differ in the presence of a yellow hindwing bar and in iridescent blue colouration (Mallet, 1986; Figure 1). The iridescence is produced by nano-structural ridges on the surface of wing scales, which are layered to produce constructive interference of blue light (Parnell et al., 2018). In a system so well-studied, little is known about how selection acts on structural colour (Sweeney, Jiggins, & Johnsen, 2003) and divergence in this trait as not been previously studied. While the yellow bar is controlled by a single major-effect gene (Mallet, 1986; Nadeau, 2016), iridescence segregates as continuous variation, with conservative estimates suggesting it is controlled by around five additive genetic loci (Brien et al., 2018).

Here, we use geographic cline analysis to test whether the colour pattern clines are maintained by direct or indirect selection, and what we can learn about the selection regimes impacting variation and convergence of iridescence. By obtaining genome-wide SNP data we can question the extent to which genome-wide barriers to gene flow may maintain the iridescence clines via indirect selection, or whether differentiation is maintained by direct selection in the face of gene flow, as is seen in other instances of colour pattern divergence in these species (Nadeau et al., 2014). This is also an opportunity to examine how the different genetic architectures of the traits influences the properties of the clines.

Methods

Butterfly specimens

Wild *Heliconius melpomene* and *Heliconius erato* individuals were collected from several sites in the Chocó rainforest corridor between the Andes and the Pacific in Ecuador and Colombia, and part way across the isthmus of Panama (Figure 1, SI Table S1). Wings were removed and stored inside envelopes. Bodies were preserved in NaCl saturated 20% dimethyl sulfoxide (DMSO) 0.25M EDTA.

Phenotypic measurements

Digital images of butterfly wings were taken with a Nikon D7000 DSLR camera fitted with an AF-S DX Micro NIKKOR 40 mm f/2.8G lens (Nikon UK Ltd., Surrey, UK), mounted on an adjustable platform. Standardised lighting conditions were achieved using two external natural daylight fluorescent lights, mounted to illuminate at 45 degrees from incident, to maximise brightness of observed iridescent colour. Photographs were taken with a shutter speed of 1/60 sec and an aperture of f/10. Each sample was photographed with an X-Rite colorchecker passport (X-Rite, Inc., MI, USA) in shot. The Nikon raw (.NEF) image files were converted to standard raw files (.DNG) using Adobe DNG converter (Adobe Systems Inc., USA). The RGB channels in the images were then linearized using the neutral grey scale on the colorchecker using GNU Image Manipulation Program, v2.8.

The mean RGB values from regions in the discal cell on the right forewing and the Cu_2 cell on the right hindwing were measured (SI Figure S1). If the wings on the right-hand side showed damage, wings on the left-hand side were used. Wing regions were selected using the polygon selection tool in ImageJ, version 1.50b (Abràmoff, Magalhães, & Ram, 2004), and mean RGB scores were measured using the Color Histogram plugin. To minimise variation in blue colour due to age and wing wear, we excluded individuals with extensive wing wear or damage.

We tested for repeatability (Whitlock & Schluter, 2009) of the RGB values on 26 individuals photographed a second time under the same conditions on a different day, with a second set of RGB

measurements taken. These individuals were selected from regions in which varying levels of iridescence is seen (20 individuals from Valle del Cauca, Colombia, and 6 individuals from Darién, Panama). Variance among individuals was calculated by taking the difference between the group mean square and the error mean square, and dividing it by the number of replicates. These components of variance were extracted from a general linear model in R v3.2.3 (R Core Team, 2015). The fraction of total variance that is due to true differences between individuals was then calculated by dividing the variance among individuals by the total variance.

A measure of relative blue reflectance (blue score) was determined for each individual by taking the mean blue channel value (B) and the mean red channel value (R) for both wing regions and calculating:

$$BR = (B-R)/(B+R)$$

This gives a standardised score of how blue an individual is, with BR = 1 being the 'bluest', and BR = -1 being the 'reddest'.

Estimation of 'yellow bar' allele frequencies

Allele frequencies for the yellow hindwing bar were estimated based on phenotype for both species. This was done for all sampling sites in Colombia and Panama with five or more individuals. The 'yellow bar' phenotype was scored categorically according to Mallet (1986), who showed that this phenotype segregates in the same way for both *Heliconius erato* and *H. melpomene*. Variation in the yellow bar across this hybrid zone is controlled by three alleles: The North Colombian yellow bar allele (Y), the West Colombian yellow bar allele (y_{wc}) and the Central American yellow bar allele (y_{ca}). Individuals of both species with a yellow bar on both sides of the wing (Figure 1A) have genotype $y_{ca}y_{ca}$. Individuals lacking a yellow bar (Figure 1B) have genotype YY. Individuals with the "shadow bar" phenotype, where the outline of the bar can be seen on the underside of the hindwing without any yellow pigment, and without a bar on the upper side of the hindwing, have genotype Yy_{wc} or Yy_{ca}.

Individuals with a yellow bar on the underside of the hindwing (Figure 1C) have genotype $y_{wc}y_{ca}$ or $y_{wc}y_{wc}$. As two of the four phenotypes can be produced by two different allele combinations we inferred the allele frequencies at each locality for each species assuming Hardy-Weinberg equilibrium for the three alleles. The frequency of Y could be directly observed from both its heterozygous and homozygous phenotypes. The frequency of y_{ca} could be inferred from the frequency of its homozygous phenotype, allowing us to infer the frequency of y_{wc}.

Phenotypic cline analysis

We modelled changes in frequency of the West Colombian yellow bar allele (y_{wc}) and level of iridescence (BR) in *Heliconius erato* and *Heliconius melpomene* along the transect by fitting a number of geographic cline models (Szymura & Barton, 1991; Szymura & Barton, 1986) implemented using the software ANALYSE v1.30 (Barton & Baird, 2002). Sampling sites with fewer than five individuals were excluded, leaving 529 *H. erato* and 126 *H. melpomene*. Blue scores were rescaled between 0 and 1 by dividing the difference between the individual's blue score (BR_i) and the species minimum, by the blue score range for the species:

$$BR_{norm} = (BR_i - BR_{min})/(BR_{max} - BR_{min})$$

Distances between sampling sites were estimated using the great circle distance, which is the shortest distance between two points on the surface of a sphere. This was calculated using the 'hzar.map.greatCircleDistance' function in the R package HZAR (Derryberry, Derryberry, Maley, & Brumfield, 2013).

ANALYSE fits cline models to marker loci and/or quantitative trait data, and can be used to compare the fit of three alternative cline models. The simplest model is a sigmoid cline described by a four-parameter hyperbolic tangent,

$$p(z) = (1 + \tanh[2(x_i - c)/w])/2$$

Where p(z) is a gene frequency (or mean score of a trait) at position x_i, c is the cline centre and w is the cline width, defined as the ratio between the total change in the frequency of an allele (Δp) or value of the trait (Δz) across the cline and gradient at the cline centre:

$$\Delta p(z)/(\partial z/\partial x)$$

The other two more complex models are 'stepped' clines. They consist of a central sigmoid step flanked by two exponential tails:

$$z = \propto \exp(-4x\sqrt{\theta}/w)$$

The tails describe the pattern of introgression from the centre into the foreign genepool; θ is the rate of decay, and the strength of the barrier to gene flow, B, can be estimated as the ratio between the difference in the allele frequency and the initial gradient in gene frequency with distance x at the edges of the central segment (Szymura & Barton, 1986). In the symmetrical 'Sstep' model, θ and B are equal on both sides. In the asymmetrical 'Astep' model, the pattern of introgression is different on the left and right side.

ANALYSE uses the Metropolis algorithm to search the likelihood surface to find the optimal solution to the model. To ensure that the likelihood surface was thoroughly explored, independent runs were conducted using a range of initial parameter estimates. After obtaining maximum likelihood (ML) solutions for the three cline models, the most likely model was identified using Likelihood Ratio Tests. As the minimum and maximum mean allele frequency or trait values ($p(z)_{min}$, $p(z)_{max}$) were allowed to vary in the tails of the cline, the sigmoid, Sstep and Astep models were described by 2 (c, w), 4 (c, w, θ, B) and 6 parameters (c, w, θ_0, θ_1, B_0, B_1), respectively.

After model selection, support limits were estimated for each parameter in the ML model. Starting with the optimum fit, and constraining the values of all other parameters, the likelihood surface for individual parameters were explored by making 10,000 random changes of their value. The range of estimates that was within 2 log-likelihood units of the maximum estimate was taken as the support limit for that parameter, and is approximately equivalent to a 95% confidence interval.

