Molecular Ecology (2008) doi: 10.1111/j.1365-294X.2008.03934.x

A hybrid zone provides evidence for incipient ecological speciation in *Heliconius* butterflies

CARLOS F. ARIAS,*‡ ASTRID G. MUÑOZ,* CHRIS D. JIGGINS,† JESÚS MAVÁREZ,‡§ ELDREDGE BERMINGHAM‡ and MAURICIO LINARES*

*Instituto de Genética, Universidad de los Andes, Carrera 1 No 18a – 70, PO Box 4976, Bogotá D.C., Colombia, †Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK, ‡Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Ancón, Panamá, República de Panamá, §Centro de Ecología, Instituto Venezolano de Investigaciones Científicas, Apartado 21827, Caracas 1020-A, Venezuela

Abstract

In *Heliconius* butterflies, it has been proposed that speciation occurs through a combination of divergence in ecological habitat preferences and mimetic colour patterns. Here we test this hypothesis by investigating a parapatric form of the widespread species *Heliconius erato*. Mendelian (colour patterns) and molecular genetic data permit us to address hypotheses about introgression and genetic differentiation between different populations. Combined analysis of colour pattern, microsatellite loci and mitochondrial DNA showed that *Heliconius erato venus* and *Heliconius erato chestertonii* form a bimodal hybrid zone implying partial reproductive isolation. In a sample of 121 individuals collected in sympatry, 25% were hybrids representing a significant deficit of heterozygotes compared to the Hardy–Weinberg expectation. Seven microsatellite loci, analysed for a subset of these individuals, showed marked differentiation between the parental taxa, and unambiguously identified two genotypic clusters concordant with our phenotypic classification of individuals. Mitochondrial DNA analysis showed *H. erato venus* as a monophyletic group well differentiated from *H. erato chestertonii*, implying a lack of historical introgression between the populations. *Heliconius erato chestertonii* is therefore an incipient species that maintains its integrity despite high levels of hybridization. Moreover, *H. erato chestertonii* is found at higher altitudes than other races of *H. erato* and has a distinct colour pattern and mimetic relationship. Hence, there are now two examples of parapatric incipient species related to *H. erato*, *H. himera* and *H. erato chestertonii*, both of which are associated with higher altitudes, more arid habitats and distinct mimetic relationships. This implies that parapatric habitat adaptation is a likely cause of speciation in this group.

Keywords: bimodal hybrid zone, colour pattern, *Heliconius*, incipient speciation, microsatellites, mtDNA

Received 1 May 2008; revision received 8 July 2008; accepted 8 August 2008

Introduction

Hybrid zones are areas where divergent lineages come into contact, mate and produce offspring of mixed ancestry (Barton & Hewitt 1989), and as such offer insights into the process of speciation. Natural hybrid zones represent a continuum of different levels of differentiation, as expected if speciation is a gradual process. At one end of the continuum are zones composed largely of recombinant individuals, known as unimodal hybrid zones, while at the other end lie zones in which discernible clusters of parental-like individuals co-exist, termed bimodal hybrid zones (Howard & Waring 1991; Harrison & Bogdanowicz 1997; Rolan-Alvarez *et al*. 1997; Jiggins & Mallet 2000; Ross & Harrison 2002; Redenbach & Taylor 2003; Vedenina & Helversen 2003). The former pattern suggests random-mating, while the latter implies that a degree of reproductive isolation has developed and speciation is partially complete. In general, unimodal hybrid zones are associated with a lack of

Correspondence: Carlos F. Arias, Fax: (57)(1)(3324069); E-mail: solracarias@gmail.com

prezygotic isolation, while bimodal hybrid zones often show strong prezygotic isolation as well as disruptive selection against hybrid genotypes (Jiggins & Mallet 2000). The existence of this continuum in the degree of reproductive isolation between hybridizing forms supports the view of speciation as a gradual and cumulative process (Jiggins & Mallet 2000). Therefore, documentation of cases in different stages of this continuum are important for our understanding of the speciation process (Grant & Grant 2003; Coyne & Orr 2004).

Heliconius butterflies exemplify this continuum and show a range of intermediate steps on the path to speciation (Mallet *et al*. 1998). These range from colour pattern races that overlap and mate randomly in hybrid zones, to sympatric species that hybridize rarely or never (Mallet *et al*. 1990, 2007; Jiggins *et al*. 1996). Indeed, most *Heliconius* species are composed of several geographical races, characterized by different wing colour patterns. One of the most diverse species is *Heliconius erato*, which is divided into over 20 geographical colour pattern races across its range in South and Central America. When such races meet in the wild, they form hybrid zones of differing width, depending on the strength of selection and gene flow (Mallet *et al*. 1990; Blum 2002). In a few cases, such hybrid zones involve incipient species rather than races. One of the best-studied cases is that between *Heliconius himera* and *Heliconius erato cyrbia* in southwestern Ecuador. Mitochondrial DNA (mtDNA), allozyme and colour pattern analysis have shown that these species form a strongly bimodal hybrid zone with little evidence for interspecific gene flow. Hybrid genotypes and phenotypes are found in low frequencies (10%) in overlapping populations, and genetic differentiation is largely explained by assortative mating and divergent habitat use (Jiggins *et al*. 1996, 1997). In particular, *H. himera* is found at a higher altitudinal range than *H. erato*, and its habitat is significantly more arid. In addition, most races of *H. erato* are involved in mimicry with *Heliconius melpomene*, but *H. himera* is not mimetic, presumably because *H. melpomene* is not found in the dry forest habitat.