Coincidence of cline centres (*c*) and concordance of cline widths (*w*) were tested using the composite likelihood method (Kawakami, Butlin, Adams, Paull, & Cooper, 2009; Phillips, Baird, & Moritz, 2004). The method involves obtaining a composite ML score for a given parameter (ML_{comp}) and comparing it with the sum of the ML estimates obtained for each profile (ML_{sum}). ML_{comp} was obtained by constructing a log-likelihood profile (10 km intervals for *c* and *w*, between 0 km and 1000 km) with all other parameters allowed to vary, summing the profiles, and obtaining the ML estimate; ML_{sum} was obtained by summing the ML estimates from each profile. If clines are not coincident or concordant, ML_{sum} is significantly smaller than ML_{comp}, as determined by a chi-squared test ($\alpha = 0.05$) with n-1 degrees of freedom, where n is the number of traits. One complication with this method for comparing cline parameters is that the profiles for each trait must be built using the same model. Although the more complex Sstep and Astep models are a significantly better fit than the sigmoid model, the parameters estimates for the cline centre and cline width were highly similar regardless of the model fit (see results). Therefore, all likelihood profiling was conducted with the sigmoid model.

To estimate the strength of selection acting on y_{wc}, the following equation was used from Barton and Gale (1993):

$$s^* = (1.782\sigma/\omega)^2$$

Where s^* is the difference in mean fitness between populations at the edge of the cline, and populations at the centre. This demonstrates the mean strength of effective selection on loci underlying a trait required to maintain a cline of width (*w*), given the dispersal distance per generation (σ). Dispersal estimates were taken from Mallet et al. (1990) and Blum (2002).

Sequencing data

Restriction-associated DNA (RAD) sequence data were generated for 265 *H. erato* (SI Table S2), and whole genome re-sequencing was carried out on 36 *H. melpomene* individuals (SI Table S3). Genomic

DNA was extracted from each individual using DNeasy Blood and Tissue Kits (Qiagen). Library preparation and sequencing was carried out by Edinburgh Genomics (University of Edinburgh).

Single-digest RAD libraries were prepared using *Pst1* restriction enzyme, with eight base-pair barcodes and sequenced on the Illumina HiSeq 2500 platform (v4 chemistry), generating an average of 554,826 125 base paired-end reads per individual (see SI Table S2 for coverage across each individual). We demultiplexed the pooled reads using the RADpools program in the RADtools package version 1.2.4 (Baxter et al., 2011).

TruSeq Nano, gel-free libraries were prepared from genomic DNA samples of 36 *H. melpomene* individuals and sequenced on Illumina HiSeq 2500 platform (v4 chemistry), generating an average of 31,484,363 125 base paired-end reads per individual (see SI Table S3 for coverage across each individual).

We checked the quality of all the raw sequencing reads using FastQC (v 0.11.5) and removed any remaining adapters using Trim Galore (v 0.4.1). We aligned the sequence data of all individuals, both RAD sequenced and WGS, to their corresponding reference genomes, either *Heliconius melpomene* version 2 (Davey et al., 2016) or *Heliconius erato* (Van Belleghem et al., 2017), obtained from lepbase (Challis, Kumar, Dasmahapatra, Jiggins, & Blaxter, 2016), using bowtie2 (v 2.3.2), with the local alignment option, and the very-sensitive pre-set parameter options to improve accuracy of the alignment. We used samtools (v 1.3.1) to sort and index the alignment files. We removed any duplicates that may have arisen during library preparation using the MarkDuplicates program in Picard tools (v 1.92).

Population structure

SNP datasets were generated using samtools mpileup (v 1.5) to compute genotype likelihoods and bcftools (v 1.5) for variant calling. For a site to be a variant, the probability that it was homozygous for the reference allele across all samples was required to be less than 0.05. Multiallelic sites, insertions

and deletions were ignored. For *H. melpomene* we identified 30,027,707 single nucleotide polymorphisms (SNPs), and for *H. erato* we identified 5,088,449 SNPs. SNPs were filtered out that had a phred quality score lower than 30, that lacked sequence data in 50% or more of the individuals, that had a minor allele frequency lower than 0.05 and that were private variants. We pruned based on linkage disequilibrium, discarding SNPs within a 20kb window with $r^2 > 0.8$, using the bcftools plugin '+prune'. This reduced the initial number of called SNPs down to 9,336,937 in *H. melpomene* and 159,405 in *H. erato*.

To examine population structure, we first estimated the ancestry of each individual using the software NGSadmix (Skotte, Korneliussen, & Albrechtsen, 2013), which estimates the proportion of each genome that can be attributed to predefined number of genetic clusters (k) using genotype likelihoods. For each species, NGSadmix was run for a range of values of k, one to ten, each being replicated ten times with a random seed. The value of k which best describes the population structure of each individual was determined using the Δk criterion (Evanno, Regnaut, & Goudet, 2005), implemented in CLUMPAK (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015).

We carried out a principal components analysis (PCA) using PCAngsd (Meisner & Albrechtsen, 2018), which estimates a covariance matrix based on genotype likelihoods. We used eigenvector decomposition to retrieve the principal components of genetic structure.

Population differentiation

To test the extent of genetic differentiation between the iridescent and non-iridescent subspecies, we first measured F_{ST} between all the individuals from the iridescent populations south of the hybrid zone, and all the non-iridescent individuals north of the hybrid zone, excluding the sampling site Jaqué, which was in the centre of the hybrid zone in both species. SNP datasets were generated for each species, using samtools mpileup and bcftools (v 1.5). In each species Hudson's F_{ST} estimator was calculated among populations (Hudson, Slatkin, & Maddison, 1992):

$$F_{ST}^{Hudson} = 1 - \frac{Hw}{Hb} = \frac{p_1(1-p_1) + p_2(1-p_2)}{p_1(1-p_2) + p_2(1-p_1)}$$

Where *Hw* is the within-population heterozygosity, *Hb* is the between-population heterozygosity, and p_1 and p_2 represent the allele frequencies in each population. This was calculated in *R* for every SNP with a custom script. Average genome-wide F_{ST} was calculated as a ratio of averages, by averaging the variance components, *Hw* and *Hb*, separately, as recommended by Bhatia *et al*. (2013). We also estimated average genome-wide F_{ST} between all pairs of populations, including those in the hybrid zone, for each species, and plotted pairwise F_{ST} against pairwise geographic distance.

Comparison of phenotypic and genomic clines

If iridescence varies across the hybrid zone in a similar way to the background genetic structure, this may suggest that variation is a result of neutral diffusion across a zone of secondary contact. If iridescence is under divergent selection, we would expect genes which contribute to this trait to introgress less between the iridescent and non-iridescent populations than neutral alleles, and the iridescence cline to be narrower than a cline in average ancestry [e.g. (Scordato et al., 2017)]. We compared the shape and position of clines in iridescence with clines in genetic structure across the hybrid zones.

To fit a cline in genetic structure, we used ancestry proportions estimated by NGSadmix for *K*=2, and calculated mean admixture proportions for each site across the hybrid zone in each species. We then fit geographic cline models to variation in the admixture proportion across sampled populations using ANALYSE, and tested whether iridescence and admixture clines had coincident centres and concordant widths, using the likelihood profiling approach described above.

Results

Phenotypic Variation

Strong phenotypic variation was observed across our range of sampling sites, with some difference apparent between *H. erato* and *H. melpomene* (Figure 2). The West Colombian yellow bar allele (y_{wc}) was fixed in all Colombian sampling sites, apart from at some of the northernmost Colombian sampling sites near Bahía Solano (Figure 2 A,B). In *H. melpomene*, the frequency of y_{wc} gradually decreased, and persisted at comparable frequencies to the North Colombian yellow bar allele (Y) for ~200 km, before the Central American yellow bar allele (y_{ca}) became predominant (Figure 2D). In contrast, in *H. erato* Y became the predominant allele, with y_{ca} approaching fixation towards the end of the transect (Figure 2C).