It would be interesting to test the conclusions derived from the *H. erato cyrbia* × *H. himera* hybrid zone in other cases of incipient speciation in the same *Heliconius* subgroup. Fortunately, another highly differentiated race of *H. erato* occurs in the Cauca Valley in Colombia, namely *Heliconius erato chestertonii*. Similar to *H. himera*, *H. erato chestertonii* is found in dry forest, has a higher altitudinal limit than other races of *H. erato* and it is not involved in a mimetic relationship with a race of *H. melpomene*, which is also absent in the Cauca Valley region. However, rather than being non-mimetic, *H. erato chestertonii* is mimetic with *Heliconius cydno* (Linares 1996). In addition, a phylogenetic analysis of mtDNA sequence variation of some *H. erato chestertonii* individuals has shown that this race is basal to the western clade of *H. erato* and is more divergent than any other race of this species (2.3% divergence) (Brower 1994, 1996). Therefore, there is preliminary evidence suggesting that *H. erato chestertonii* is more than just another race of *H. erato*.

In the Calima River Valley in the Western Cordillera of the Andes, *H. erato chestertonii* comes into contact and forms a hybrid zone with *Heliconius erato venus*. The latter is found in the lowland wet forest and associated disturbed habitats on the pacific coast of Colombia and mimics *Heliconius melpomene vulcanus*. Mitochondrial DNA haplotypes sampled from *H. erato venus* fall into the western clade of *H. erato* most closely related to adjacent populations from western Ecuador and Central America (Brower 1994, 1996). Thus, the hybrid zone between *H. erato chestertonii* and *H. erato venus* gives an opportunity to test the hypothesis that the unusual ecological characteristics of *H. erato chestertonii* might be associated with incipient reproductive isolation.

Here we take advantage of the simple genetic basis of colour pattern to investigate linkage disequilibrium (LD) and heterozygote deficits in the hybrid zone. Most races of *H. erato* show colour pattern alleles in approximate Hardy–Weinberg equilibrium and with only weak LD between physically unlinked genes. Hence, marked heterozygote deficits at colour pattern loci in a hybrid zone imply reproductive isolation. Second, we use microsatellite loci and mtDNA sequences to examine genotypic data for regions unlinked, as far as we are aware, to genes under selection. Alleles that introgress freely will have their frequencies homogenized across populations, whereas alleles whose introgression is blocked by reproductive isolation are expected to show strong frequency differences between parental forms and a bimodal pattern in the hybrid zone. If there is significant reproductive isolation, we therefore expect to find a strong deficit of heterozygotes, linkage disequilibrium between loci diagnostic for parental forms, and two different genotypic clusters at the contact zone. Evidence for reproductive isolation between these forms in sympatry would support the hypothesis that ecological habitat adaptation contributes to speciation.

Materials and methods

To describe the hybrid zone, we collected 192 adult butterflies between 1997 and 2004 in four localities throughout the Cauca River Valley and the Pacific coast of Colombia (Fig. 1, Table 1). The Km 15 Calima River Valley locality lies at a gradual ecotone between the dry forest of the Cauca Valley and the wet forest of the Pacific slopes (Fig. 1). In total, 121 individuals were collected in this sympatric locality, which probably corresponds to the centre of the hybrid zone given the roughly equal frequency of the two parental forms (see Results). In addition, we also collected *Heliconius erato chestertonii* and *Heliconius erato venus* from allopatric populations (i.e. geographical localities where all individuals show fixed colour patterns

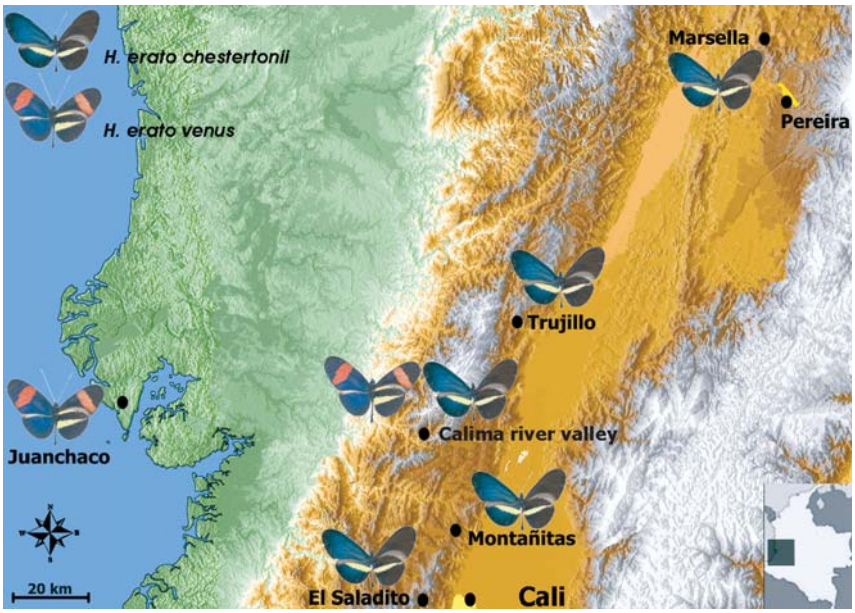

Fig. 1 Sample localities and races used in this study. The map shows the partial distribution of *Heliconius erato venus* and *Heliconius erato chestertonii* on the pacific wet forest and the dry forest of the Cauca river valley of Colombia, respectively. The pattern of coloration of each race is presented. The dorsal wing is shown on the left and ventral view on the right of each butterfly. Detailed explanation of the colour pattern is given in the text.