In both species the blue score, used as a proxy measure for iridescence, decreased across the transect (Figure 2 A,B). The colour measurements used to calculate the blue score were highly repeatable ($p<0.001$ for both red and blue values in both wing patches measured, Table S5). The bluest *H. melpomene* individuals were less blue than the bluest *H. erato* (Figure 2), which is consistent with reflectance spectrometry data from *H. erato cyrbia* and *H. melpomene cythera* (Parnell et al., 2018)

Geographic clines in colour patterns

To compare the selection regime acting on iridescence and the yellow bar in *H. erato* and *H. melpomene*, clines were modelled on variation in the blue score (BR) (Figure 3A, B), and the frequency of the y_{wc} allele (Figure 3C, D) across the transect. The cline fitting revealed that an asymmetrical stepped cline best described the variation in iridescence in *H. erato*, with a steeper right tail extending north into Panama. Neither stepped model was a significantly better fit than sigmoidal clines for the yellow bar in *H. erato*, and both colour traits in *H. melpomene* (Table 1; SI Table S6).

Likelihood profiling revealed that both iridescence and the yellow bar clines were coincident (i.e. the same cline centre) between species, however the iridescence cline was significantly wider in *H. melpomene* (Table 2). The yellow bar widths were not significantly different between species (Table 2), despite non-overlapping support limits (Table 1), possibly due to gaps in sampling. Within each species, clines in both colour traits were coincident and concordant (Table 2).

We estimated effective selection (s*) on y_{wc} across the hybrid zone in both species using the ML estimates and two log-likelihood support limits of cline width (Table 1), and dispersal estimates from Mallet et al. (1990), 2.6 km for *H. erato* and 3.7 km for *H. melpomene.* For *H. erato* s*=0.00203 (0.00102–0.00427), and for *H. melpomene* s*=0.000213 (0.000165–0.000303). Blum (2002) estimates higher dispersal for *H. erato*, 10 km, which increases the value of s* to 0.0300 (0.0151–0.0632). Given that the widths of the yellow bar and iridescence clines were not different within each species, similar estimates are found for iridescence. For *H. melpomene*, s*=0.000200 (0.000135–0.000300), for *H. erato*, s*= 0.00208 (0.00167–0.00267) if the dispersal distance is 2.6 km, and 0.0307 (0.0247–0.0395) if the dispersal distance is 10 km. However, it should be noted that in the case of iridescence, s* is the average strength of selection acting across loci controlling iridescence.

Population Structure

We investigated population structure using genome-wide SNP data in the programs NGSadmix, to estimate ancestry proportions from a varying number of genetic clusters (K), and PCAngsd to confirm population clustering by principal components (PCA). This revealed different patterns of population structure in the co-mimics. In *H. erato,* NGSadmix supported K=2 (SI Figure S2B), representing a "Panama-like" and a "Colombia-like" genetic background (Figure 4B), with individuals of consistently mixed ancestry found in the site closest to the centre of the iridescence cline. Introgression from Panamanian populations could be detected in northern Colombian populations. The PCA supported this, with three clusters separated by geography along the first axis of variation, representing the Colombian populations, the Panamanian populations, and individuals with mixed

ancestry and intermediate levels of iridescence clustered between them (Figure 4C). PC1 explained 5.84% of genetic variation in *H. erato*, with all subsequent eigenvectors explaining 0.7% or less of the genetic variation (SI Figure S3).

NGSadmix also supported K=2 for *H. melpomene* (although K=1 cannot be tested, SI Figure S2D), but revealed a less straightforward population structure. While a "Colombia-like" genetic background could be seen, Panamanian individuals showed mixed ancestry, with the exception of two individuals from the site closest to the centre of the iridescence cline (Figure 4D). This is supported by the PCA. PC1 explained 5.28% of genetic variation, separating Colombian and Panamanian individuals. Individuals with intermediate levels of iridescence do not form an intermediate cluster between Panamanian and Colombian individuals, as is seen in *H. erato* (Figure 4E). The percent of genetic variation explained by PC1 and subsequent principal components show a more uniform distribution than in *H. erato* (SI Figure S3) consistent with weaker population structure.

Using SNPs from individuals either side of the hybrid zone, genome-wide average Hudson's F_{ST} was estimated for each species, using the ratio of averages approach. This revealed that genome-wide divergence across the hybrid zone is greater in *H. erato* (F_{ST}=0.188), compared to *H. melpomene* (F_{ST}=0.0739). The difference in genetic structure is also apparent in the plots of the pairwise genetic distance between sampling locations, plotted against their geographic distance. In *H. erato*, within-race comparisons that span distances of 195 – 325 km show a range of F_{ST} values between 0.063 – 0.129. However, between-race comparisons made over a similar range of distances (188 – 345 km) have substantially higher F_{ST} (0.226 – 0.271), suggesting that the genetic structure is much stronger than would be expected based on geography alone (Figure 5). The pattern in *H. melpomene* is very different, as the between-race comparisons span a similar range of F_{ST} values to the within-race comparisons.

Clines in genomic admixture proportion

To test whether variation in iridescence was independent from genetic structure across the hybrid zone, we compared clines in iridescence and admixture proportion across the transect (Figure 3E, F). An asymmetrical stepped cline model best fits the variation in admixture proportions across the transect for *H. erato* (Table 1, SI Table S6), with a steeper right tail, similar to the iridescence cline. However, the admixture proportion cline is not coincident with the cline in iridescence in *H. erato* (Table 2), with the centre of the iridescence cline estimated to be located significantly further North (Table 2). The admixture proportion cline is also significantly wider than the iridescence cline in *H. erato.*

Neither stepped cline model is a significantly better fit than a sigmoidal model for variation in admixture proportions in *H. melpomene* (Table 1). The clines in iridescence and admixture proportion are coincident, but not concordant in *H. melpomene*, with a wider iridescence cline (Table 2). However, the likelihood profiles reveal a wide range of values for admixture cline centres and widths with similar likelihoods (SI Figure S4), suggesting the cline model fits poorly to the data. This may be due to gaps in sampling.

Discussion

Here, we investigate hybrid zones between different colour pattern races of the Müllerian mimics *H. erato* and *H. melpomene*, across which both species display continuous variation in iridescent structural colour. Using comparisons of phenotypic clines, and population genomic approaches, we investigate the selection regimes acting on iridescence, and the genetic structure of the hybrid zones.

Phenotypic clines and selection

Our geographic cline analysis of colour pattern variation in *H. erato* and *H. melpomene* revealed that clines in quantitative iridescence and discrete yellow bar traits were not only coincident within both species, but all colour patterns were coincident between species. Physical linkage is unlikely to

play a role in the coincidence of the traits in *H. erato*, as crosses between iridescent and non-iridescent subspecies of *H. erato* show independent segregation of iridescence and yellow bar phenotypes (Brien et al., 2018).

Under a scenario where the colour pattern clines are maintained by a balance between migration and direct selection, shared cline centres would be expected if both traits experience a similar strength of selection, and either the same selective pressure, or any other which changes in approximately the same location (Barton & Hewitt, 1985). Local warning colour patterns in *Heliconius* are maintained by predator-mediated positive frequency-dependent selection, with rare colour morphs experiencing increased predation (Benson, 1972; Dell'aglio, Stevens, & Jiggins, 2016; Langham, 2004; Mallet & Barton, 1989). The centre of colour pattern clines represents the location where the most effective warning pattern shifts to that of a neighbouring subspecies. The coincidence of cline centres for iridescence and y_{wc}, which is observed in both *H. melpomene* and *H. erato*, is consistent with both traits being involved in a warning signal. The coincidence of iridescence clines between species supports a common source of selection, and a role for aposematism and mimicry in the evolution of iridescence.

The shape of clines can provide insight into the balance of direct versus indirect selection acting on traits (Vines et al., 2016). Concordant, stepped clines (steep clines with flanking tails) indicate that indirect selection is dominant (Kruuk et al., 1999). In *H. erato*, the two colour pattern clines are coincident, and the best fitting cline model for variation in iridescence is stepped, however the stepped cline models are not a significantly better fit than a simple sigmoidal cline for the yellow bar, lending less support to indirect selection. In addition, while the best fitting cline model for the admixture proportions is stepped, its centre is significantly different from the colour pattern clines, which again suggests against indirect selection being the dominant force. Finally, while we do see stepped clines in *H. erato*, they are both asymmetrical clines, with a very shallow left tail, and a more prominent right tail. It is therefore more likely that these tails reflect genuine asymmetry, due to hybrid zone movement, which has been predicted and documented in these species (Blum, 2002; Mallet, 1986;

Thurman, Szejner-Sigal, & McMillan, 2019), or asymmetric selection, rather than being due to indirect selection.