Table 1 Populations sampled and number of individuals used in each analysis

Species (Subspecies)	No. of individuals Microsa/mtDNA/colour pattern	Locality		
Heliconiuserato chestertonii	0/1/0	Marsella	Cauca valley	Allopatric
H. erato chestertonii	0/2/0	Trujillo	Cauca valley	Allopatric
H. erato chestertonii	0/2/0	Saladito	Cauca valley	Allopatric
H. erato chestertonii	30/5/0	Montañitas	Cauca valley	Allopatric
H. erato chestertonii	11/0/0	Km 4 Calima river valley	Cauca valley	Sympatric
H. erato chestertonii	16/5/32	Km 15 Calima river valley	Cauca valley	Sympatric
Hybrids	16/2/31	Km 15 Calima river valley	Cauca valley	Sympatric
Heliconius erato venus	18/5/58	Km 15 Calima river valley	Cauca valley	Sympatric
H. erato venus	30/5/0	Juanchaco	Pacific coast	Allopatric

characteristics of a single race). Allopatric *H. erato chestertonii* were collected in Montañitas (30 individuals; Fig. 1) in the dry forest habitats of the Cauca Valley, while allopatric samples of *H. erato venus* were collected in Juanchaco (30 individuals; Fig. 1), in wet forest habitats in western (Pacific) Colombia. Eleven more *H. erato chestertonii* individuals were collected in another ecotonal locality, the Km 4 Calima River Valley. Although this locality is considered sympatric here, it is clearly situated on the *H. erato chestertonii* side of the hybrid zone and no *H. erato venus* or hybrids were collected there in this study. Five additional individuals, used for the mtDNA analysis, were sampled in another three localities in the Cauca Valley (Fig. 1, Table 1).

Wings were removed and the bodies preserved in dimethyl sulphoxide or ethanol (96%). All the specimens are stored in the collection of Instituto de Genética de la Universidad de los Andes in Colombia. DNA was extracted from one-third of a thorax of each individual using the DNeasy tissue Kit (QIAGEN), following the manufacturer's protocol.

Colour pattern loci

The study of *Heliconius* hybrid zones is facilitated by the simple genetic basis of their colour patterns (i.e. allele frequencies at colour pattern genes can be determined by visual inspection without the need for molecular markers) (Mallet *et al*. 1990; Jiggins & McMillan 1997; Blum 2002; Merchan *et al*. 2005). In the case of *H. erato chestertonii* and *H. erato venus*, marked differences in colour pattern are largely controlled by just two major loci that segregate in a Mendelian manner (Fig. 1; Fig. S1) and are homologous to those identified in previous studies of *H. erato* races (Sheppard *et al*. 1985; Mallet *et al*. 1990; Jiggins & McMillan

1997): (i) *Heliconius erato venus* has a red band on the forewing against a black background, while *H. erato chestertonii* has an entirely melanic forewing. This difference is controlled by the red band locus, *D*. (ii) On the hindwing, *H. erato venus* has a broad yellow bar on the underside, a white margin visible on both upper and underside and a narrow yellow band on the overlapping region between the forewing and hindwing on the underside, while *H. erato chestertonii* has a broad yellow bar on both the upper and undersides. The yellow bar locus, *Cr*, controls the presence or absence of this broad yellow bar on the upperside of the hindwing, and the white hindwing margin (Sheppard *et al*. 1985; Mallet *et al*. 1990; Jiggins & McMillan 1997). Finally, the upperside background colour of both *H. erato venus* and *H. erato chestertonii* is iridescent, with a slight difference in colour, the latter being more greenish. However, this colour difference does not appear to have a simple genetic basis and is difficult to characterize so is not considered further.

The whole sample from Km 15 Calima River Valley (121 individuals) was analysed for colour pattern loci *D* and *Cr* (the other populations being monomorphic). Wing colour patterns were scored under a binocular microscope and the genotypic data of each individual were recorded based on the differences in wing colour pattern between the forms. Based on genotype scores at the two colour pattern loci, individuals were also classified into four categories: parental forms, F_1, backcrosses, and F_2 or further crosses (Fig. S1). Nonetheless, it is important to note that with only two unlinked Mendelian loci, these classes cannot be distinguished with complete confidence. For example, 25% of individuals in a first generation backcross family are expected to have genotypes identical to an F_1 individual.

Microsatellite genotyping

A total of 60 individuals from the allopatric populations (30 *H. erato chestertonii* and 30 *H. erato venus*; Table 1), 50 individuals from the sympatric population of Km 15 Calima River Valley (16 *H. erato chestertonii*, 18 *H. erato venus*, 16 hybrid individuals; Table 1) and 11 more individuals from Km 4 Calima River Valley were analysed for microsatellite variation (Table 1). Nuclear DNA variation was characterized at seven microsatellite loci (Hel1, Hel4, Hel5, Hel10, Hel12, Hel13, Hel14) using primers and polymerase chain reaction (PCR) conditions outlined by Flanagan *et al*. (2002) and Mavarez & Gonzalez (2006). Loci were amplified with two different conditions: (i) using 0.1 μM dye-labelled (6-FAM, NED, HEX, PET or VIC) forward primer and 0.5 μM reverse primers, or (ii) using 0.01 μM M13-tailed forward primer, 0.4 μM dye-labelled (6-FAM, NED, HEX, PET or VIC) M13 primer (5′CACGACGTTGTAAAACGAC3′) and 0.4 μM reverse primer. Reaction products were diluted (usually 1:20 for NED or HEX, 1:30 for PET and VIC, and 1:40 for 6-FAM-labelled products) and resolved in an ABI PRISM 3100 Genetic Analyser (PE Applied Biosystems). Allele sizes were determined using GeneScan 3.7 and GenoTyper 3.7 (PE Applied Biosystems) with GeneScan ROX 500 or LIZ 500 (Applied Biosystems) as size standards.