In *H. melpomene* the colour pattern clines are estimated to be more than four times wider than the corresponding traits in *H. erato*, and are not characteristic of the steep, stepped clines which result from strong LD and indirect selection (e.g. Szymura & Barton, 1991). Wider clines could be a result of greater dispersal distances. While dispersal distances have not been directly measured in *Heliconius* butterflies, estimates have been made (Blum, 2002; Mallet et al., 1990), and the only study which compares the two species reports higher dispersal distances in *H. melpomene* (Mallet et al., 1990). However our estimates of the selection coefficient s^* (Barton & Gale, 1993) show that even with higher dispersal, colour pattern traits in *H. melpomene* appear to be under much weaker selection. Other studies on parallel hybrid zones between neighbouring races in this species pair show that *H. melpomene* tend to have wider clines than *H. erato*, but not to the degree seen in the present study (Mallet et al., 1990; Salazar, 2012). *H. melpomene* displays less vivid iridescence than its co-mimic, and the colour difference between iridescent and non-iridescent *H. melpomene* is less pronounced than the colour difference between *H. erato* races (Parnell et al., 2018, Figure 1). Hybrid phenotypes are therefore less distinct from the parental populations in *H. melpomene* which could result in weaker selection against hybrid offspring.

Genetic structure and population history

Overall, we found more defined population structure in *H. erato* compared to *H. melpomene.* In *H. erato*, the PCA and estimated admixture proportion support two main genetic clusters, one representing the non-iridescent Panamanian populations, the other representing the iridescent Colombian populations. Hybrid zone individuals are consistently of mixed ancestry, indicating admixture between the two genetic clusters. In addition, pairwise F_{ST} in between-subspecies population pairs is consistently higher than within-subspecies population pairs, even when accounting for geographic distance. This contrasts with Peruvian and Ecuadorian hybrid zones between different

H. erato races, which comprise of single genetic clusters, indicating hybrid zone formation via primary intergradation, or ancient secondary contact (Nadeau et al., 2014). It is possible that the *H. erato* hybrid zone in the present study also formed via primary intergradation, with genetic structure forming as a result of reduced effective migration across the genome due to the combined effects of divergent selection and the build-up of statistical associations between loci (Flaxman et al., 2014). Alternatively, the two ancestral groups may have diverged during isolation, with gene flow resuming upon secondary contact, causing an abrupt discontinuity between genetically differentiated populations (Barton, 1983). When several traits under selection change simultaneously across a contact zone, linkage disequilibrium (LD) can be maintained between them, and indirect selection can therefore predominate (Szymura & Barton, 1986). However, low genome-wide F_{ST} (0.188), and traces of introgression of the Panamanian genetic cluster into Colombian populations (from the admixture proportions), and vice versa, suggests ongoing gene-flow, which would likely break down LD, unless secondary contact was relatively recent.

There is generally weaker population structure in *H. melpomene*. This could be due to greater levels of dispersal, weaker selection, and/or a primary intergradation scenario. Secondary contact cannot be ruled out, since there is a genetic discontinuity near the centre of the clines, but if so these populations were less divergent than the *H. erato* populations, as there is very low genome-wide F_{ST} (0.0739) between populations either side of the hybrid zone. In addition, the admixture proportions do not clearly define a Panamanian genetic background, as seen in *H. erato*. Our methods do not allow us to explicitly test for an absence of genetic structure. The method used to determine the number of genetic clusters (ΔK) is a based on the plateauing of explanatory power with increasing numbers of genetic clusters (Evanno et al., 2005; Janes et al., 2017), and therefore does not allow us to distinguish between K=1 and K=2. However, pairwise F_{ST} is not greater in between-subspecies comparisons compared to within-subspecies comparisons, supporting lack of differentiation across the hybrid zone. It must be noted, however, that we had fewer *H. melpomene* samples, and fewer sampling locations than for *H. erato*, which may have impacted some of these results.

Genetic differentiation and divergence in a quantitative trait

While the colour pattern traits studied here likely have different genetic architectures, with the yellow bar being controlled by a single major-effect, Mendelian locus (Joron et al., 2006; Mallet, 1986; Nadeau et al., 2016), and iridescence, a quantitative trait, likely being controlled by multiple genes (Brien et al., 2018), phenotypic clines in these traits show close concordance and coincidence within each species. Quantitative trait divergence may require a more complex pattern of genetic differentiation than traits with a Mendelian pattern of inheritance, so reduced gene flow could have facilitated the evolution of iridescence. However, within both species we find evidence that the clines in iridescence are not tightly coupled to clines in genome-wide ancestry. Clines in morphological traits can be uncoupled from background genetic clines when direct natural (Vines et al., 2016), or sexual (Baldassarre, White, Karubian, & Webster, 2014) selection outweighs the effects of indirect selection from other loci in LD. In *H. erato*, the iridescence and admixture proportion clines have significantly different centres. This is particularly striking given the similarities in the phenotypic clines within this species, and suggests that coincidence of colour pattern clines is not simply due to a pattern of genome-wide differentiation, but more likely due to common selective forces maintaining closely correlated clines. Spatial variation in quantitative traits can also be non-adaptive, arising via purely stochastic processes (Storz, 2002). However, the iridescence cline is significantly narrower than the admixture proportion cline suggesting reduced introgression of genes controlling iridescence relative to the genome-wide average, likely due to selection on colour. It is therefore unlikely that the iridescence cline in *H. erato* is a result of neutral diffusion following secondary contact.

In *H. melpomene*, we find low overall genetic differentiation. A cline model of admixture proportions based on two genetic clusters has a centre and width not significantly different to the colour pattern clines. However, this is likely due to the poor fit of the cline models to variation in admixture proportions (Figure 4; SI Figure S4), illustrated by the large confidence intervals (Table 1). This is possibly due to gaps in sampling, but the poor fit can also be explained by the less defined

population structure in this species. Also, clear phenotypic intermediates in the hybrid zone are not predicted to be of mixed ancestry (Figures 3, 4). This suggests that divergence in iridescence in this species is independent of any extensive genomic differentiation and that the broad phenotypic clines that we see are due to weak selection over large distances.

Conclusions

This unique system allowed us to compare genetic and phenotypic patterns of divergence in convergent traits that show both discrete and quantitative variation. Despite different patterns of population structure and genomic divergence, the co-mimics *H. erato* and *H. melpomene* have formed parallel clines in iridescence, and the evidence appears to support direct selection plays a role in maintaining these clines. This is likely due to iridescence playing a role in the mimetic warning signal, despite the weaker selection which appears to be operating in *H. melpomene.*

While pigmentation has a long history in our understanding of the link between genotype and phenotype, we have barely scratched the surface of our understanding of the genetic control of structural colour. This natural system is a promising candidate for association or admixture mapping, and could allow us to identify genomic regions controlling structural colour for the first time.

Acknowledgements

Thanks to the governments of Panama and Colombia (ANLA- Permit 0530) for giving permission to collect butterfly specimens. Thanks to the McMillan and Jiggins labs for providing access to samples. Thanks also to Patricio Salazar, Juan Enciso, Juan Camilo Dumar, Melanie Brien, Carlos Arias, Agata Surma and others in Panama and Colombia, in particular the residents of Jaqué, Darién, for help with logistics and collecting in the field. Thanks to Roger Butlin for valuable comments on this manuscript. This work was funded by the UK Natural Environment Research Council (NERC) through an Independent Research Fellowship (NE/K008498/1) to NJN, and by The Royal Society through an

International Exchange Scheme grant. EVC was funded by the NERC doctoral training partnership, ACCE. CS and CP were funded by COLCIENCIAS (Grant FP44842-5-2017).

References

Abràmoff, M. D., Magalhães, P. J., & Ram, S. J. (2004). Image processing with imageJ. *Biophotonics International*, *11*(7), 36–41. https://doi.org/10.1117/1.3589100

Arias, C. F., Muñoz, A. G., Jiggins, C. D., Mavárez, J., Bermingham, E., & Linares, M. (2008). A hybrid zone provides evidence for incipient ecological speciation in Heliconius butterflies. *Molecular Ecology*, *17*(21), 4699–4712. https://doi.org/10.1111/j.1365-294X.2008.03934.x

Barton, N. H. (1983). Multilocus Clines. *Evolution*, *37*(3), 454–471. https://doi.org/10.2307/2408260

Barton, N. H, & Hewitt, G. M. (1985). Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics*, Vol. 16, pp. 113–148. https://doi.org/10.1146/annurev.es.16.110185.000553

Barton, N. H., & Baird, S. J. E. (2002). *Analyse. Version 1.30*. Retrieved from http://archive.bio.ed.ac.uk/software/Mac/Analyse/index.html

Barton, N. H., & Gale, K. S. (1993). Genetic analysis of hybrid zones. In R. G. Harrison (Ed.), *Hybrid zones and the evolutionary process* (pp. 13–45). Retrieved from https://books.google.co.uk/books?id=aFJFkVKskYIC&pg=PA13&source=gbs_toc_r&cad=4#v=onepage&q&f=false

Baxter, S W, Johnston, S. E., & Jiggins, C. D. (2009). Butterfly speciation and the distribution of gene effect sizes fixed during adaptation. *Heredity*, *102*, 57–65. https://doi.org/10.1038/hdy.2008.109

Baxter, S. W., Davey, J. W., Johnston, J. S., Shelton, A. M., Heckel, D. G., Jiggins, C. D., & Blaxter, M. L. (2011). Linkage Mapping and Comparative Genomics Using Next-Generation RAD Sequencing of a Non-Model Organism. *PLoS ONE*, *6*(4), e19315. https://doi.org/10.1371/journal.pone.0019315

Benson, W. (1972). Natural Selection for Mullerian Mimicry in Heliconius erato in Costa Rica. *Science*, *176*(4037), 936–939. Retrieved from http://www.sciencemag.org/content/176/4037/936.short

Bhatia, G., Patterson, N., Sankararaman, S., & Price, A. L. (2013). Estimating and interpreting F_{ST}: The impact of rare variants. *Genome Research*, *23*(9), 1514–1521. https://doi.org/10.1101/gr.154831.113

Blum, M. J. (2002). Rapid movement of a Heliconius hybrid zone: evidence for phase III of Wright's shifting balance theory? *Evolution; International Journal of Organic Evolution*, *56*(10), 1992–1998. https://doi.org/10.1554/0014-3820(2002)056[1992:RMOAHH]2.0.CO;2

Brien, M. N., Enciso-Romero, J., Parnell, A. J., Salazar, P. A., Morochz, C., Chalá, D., … Nadeau, N. J. (2018). Phenotypic variation in Heliconius erato crosses shows that iridescent structural colour is sex-linked and controlled by multiple genes. *Interface Focus*, *9*(1), 20180047. https://doi.org/10.1098/rsfs.2018.0047

Brown, K. S. (1981). The Biology of Heliconius and Related Genera. *Annual Review of Entomology*, *26*(1), 427–457. https://doi.org/10.1146/annurev.en.26.010181.002235

Challis, R. J., Kumar, S., Dasmahapatra, K. K., Jiggins, C. D., & Blaxter, M. (2016). Lepbase: The Lepidopteran genome database. *BioRxiv*, 056994. https://doi.org/10.1101/056994

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*(3), 695–708. https://doi.org/10.1534/g3.115.023655

Dell'aglio, D. D., Stevens, M., & Jiggins, C. D. (2016). Avoidance of an aposematically coloured butterfly by wild birds in a tropical forest. *Ecological Entomology*, *41*(5), 627–632. https://doi.org/10.1111/een.12335

Derryberry, E. P., Derryberry, G. E., Maley, J. M., & Brumfield, R. T. (2013). Hzar: Hybrid zone analysis using an R software package. *Molecular Ecology Resources*, *14*(3), 652–663. https://doi.org/10.1111/1755-0998.12209

Endler, J. A. (1977). *Geographic variation, speciation, and clines*. Princeton, New Jersey: Princeton University Press.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Flaxman, S. M., Wacholder, A. C., Feder, J. L., & Nosil, P. (2014). Theoretical models of the influence of genomic architecture on the dynamics of speciation. *Molecular Ecology*, *23*(16), 4074–4088. https://doi.org/10.1111/mec.12750

Gay, L., Crochet, P. a., Bell, D. a., & Lenormand, T. (2008). Comparing clines on molecular and phenotypic traits in hybrid zones: A window on tension zone models. *Evolution*, *62*(11), 2789–2806. https://doi.org/10.1111/j.1558-5646.2008.00491.x

Gompert, Z., Mandeville, E. G., & Buerkle, C. A. (2017). Analysis of Population Genomic Data from Hybrid Zones. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 207–229. https://doi.org/10.1146/annurev-ecolsys-110316-022652

Hudson, R. R., Slatkin, M., & Maddison, W. P. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics*, *132*(2), 583–589. https://doi.org/PMC1205159

Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The K = 2 conundrum. *Molecular Ecology*, *26*(14) , 3594–3602. https://doi.org/10.1111/mec.14187

Joron, M., Papa, R., Beltrán, M., Chamberlain, N., Mavárez, J., Baxter, S., … Jiggins, C. D. (2006). A conserved supergene locus controls colour pattern diversity in Heliconius butterflies. *PLoS Biology*, *4*(10), 1831–1840. https://doi.org/10.1371/journal.pbio.0040303

Kawakami, T., Butlin, R. K., Adams, M., Paull, D. J., & Cooper, S. J. B. (2009). Genetic analysis of a chromosomal hybrid zone in the Australian morabine grasshoppers (Vandiemenella, viatica species group). *Evolution*, *63*(1), 139–152. https://doi.org/10.1111/j.1558-5646.2008.00526.x

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Kronforst, M. R., Young, L. G., Kapan, D. D., Mcneely, C., Neill, R. J. O., & Gilbert, L. E. (2006). Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. *Proceedings of the National Academy of Sciences*, *103*(17), 6575–6580.

Kruuk, L. E. B., Baird, S. J. E., Gale, K. S., & Barton, N. H. (1999). A Comparison of Multilocus Clines Maintained by Environmental Adaptation or by Selection Against Hybrids. *Genetics*, *153*(4), 1959–1971.

Langham, G. M. (2004). Specialized Avian Predators Repeatedly Attack Novel Color Morphs of Heliconius Butterflies. *Evolution*, *58*(12), 2783–2787.

Lango Allen, H., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., … Hirschhorn, J. N. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, *467*(7317), 832–838. Retrieved from http://dx.doi.org/10.1038/nature09410

Le Corre, V., & Kremer, A. (2012). The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, *21*(7), 1548–1566. https://doi.org/10.1111/j.1365-294X.2012.05479.x

Mallet, J. (1989). The Genetics of Warning Colour in Peruvian Hybrid Zones of Heliconius erato and H. melpomene. *Proceedings of the Royal Society B: Biological Sciences*, *236*(1283), 163–185. https://doi.org/10.1098/rspb.1989.0019

Mallet, J. (1986). Hybrid zones of Heliconius butterflies in Panama and the stability and movement of warning colour clines. *Heredity*, *56*(2), 191–202. https://doi.org/10.1038/hdy.1986.31

Mallet, J. (1993). Speciation, raciation, and color pattern evolution in Heliconius butterflies: evidence from hybrid zones. In Richard G Harrison (Ed.), *Hybrid zones and the evolutionary process* (pp. 226–260). New York: Oxford University Press.

Mallet, J., Barton, N. H., Gerardo Lamas, M., Jose Santisteban, C., Manuel Muedas, M., & Eeley, H. (1990). Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in Heliconius hybrid zones. *Genetics*, *124*, 921–936.

Mallet, J., & Barton, N. H. (1989). Strong Natural Selection in a Warning-Color Hybrid Zone. *Evoution*, *43*(2), 421–431. https://doi.org/10.2307/2409217

Martin, A., Papa, R., Nadeau, N. J., Hill, R. I., Counterman, B. A., Halder, G., … Reed, R. D. (2012). Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *Proceedings of the National Academy of Sciences*, *109*(31), 12632–12637. https://doi.org/10.1073/pnas.1204800109

Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Slazar, C., Walters, J. R., Simpson, F., … Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in Heliconius butterflies. *Genome Research*, *23*, 1817–1828. https://doi.org/10.1101/gr.159426.113.