Population genetic analysis of colour pattern and microsatellite loci

The Arlequin 2000 package (Schneider *et al*. 2000) was used (i) to test for deviations from Hardy–Weinberg equilibrium for each population, at each locus using the Guo & Thompson's (1992) analogue of Fisher's exact test (HWE, 10^5 Markov chain steps, 10^3 dememorization steps), (ii) to test for linkage disequilibrium (LD) between all pairs of loci and populations using a likelihood-ratio test (10^4 permutations), (iii) to estimate overall levels of genetic differentiation by calculating the estimator Θ of Wright's F_{ST} for each locus (Weir & Cockerham 1984), tested against the null hypothesis of no differentiation by permuting genotypes between populations (10^5 permutations), and (iv) to perform an AMOVA (Excoffier *et al*. 1992) with 10^4 permutations. We tested the hierarchies 'among races', 'among populations within races', and 'within populations', using a model that contained populations and individuals of the two races. In addition, we also calculated Wright's F_{IS} index for each population as estimated by Weir & Cockerham (1984) using GenePop 3.3 (Raymond & Rousset 1995). Positive, negative and zero values of F_{IS} indicate a heterozygote deficit, excess and random union of gametes, respectively. Also, the total sample genotyped for microsatellites from the Km 15 Calima River Valley (50 individuals; *H. erato venus* + *H. erato chestertonii* + hybrids) was analysed as a single population and studied for deviations from Hardy–Weinberg equilibrium, linkage equilibrium and the F_{IS} statistic.

Lepidoptera microsatellites are known to show a high frequency of null alleles (Flanagan *et al*. 2002; Meglecz *et al*. 2004; Jiggins *et al*. 2005; Tobler *et al*. 2005; Mavarez & Gonzalez 2006), which can have a significant impact on estimations of population structure (Chapuis & Estoup 2007). Null allele frequencies were estimated for each locus and population following the expectation-maximization (EM) algorithm of Dempster *et al*. (1977). F_{ST} values were also calculated following the C2 method described in Chapuis & Estoup (2007) and using the FreeNA software, which corrects for the positive bias induced by the presence of null alleles. This method works by calculating Weir's (1996) F_{ST}, where the null allele state is ignored in calculation and the sums of allele and genotype frequencies are not adjusted to 1. Additionally, Micro-Checker software was used to exclude other genotyping errors such as stuttering and drop-out (van Oosterhout *et al*. 2003). We also estimated Cavalli-Sforza & Edwards (1967) pairwise differences, D_C, between all the populations. Chapuis & Estoup (2007) have

shown that genetic distances are less influenced by null alleles than F_{ST} values, with the Cavalli-Sforza & Edwards (1967) genetic distance, D_C, slightly better than Nei's genetic distance, D_S (Nei 1972). The D_C genetic distance was calculated using the software Genetix 4.05.2 (Belkhir *et al*. 2001). No correction for null alleles was applied to any within-population analysis (i.e. HWE, LD or F_{IS}).

Assignment test

In order to test the assignment of individuals to the two *H. erato* forms, we used a Bayesian model–based clustering algorithm implemented in the program Structure 2.1.4 (Pritchard *et al*. 2000). The admixture model and the option of correlated allele frequencies between populations were used, as advised by the authors for cases of subtle population structure. We determined the number of ancestral clusters, *K*, by comparing the likelihood ratios in 10 replicate runs for *K* values between 1 and 10. Each run consisted of 10^6 iterations, after a burn-in period of 10^4 iterations. To use this program, Hardy–Weinberg and linkage equilibrium are assumed, which is not expected in bimodal hybrid zones. However, the software differentiates mixed populations on the basis of allele frequency at each locus. The best estimate of *K* was calculated using the ad hoc statistic ΔK (Evanno *et al*. 2005). This method is based on the estimated rate of change in the log probability of data between successive *K* values, making salient the break in slope of the distribution of $L(K)$ at the true *K* value (Evanno *et al*. 2005). Additionally, we used the Bayesian clustering method of Anderson & Thompson (2002), implemented in the program NewHybrid (Anderson & Thompson 2002) to assign individuals to different genotypic classes (parental, F_1, F_2 or backcrosses). This model considers a case of hybridization between two parental forms A and B, and hybrids formed after *n* generations of interbreeding. The hybridizing population is modelled as a mixture of unknown proportions of individuals of the different genotype frequency classes, corresponding to parental pure species A and pure species B, and various classes of hybrids (F_1, F_2, and backcrosses) from which the sample is drawn randomly. The method computes, by Markov chain Monte Carlo, the Bayesian posterior probability that an individual in a sample belongs to each of different hybrid classes while simultaneously estimating allelic frequencies for parental species. The program was run several times with varying lengths of burn-in period and numbers of sweeps, as recommended by the authors (Anderson & Thompson 2002).

Mitocondrial DNA analysis

We amplified a region of mtDNA spanning the 3′ end of cytochrome oxidase I (COI), leucine tRNA, and cytochrome oxidase II (COII) of 27 specimens of *H. erato venus* and *H. erato chestertonii* from seven populations (Table 1). The PCR was performed using primers and conditions as outlined by Beltrán *et al*. (2002). The PCR products were electrophoretically separated on 1.5% low melting point agarose (NuSieve), and the bands were cut from the gel and dissolved in gelase (Invitrogen). Clean PCR products were sequenced using the DNA sequence kit (BigDye 3.1, PE Applied Biosystems), in an ABI PRISM 3100 Genetic Analyser (PE Applied Biosystems).

Sequences were edited and aligned using Sequencher 3.1 (Gene Codes Corporation, Inc.). The DNA sequences from protein-coding regions were checked for reading-frame errors and termination codons and translated to functional peptide sequences in MacClade 4.02 (Maddison & Maddison 2001). Sequences have been deposited in GenBank (Accession nos EU707581–EU707607). For comparison, we also included seven sequences from Beltrán *et al*. (2002) (GenBank Accession nos AF413684, AF413685, AF413688–AF413691, AY748069) and 16 sequences from Brower (1994) (GenBank Accession nos UO8530, UO8543, UO8565, UO8568–UO8571, U08576, UO8580–UO8582, UO8587–UO8590). These sequences correspond to available haplotypes of *H. erato chestertonii*, *H. erato venus*, six additional races of *H. erato*, *H. himera* and *H. telesiphe* (outgroup).

Phylogenetic analyses were performed with PAUP*version 4.0b8 (Swofford 2000) for maximum parsimony analysis and MrBayes 3.0 (Huelsenbeck & Ronquist 2001) for Bayesian inference. For maximum parsimony (MP), trees were obtained using a heuristic search with tree-bisection–reconnection (TBR) branch swapping. The consensus tree was calculated using a majority rule at 50% and the branch support was calculated by bootstrapping (10^3 replicates, heuristic search with TBR branch swapping). Models of sequence evolution were compared using MrModelTest 2.1 (Nylander 2004). MrModelTest selected GTR + I + G as the most likely model of nucleotide substitution. This pattern of substitution was similar to that already described for a more complete sampling across more taxa in the genus *Heliconius* (Beltrán *et al*. 2002). The Bayesian inference was made using the best-fit model (GTR + I + G) and the analysis was carried out with four Markov chains for 10^6 generations. The consensus tree was calculated using a majority rule at 50%.

Results

Colour pattern. We found that 31 out of 121 (25.6%) individuals analysed for colour pattern loci in the Km 15 Calima River Valley showed a mixed genotype. On the basis of genotypes at the two major loci, the wild hybrids were classified as: 5 F_1 (4.13%), 5 backcrosses to *Heliconius erato chestertonii* (4.13%), 19 backcrosses to *Heliconius erato venus* (15.7%) and 2 F_2 or further cross individuals (1.66%). When the sympatric populations were analysed together, there were significant heterozygote deficits at both colour pattern loci

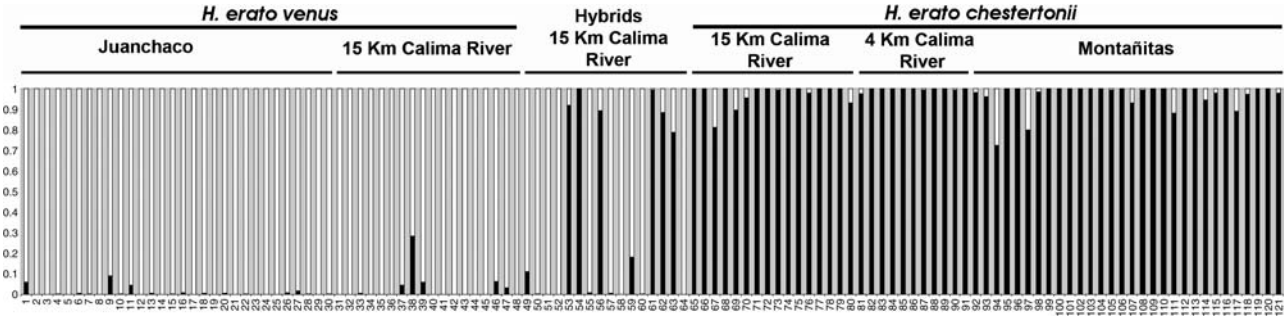

Fig. 2 Bayesian population assignment test. Bar plots showing Bayesian assignment probabilities from the software Structure 2.1.4 for two clusters. Each vertical bar corresponds to one individual. The proportion of each bar that is white (*Heliconius erato venus*) and black (*Heliconius erato chestertonii*) represents an individual's assignment probability to the different clusters. Horizontal bar on the figure shows the collection localities and the phenotypic classification from colour pattern loci D and Cr.

(F_{IS} for $D = 0.59$ $P < 0.01$, F_{IS} for $Cr = 0.8$ $P < 0.01$; Table 2). The expected frequency of F_1-like hybrids (heterozygotes at both *Cr* and *D* loci, calculated by multiplying expected heterozygosity at the two loci) is 26%, which contrasts with the observed frequency of 4.13%. Additionally, we found strong LD between these two loci ($D = +0.178$; $P < 0.01$). The high frequency of parental individuals, heterozygote deficiency and LD together indicate that this hybrid zone shows a substantial degree of bimodality.