Meisner, J., & Albrechtsen, A. (2018). Inferring Population Structure and Admixture Proportions in Low Depth Next-Generation Sequencing Data. *BioRxiv*, 302463. https://doi.org/10.1101/302463

Milano, E. R., Kenney, A. M., & Juenger, T. E. (2016). Adaptive differentiation in floral traits in the presence of high gene flow in scarlet gilia (Ipomopsis aggregata). *Molecular Ecology*, *25*(23), 5862–5875. https://doi.org/10.1111/mec.13752

Nadeau, N. J. (2016). Genes controlling mimetic colour pattern variation in butterflies. *Current Opinion in Insect Science*, Vol. 17, pp. 24–31. https://doi.org/10.1016/j.cois.2016.05.013

Nadeau, N. J., Pardo-Diaz, C., Whibley, A., Supple, M. A., Saenko, S. V., Wallbank, R. W. R., … Jiggins, C. D. (2016). The gene cortex controls mimicry and crypsis in butterflies and moths. *Nature*, *534*(7605), 106–110. https://doi.org/10.1038/nature17961

Nadeau, N. J., Ruiz, M., Salazar, P., Counterman, B., Medina, J. A., Ortiz-Zuazaga, H., … Papa, R. (2014). Population genomics of parallel hybrid zones in the mimetic butterflies, H. melpomene and H. erato. *Genome Research*, *24*(8), 1316–1333. https://doi.org/10.1101/gr.169292.113

Nadeau, N. J., Whibley, A., Jones, R. T., Davey, J. W., Dasmahapatra, K. K., Baxter, S. W., …

Jiggins, C. D. (2012). Genomic islands of divergence in hybridizing Heliconius butterflies identified by large-scale targeted sequencing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1587), 343–353. https://doi.org/10.1098/rstb.2011.0198

Parnell, A. J., Bradford, J. E., Curran, E. V., Washington, A. L., Adams, G., Brien, M. N., … Nadeau, N. J. (2018). Wing scale ultrastructure underlying convergent and divergent iridescent colours in mimetic *Heliconius* butterflies. *Journal of The Royal Society Interface*, *15*(141), 20170948. https://doi.org/10.1098/rsif.2017.0948

Phillips, B. L., Baird, S. J. E., & Moritz, C. (2004). When Vicars Meet: A Narrow Contact Zone between Morphologically Cryptic Phylogeographic Lineages of the Rainforest Skink, Carlia rubrigularis. *Evolution*, *58*(7), 1536–1548.

Poelstra, J. W., Vijay, N., Bossu, C. M., Lantz, H., Ryll, B., Muller, I., … Wolf, J. B. W. (2014). The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, *344*(6190), 1410–1414. https://doi.org/10.1126/science.1253226

Reed, R. D., Papa, R., Martin, A., Hines, H. M., Counterman, B. a, Pardo-Diaz, C., … McMillan, W. O. (2011). Optix Drives the Repeated Convergent Evolution of Butterfly Wing Pattern Mimicry. *Science*, *333*(6046), 1137–1141. https://doi.org/10.1126/science.1208227

Rosser, N, Phillimore, A. B., Huertas, B. C., Willmott, K. R., & Mallet, J. (2012). Testing historical explanations for gradients in species richness in heliconiine butterflies of tropical America. *Biological Journal of the Linnean Society*, *105*, 479–497. https://doi.org/10.1111/j.1095-8312.2011.01814.x

Rosser, N., Dasmahapatra, K. K., & Mallet, J. (2014). Stable Heliconius butterfly hybrid zones are correlated with a local rainfall peak at the edge of the Amazon basin. *Evolution*, *68*(12), 3470–3484. https://doi.org/10.1111/evo.12539

Scordato, E. S. C., Wilkins, M. R., Semenov, G., Rubtsov, A. S., Kane, N. C., & Safran, R. J. (2017). Genomic variation across two barn swallow hybrid zones reveals traits associated with divergence in sympatry and allopatry. *Molecular Ecology*, *26*(20), 5676–5691. https://doi.org/10.1111/mec.14276

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, *195*(3), 693–702. https://doi.org/10.1534/genetics.113.154138

Stankowski, S., Sobel, J. M., & Streisfeld, M. A. (2015). The geography of divergence with gene flow facilitates multitrait adaptation and the evolution of pollinator isolation in Mimulus aurantiacus. *Evolution*, *69*(12), 3054–3068. https://doi.org/10.1111/evo.12807

Supple, M. A., Hines, H. M., Dasmahapatra, K. K., Lewis, J. J., Nielsen, D. M., Lavoie, C., … Counterman, B. A. (2013). Genomic architecture of adaptive color pattern divergence and convergence in Heliconius butterflies. *Genome Research*, *23*, 1248–1257. https://doi.org/10.1101/gr.150615.112.1248

Sweeney, A., Jiggins, C., & Johnsen, S. (2003). Polarized light as a butterfly mating signal. *Nature*, *423*(6935), 31–32. https://doi.org/10.1038/423031a

Szymura, J. M., & Barton, N. H. (1991). The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and between loci. *Evolution*, *45*(2), 237–261. https://doi.org/10.2307/2408943

Szymura, J. M., & Barton, N. H. (1986). Genetic Analysis of a Hybrid Zone Between the Fire-Bellied Toads, Bombina bombina and B. variegata, Near Cracow in Southern Poland. *Evolution*, *40*(6),

1141–1159.

Thurman, T. J., Szejner-Sigal, A., & McMillan, W. O. (2019). Movement of a Heliconius hybrid zone over 30 years: A Bayesian approach. *Journal of Evolutionary Biology*, *32*(9), 974–983. https://doi.org/10.1111/jeb.13499

Toews, D. P. L., Taylor, S. A., Vallender, R., Brelsford, A., Butcher, B. G., Messer, P. W., & Lovette, I. J. (2016). Plumage Genes and Little Else Distinguish the Genomes of Hybridizing Warblers. *Current Biology*, *26*(17), 2313–2318. https://doi.org/10.1016/j.cub.2016.06.034

Van Belleghem, S. M., Rastas, P., Papanicolaou, A., Martin, S. H., Arias, C. F., Supple, M. A., … Papa, R. (2017). Complex modular architecture around a simple toolkit of wing pattern genes. *Nature Ecology and Evolution*, *1*(3), 1–12. https://doi.org/10.1038/s41559-016-0052

Vines, T. H., Dalziel, A. C., Albert, A. Y. K., Veen, T., Schulte, P. M., & Schluter, D. (2016). Cline coupling and uncoupling in a stickleback hybrid zone. *Evolution*, *70*(5), 1023–1038. https://doi.org/10.1111/evo.12917

Westerman, E. L., VanKuren, N. W., Massardo, D., Tenger-Trolander, A., Zhang, W., Hill, R. I., … Kronforst, M. R. (2018). Aristaless Controls Butterfly Wing Color Variation Used in Mimicry and Mate Choice. *Current Biology*, *0*(0), 3469–3474. https://doi.org/10.1016/J.CUB.2018.08.051

Whitlock, M. C., & Schluter, D. (2009). *The Analysis of Biological Data*. Greenwood Village, Colorado: Roberts and Company Publishers.

Data Accessibility Statement

Sequence data have been deposited in the European Nucleotide Archive with the project number PREJEB32848.

Author Contributions

EVC and NJN conceived and designed the study. EVC, NJN, CPD, CAS and ML carried out field work. EVC generated and analysed the data with the help of SS. EVC wrote the manuscript, and all co-authors revised the manuscript.

Figures and tables

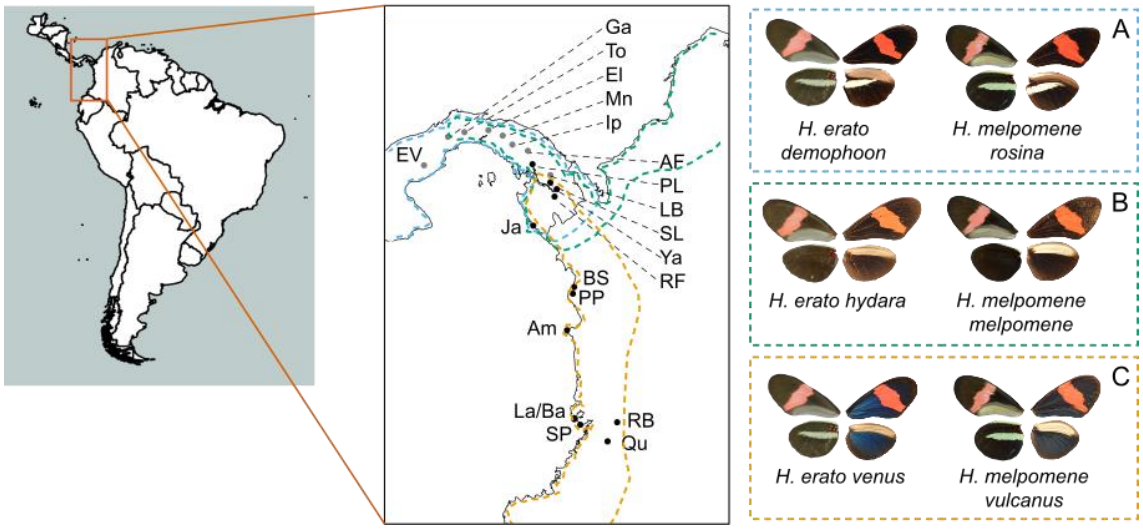

Figure 1 – Sampled populations in Colombia and Panama. Sites are labelled with abbreviations (further information about sites and collections are in Table S1). Photographs show the phenotypes of mimetic races of *H. erato* and *H. melpomene* from Central America (**A**), North Colombia (**B**), and Western Colombia (**C**). For each pictured phenotype, the wings on the left-hand side show the ventral wing pattern, and the wings on the right-hand side show the dorsal wing pattern. Approximate ranges for the mimetic race pairs are outlined with dashed lines (Rosser, Phillimore, Huertas, Willmott, & Mallet, 2012). Populations that are included in the phenotypic analysis only are shown in grey, populations that are included in both the phenotypic and genetic analysis are shown in black.

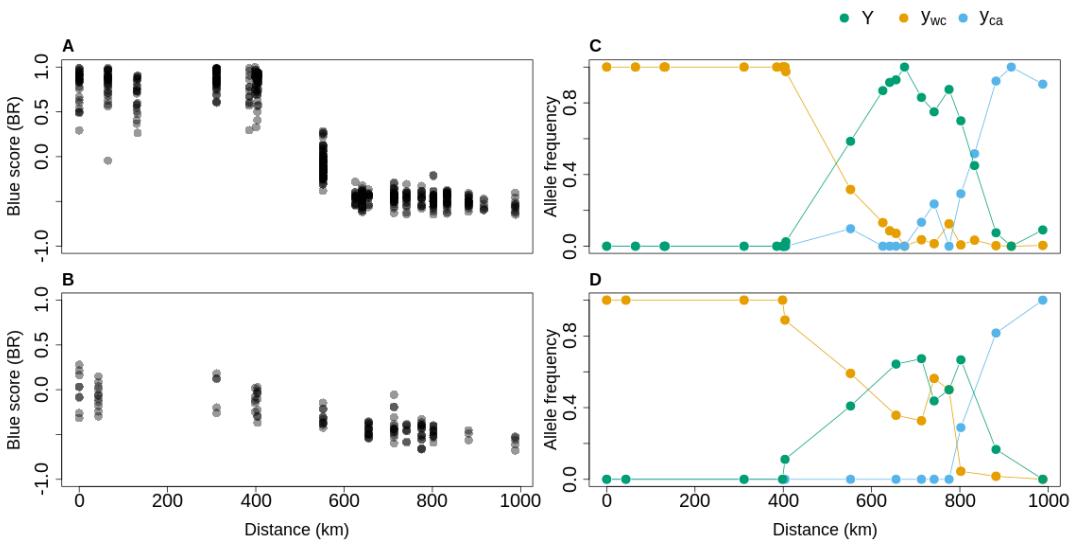

Figure 2. Variation in wing pattern phenotypes across the transect, from 0 km at Queremal, Colombia, to 987.85 km at El Valle, Panama. Blue score (BR) values of each individual for *H. erato* (**A**) and *H. melpomene* (**B**). Frequency of the North Colombian (Y, green), West Colombian ($y_{wc,}$ orange) and Central American ($y_{ca,}$ blue) yellow bar allele at each site with 5 or more samples for *H. erato* (**C**) and *H. melpomene* (**D**).

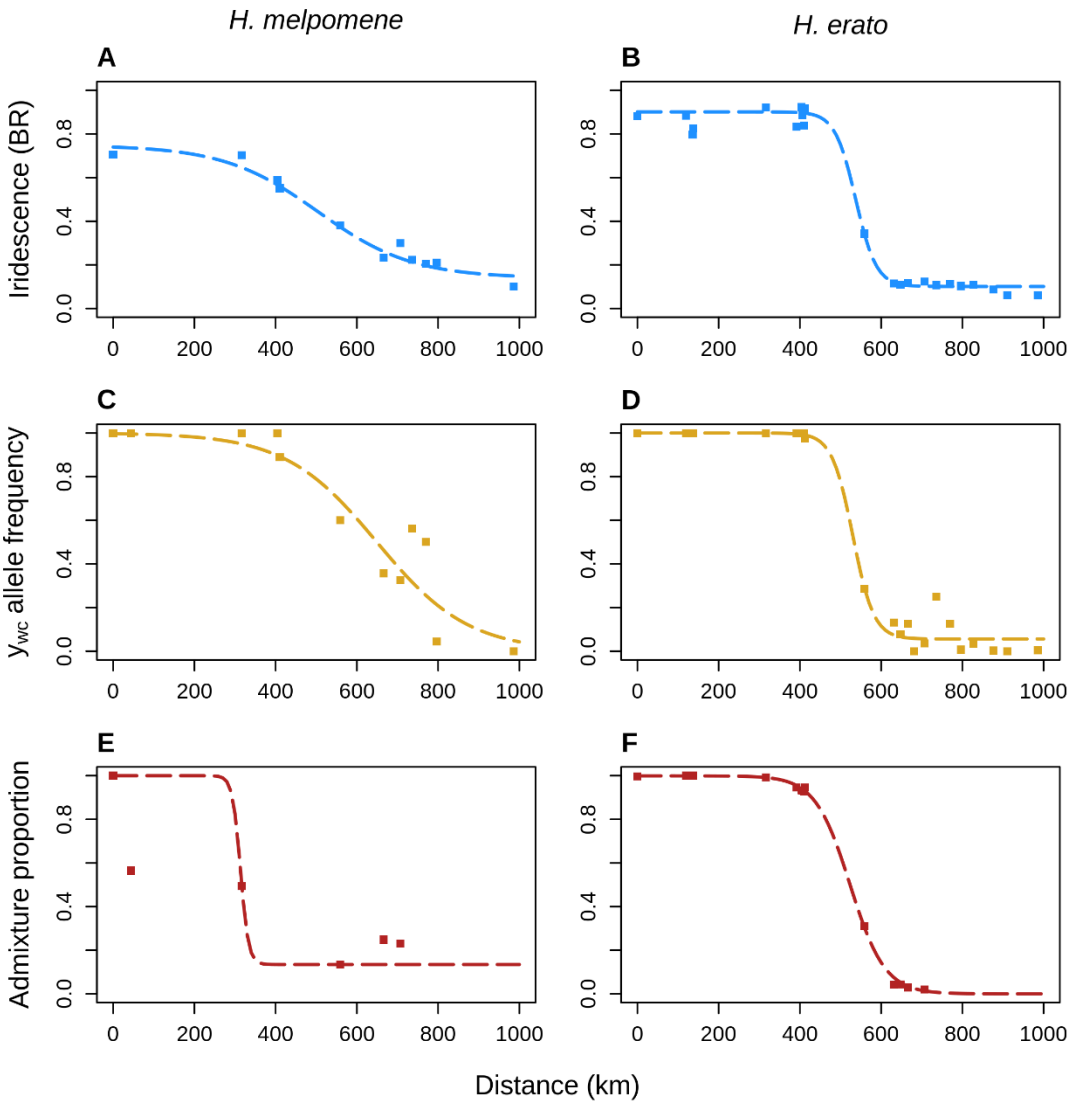

Figure 3 – The best fitting geographic clines (dashed lines) of iridescence (**A, B**; blue), the West Colombian yellow bar allele frequency (y_{wc}; **C, D**; yellow), and admixture proportions (**E, F**; red), across a transect of sampling sites (points) for *Heliconius erato* (**A, C, E**) and *Heliconius melpomene* (**B, D, F**). The transect begins (at 0 km) in the Queremal locality, in the Cauca Valley region of Colombia.