Microsatellite loci. Deviations from Hardy–Weinberg equilibrium were found in 17 of 35 comparisons within populations and loci. In all these cases, heterozygote deficiencies produced positive F_{IS} with no consistent pattern of particular loci or populations showing deviations (Table 2). However, the likely presence of null alleles should therefore be taken into account in interpreting our analyses, although differences in the allele frequencies were sufficient to genetically differentiate all populations. Corrections for null alleles were performed (Table 2) and F_{ST} estimations carried out with and without the null allele correction showed very similar results (Table 3). F_{ST} estimates showed significant structure between populations: (i) *Heliconius erato venus* was strongly differentiated from *H. erato chestertonii* ($F_{ST} = 0.324$ $P < 0.01$), (ii) F_{ST} estimates among populations within each race were often significant, but considerably lower than that between the two races (Table 3). Furthermore, the same pattern was observed in pairwise distance analyses, D_C (Table 3). On the other hand, when all individuals captured at the Km 15 Calima River Valley in the centre of the hybrid zone were considered as a single population (*H. erato venus* + *H. erato chestertonii* + hybrids), there were significant heterozygote deficits at all seven microsatellite loci ($P < 0.05$; Table 2), a situation not seen at any population within a race. This heterozygote deficit is very unlikely to be due to within-species pooling of geographical populations (Wahlund effects) given that the spatial scale of our sampling is smaller than the known per-generation dispersal distance of *H. erato* (Mallet *et al*. 1990). The observed deficit is therefore best explained as a result of pooling two sympatric and genetically differentiated forms. This is supported by the strong LD between colour pattern loci (D and Cr) and some of the microsatellite loci (Table 4). For, instance, D showed significant LD with four of the microsatellite loci, while Cr showed LD with three of the loci (Table 4). Despite the caution that should be exercised due to the undoubted presence of null alleles, the microsatellites therefore provide independent support both for the bimodality of the hybrid zone and for a high level of genetic differentiation between the two races. This was further supported by the hierarchical analyses (AMOVA) which indicated that nearly 64.7% of the total variation was due to variation within populations, whereas variation among races accounted for 32.46%, with only 2.78% of the variation occurring between populations within races.

Assignment test. The model-based clustering method implemented in Structure found similar results with different values of *K* ranging from 1 to 10. Moreover, the likelihood values showed an increase with the increase of *K*. Evanno *et al*. (2005) showed that this pattern can lead to an overestimate of the number of populations. To address this problem, we calculate the best estimate of *K* by using the ad hoc statistic ΔK (Evanno *et al*. 2005). The modal value of ΔK was found at $K = 2$ (log ln = –3414.14), with the clusters corresponding to the two *H. erato* races (Fig. 2). Individuals from allopatric samples exhibited q_i average values of 0.99 (*H. erato venus*–Juanchaco) and 0.97 (*H. erato chestertonii*–Montañitas). Additionally, sympatric populations exhibited q_i average values of 0.97 (*H. erato venus*–Calima River Valley) and 0.97 (*H. erato chestertonii*–Calima River Valley). Moreover, the individuals phenotypically classified as hybrids tended to have high q_i values, that were not significantly lower than in parental forms (average 0.948 with a minimum value of 0.787; $P > 0.05$), as expected for backcross or further crosses rather than F_1s.

Table 2 Hardy-Weinberg estimates for colour pattern and microsatellite loci. The estimates were performed with and without null allele corrections for each locality and taking as a single population all individuals captured within the Km 15 transect in the Calima River Valley (*Heliconius erato venus* Km 15 Calima River Valley + *Heliconius erato chestertonii* Km 15 Calima river valley + hybrids)

	Heliconius erato venus Juanchaco		*Heliconius erato venus* Km 15 Calima River Valley		*Heliconius erato chestertonii* Km 15 Calima River Valley		*Heliconius erato chestertonii* Km 4 Calima River Valley		*Heliconius erato chestertonii* Montañitas		All individuals at Km 15 Calima River Valley
	Without null allele correction	With null allele correction	Without null allele correction	With null allele correction	Without null allele correction	With null allele correction	Without null allele correction	With null allele correction	Without null allele correction	With null allele correction	Without null allele correction
Locus	F_{IS}	F_{IS}	F_{IS}	F_{IS}	F_{IS}	F_{IS}	F_{IS}	F_{IS}	F_{IS}	F_{IS}	F_{IS}
D	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	0.59**
Cr	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	Monomorphic	0.8**
Hel1	0.06	–0.01	0.05*	0.05**	0.06	0.06	–0.07	–0.07	0.01	–0.03	0.07*
Hel4	0.34**	–0.13	0.39**	0.08	–0.03	–0.03	0	0	0	0	0.3**
Hel5	0.2	–0.09	0.7**	–0.11	0.27**	–0.15	0.86**	–0.09	0.44**	0.13	0.49**
Hel10	0.52**	–0.12**	0.38**	–0.11	0.18**	–0.12**	0.35	0.12	0.09	–0.08	0.37**
Hel12	0.29**	0.07**	0.18	–0.01	0.46**	0.02	0.13	0.13	0.26**	0.08*	0.28**
Hel13	0.13	0.03	0.19	0.06	0.47**	–0.07	0.48*	–0.05	0.15	0	0.27**
Hel14	–0.15	–0.15	0.21	0.03	0.45**	0.02	0.15	–0.06	0.24*	–0.07	0.22**

*$P < 0.05$, **$P < 0.01$.

Populations	Null alleles	V J	V RC 15	C RC 15	C RC 4	C M
V J	With	—	0.050**	0.120**	0.136**	0.122**
	Without					
V RC 15	With	0.04469**	—	0.100**	0.115**	0.115**
	Without	0.03422**				
C RC 15	With	0.37183**	0.34977**	—	0.041	0.036**
	Without	0.3639**	0.34361**			
C RC 4	With	0.38136**	0.35698**	0.01556	—	0.038**
	Without	0.36735**	0.34383**	0.00073		
C M	With	0.41153**	0.39151**	0.07329**	0.04852**	—
	Without	0.38561**	0.3687**	0.04676**	0.03361**	

Table 3 Matrix of pairwise comparisons of F_{ST} and genetic distance. Values below diagonal are F_{ST} estimates for microsatellite loci (with and without null allele corrections) and values above the diagonal are Cavalli-Sforza & Edwards (1967) pairwise distances between populations

*$P < 0.05$, **$P < 0.01$. V J, *H. erato venus*, Juanchaco; V RC 15, *H. erato venus*, Km 15 Calima River Valley; C RC 15, *H. erato chestertonii*, Km 15 Calima River Valley; C RC 4, *H. erato chestertonii* Km 4 Calima River Valley and C M, *H. erato chestertonii*, Montañitas.