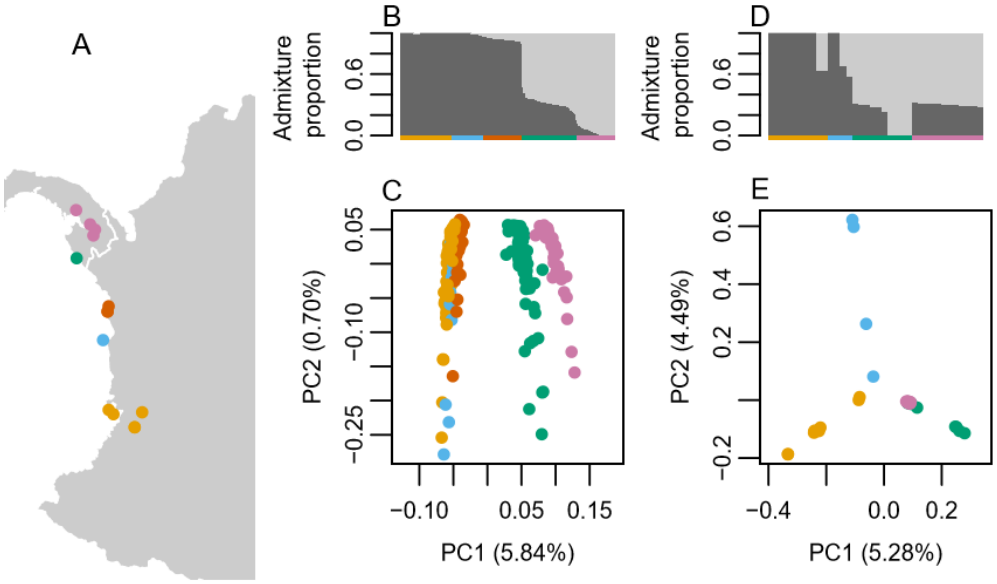

Figure 4 – Population structure across the hybrid zones in *H. erato* (**B,C**) and in *H. melpomene* (**D, E**). Sampling locations across the hybrid zone (**A**). Individual admixture proportions estimated using NGSadmix, with k = 2 (**B, D**). Each horizontal bar represents an individual, bar colour represents the estimated proportion of ancestry derived from population 1 (dark grey) or population 2 (light grey). Vertical bars indicate the population of origin, colours match those on the map. Principal components analysis (**C, E**). Colour of points indicate the population of origin, as shown on the map.

Figure 5 – Relationship between geographic distance and genetic differentiation (genome-wide average F_{ST}) between sampling sites in *H. erato* (**A**) and *H. melpomene* (**B**). Pairwise comparisons are colour-coded to indicate comparisons between populations of the same colour pattern race (blue), between populations of different colour pattern races (yellow), and comparisons where one population is from the hybrid zone centre (green).

Table 1 – Cline parameter estimates for variation in iridescence, y_{wc} allele frequency, and admixture proportion across transects for *H. erato* and *H. melpomene*, which begin at the Queremal locality. ML estimates for sigmoid models (Sig), symmetrical stepped models (Sstep), and asymmetrical stepped models (Astep) were estimated for each trait. If a model is a significantly better fit as determined by a likelihood ratio test (details in Table S6) it is denoted with *. Parameters are log-likelihood (LnL) cline centre (c), width (w), barrier strength for either side of stepped models (B_0/w, B_1/w), the rate of exponential decay for either tail (θ_0, θ_1).

Species	Trait	Model	LnL	Centre (km)	Width (km)	pmin	pmax	B_0/w	θ_0	B_1/w	θ_1
Heliconius erato	Iridescence	Sig	-38.62	537.38 (539.92-34.97)	101.66 (89.66-113.50)	0.101 (0.096-0.106)	0.901 (0.889-0.914)	--	--	--	--
		Sstep	-34.12	549.48 (508.87-562.98)	46.11 (38.78-74.86)	0.084 (0.020-0.092)	0.917 (0.888-0.913)	63.63 (37.24-80.76)	0.012 (0.0003-0.9998)	--	--
		Astep*	-27.34	546.38 (547.28-550.91)	78.62 (69.46-90.07)	0.089 (0.021-0.0903)	0.90 (0.089-0.92)	40.63 (38.69-56.57)	0.005 (0.002-0.006)	9.56E+09 (2E+06 – 1E+10)	0.73 (0.001-0.999)
	y_{wc}	Sig	-9.29	530.53 (515.48-543.70)	102.87 (70.90-145.19)	0.056 (0.013-0.088)	1.000 (0.989-1.00)	--	--	--	--
		Sstep	-9.83	530.53 (510.32-543.20)	102.81 (66.99-167.62)	0.056 (0.011-0.087)	0.999 (0.989-1.000)	7.44E+09 (195247-1E+10)	0.35 (0.00-0.99)	--	--
		Astep	-6.59	536.09 (523.98-547.17)	98.19 (70.48-145.52)	1.00E-04 (0.000-0.087)	1.00 (0.989-1.000)	8.41 (4.17-13.34)	0.075 (0.052-0.144)	6.86E+09 (3886-1E+10)	0.51 (0.000-0.991)
	Admixture proportion	Sig	-25.83	523.60 (521.26-525.76)	171.36 (167.33-175.56)	0.000017 (0.000015-0.010190)	0.9995 (0.996-0.9998)	--	--	--	--

		Sstep	-20.53	524.28 (520.34-527.59)	165.84 (163.23-169.74)	0.000012 (0.00001-0.00078)	0.99 (0.996-1.000)	1.01 (0.83-1.54)	0.98 (0.932-0.999)	--	--
		Astep*	-13.63	536.79 (522.65-537.43)	101.30 (97.54-110.43)	0.021 (0.018-0.023)	0.99 (0.9997-1.000)	4144656896 (6775875-778645878)	0.11 (0.003-0.999)	1.88 (0.994-37.975)	0.37 (0.00-0.873)
Heliconius melpomene	Iridescence	Sig	-5.82	504.27 (474.68-532.89)	466.31 (380.85-567.89)	0.14 (0.12-0.17)	0.75 (0.70-0.79)	--	--	--	--
		Sstep	-3.62	553.87 (468.98-557.76)	0.0001 (0.00001-0.0118)	0.0001 (0.00001-0.0118)	0.82 (0.79-0.83)	2.10 (0.17-12.34)	0.021 (0.001-0.046)	--	--
		Astep	-3.40	572.18 (542.77-584.88)	454.48 (398.87-528.72)	0.0001 (0.00001-0.0087)	0.72 (0.70-0.73)	1.72 (1.23-17.35)	0.17 (0.000-0.65)	136571824 (943923-232068365)	0.89 (0.67-0.999)
	y_{wc}	Sig	-6.70	649.14 (597.78-666.74)	451.92 (378.88-513.53)	0.0001 (0.0001-0.14)	1.000 (0.999-1.000)	--	--	--	--
		Sstep	-6.70	649.17 (587.88-668.80)	451.78 (382.81-498.02)	0.0001 (0.0001-0.14)	1.000 (0.999-1.000)	7857362432 (4730095-9998937088)	0.78 (0.29-0.998)	--	--
		Astep	-6.05	511.27 (505.82-527.64)	156.12 (135.76-190.54)	0.0001 (0.0001-0.14)	1.00 (0.999-1.000)	0.70 (0.12-0.999)	0.078 (0.001-0.64)	2958797824 (4729519 - 8986493729)	0.999 (0.000-0.999)
	Admixture proportion	Sig	-0.062	313.45 (126.86-363.68)	39.30 (0.43-123.81)	0.13 (0.0001-0.17)	1.000 (0.900-1.000)	--	--	--	--

	Sstep	-0.062	306.36 (150.64-313.85)	122.47 (0.37-152.64)	0.13 (0.0001-0.16)	1.000 (0.900-1.000)	7427586048 (1567365-9867457635)	0.32 (0.28-0.999)	--	--
	Astep	-0.062	311.74 (147.74-343.76)	59.34 (0.52-76.78)	0.13 (0.0001-0.16)	1.000 (0.900-1.000)	1764016384 (863863-8223565965)	0.79 (0.67-0.999)	8084597760 (878456-9864222189)	0.51 (0.000-0.97)

Table 2 – Likelihood ratio tests for coincidence (*c*) and concordance (*w*) of iridescence, yellow bar, and admixture proportion clines. ΔML is the test statistic, d.f. is degrees of freedom. The combination of clines being compared is noted under Trait(s).

	c			***w***		
Trait(s)	ΔML	d.f.	*P*	ΔML	d.f.	*P*
H. erato						
iridescence, yellow bar	0.75	1	0.386	0	1	1.000
iridescence, admixture proportion	11.22	1	< 0.001	16.85	1	< 0.001
iridescence, yellow bar, admixture proportion	11.93	2	0.003	20.47	2	< 0.001
H. melpomene						
iridescence, yellow bar	0.80	1	0.370	0.004	1	0.950
iridescence, admixture proportion	0.22	1	0.639	16.85	1	< 0.001
iridescence, yellow bar, admixture proportion	1.02	2	0.599	0.01	2	0.994
Both species						
iridescence	0.62	1	0.430	6.42	1	0.011
yellow bar	0.59	1	0.443	1.70	1	0.192
admixture proportion	0.22	1	0.639	0	1	1.000
iridescence, yellow bar	1.88	3	0.598	8.13	3	0.043
iridescence, yellow bar, admixture proportion	12.95	5	0.024	25.92	5	< 0.001