Loci	D	Cr	Hel1	Hel4	Hel5	Hel10	Hel12	Hel13	Hel14
D	HEC3	+	+	+	+	+	–	–	–
Cr	HEC2	*	+	+	+	+	–	–	–
Hel1	HEE13		*	–	+	+	+	–	–
Hel4	HEE7			*	–	–	–	–	–
Hel5	HEC13				*	–	–	–	+
Hel10	HEC10					*	–	–	–
Hel12	–						*	–	–
Hel13	HEC2							*	+
Hel14	HEC2								*

Table 4 Patterns of linkage disequilibrium between colour pattern and microsatellite loci at Km 15 Calima River Valley

+, Markers show significant linkage disequilibrium at $P < 0.05$; –, markers show no significant linkage disequilibrium. Markers *D* and *Cr* are colour pattern loci. In the left hand column are shown the linkage groups of these markers where known (Kapan *et al*. 2006 and Tobler *et al*. 2005). HEE and HEC refer to linkage groups inferred from two different crossing schemes.

Analysis with the software NewHybrid showed similar results as Structure (Fig. 2; Table S1; Fig. S2) with two major genotypic clusters. The genotypic assignment by NewHybrid in the Km 15 Calima River Valley assigned 15 individuals to *H. erato venus* (30%), 20 individuals to *H. erato chestertonii* (40%), and identified 15 individuals (30%) with mixed genotype: 13 backcrosses to *H. erato venus* (24%) and 2 individuals that showed equal probability of being either F_2 or backcrosses to *H. erato chestertonii* (4%). No individuals were assigned as F_1s (Table S1; Fig. S2).

Mitochondrial DNA. We sequenced a mitochondrial region of 1551 pb from 27 individuals of *H. erato venus* and *H. erato chestertonii* representing 795 pb of COI, 62 pb of $tRNA^{leu}$ and 694 pb of COII. With the addition of the available GenBank accessions, the final alignment included sequences from 68 individuals, had 1551 nucleotide sites of which 190 (12.2%) were variable. The parsimony and Bayesian analyses were congruent (Fig. 3). Both analyses showed that all individuals of *H. erato venus* fell into derived positions in the western clade of *H. erato* races (Bayesian $P = 79\%$; Fig. 3). Also, there is some population structure within *H. erato venus*, with the hybrid zone haplotypes falling together in a well-supported clade (bootstrap = 100%, Bayesian $P = 100\%$; Fig. 3). On the other hand, *H. erato chestertonii* was strongly differentiated (and itself paraphyletic) with respect to *H. erato venus*, a result consistent with previous observations based on fewer individuals (Brower 1994, 1996). *Heliconius erato chestertonii* forms three distinct clades: (i) the first is basal to all western *H. erato* races and includes six haplotypes, two from Km 15 at the Calima hybrid zone, two from the study of Brower (one from an unknown location and one from La Moralia, Brower 1994, 1996) and two from Trujillo, north of the hybrid zone (Fig. 3), (ii) a basal clade well differentiated from the other

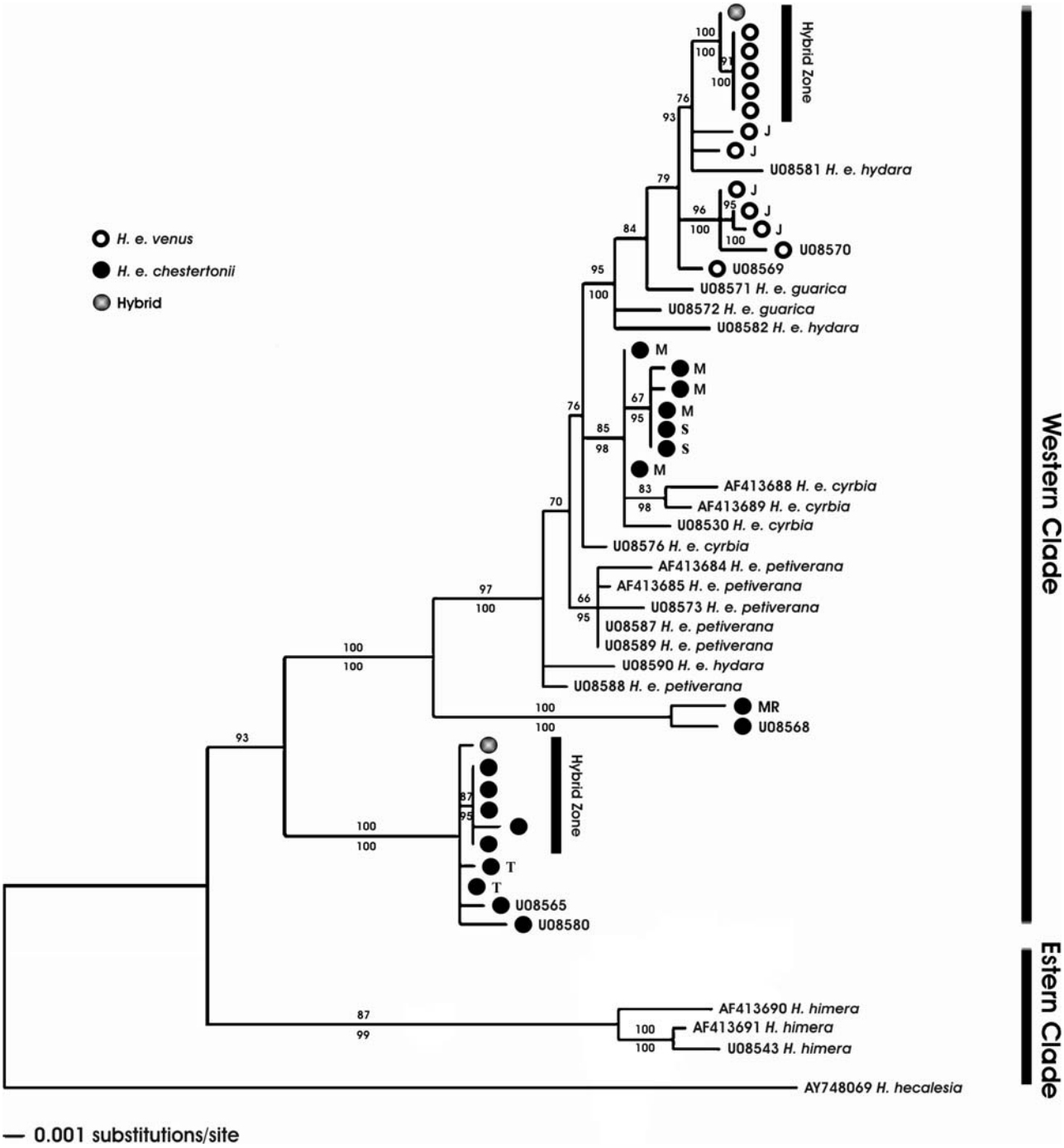

Fig. 3 Phylogenetic relationships of *Heliconius erato venus* and *Heliconius erato chestrertonii* with other races of *H. erato* based on COI and COII sequences. *Heliconius erato venus*, open circles; *H. erato chestertonii*, filled circles. *Heliconius hecalesia* was used as the out-group. Probability values over branches were estimated using Bayesian analysis and values below the branches using parsimony. Hybrid zone, individuals from Km 15 Calima river valley; J, Juanchaco; M, Montañitas; T, Trujillo; S, Saladito; MR, Marsella (Risaraldas).

western *H. erato* races, including one haplotype from Marsella, to the North of the Cauca Valley and a haplotype from Brower´s study (Km 26 Old Cali-Buenaventura Road, Brower 1994, 1996, Fig. 3), and (iii) a clade that includes five haplotypes from Montañitas, two haplotypes from El Saladito, both populations to the south of the hybrid zone, and three haplotypes of *H. erato cyrbia*, a western race of *H. erato* from Ecuador (bootstrap = 98%, Bayesian P = 85%, Fig. 3).

Discussion

The results presented here show strong evidence for a bimodal hybrid zone between *Heliconius erato venus* and *Heliconius erato chestertonii*. This provides support for the hypothesis that parapatric ecological adaptation contributes to speciation in this group, which was originally proposed based on study of another bimodal hybrid zone in Ecuador between *Heliconius erato* and *Heliconius himera* (Jiggins *et al*. 1997). Bimodal hybrid zones are regions where hybridizing forms overlap, but nonetheless retain their genetic integrity as distinct lineages. In other words, they are hybrid zones between forms already showing high levels of reproductive isolation, and as such represent an important intermediate step towards speciation (Harrison 1993; Harrison & Bogdanowicz 1997; Jiggins & Mallet 2000). The contact zone studied here shows a slow but well-defined ecotonal change from dry to wet forest. Thus, *H. erato chestertonii* is similar to the previously studied *H. himera* in being a parapatric isolate of *H. erato* which is associated with dryer forest habitats and a higher altitudinal range. The partial reproductive isolation between both of these taxa and adjacent populations of *H. erato* supports the suggestion that adaptation to these habitats contributes to speciation.

Although we have not been able to map this hybrid zone in detail, field observations suggest that the zone has a width of ≈ 4 km. This contrasts with the width of other hybrid zones between *H. erato* races which varies from ≈ 10 km to ≈ 140 km (Mallet 1986; Mallet *et al*. 1990, 1998; Blum 2002), while the interspecific hybrid zones between *H. himera* and *Heliconius erato cyrbia* shows a width of 5 km (Jiggins *et al*. 1997). Thus, the width of the hybrid zone between *H. erato venus* and *H. erato chestertonii* seems more similar to interspecific than interracial hybrid zones in *H. erato*. The hybrid zone also appears clinal, similar to other *Heliconius* hybrid zones and shows no evidence of a mosaic distribution of genotypes, which is commonly associated with bimodality (Howard & Waring 1991; Bridle *et al*. 2001; Ross & Harrison 2002). Mosaic hybrid zones are common in organisms with low dispersal ability and highly patchy habitat distributions, neither of which is the case for *H. erato* (Mallet *et al*. 1990). Obviously, a more detailed study of the geographical structure and extent of the hybrid zone would be of considerable interest, but is currently precluded by political unrest in the study area.

Analysis of microsatellite data using the model-based clustering method differentiated two genotypic clusters that were largely concordant with the classification of individuals based on colour pattern. Individuals that were classified as phenotypic hybrids generally also had genomes of mixed origin although this was not always the case. Some discordance between the two is unsurprising, given that the colour pattern is controlled by just two independently segregating loci. Therefore, even first-generation backcross individuals can be phenotypically 'pure' but nonetheless have a very mixed genomic ancestry (25% introgressed). Conversely, single pattern alleles could introgress from one race to another over many generations, giving rise to individuals with a relatively 'pure' genome but a mixed phenotype. Both of these types of individuals occur in our samples, with some phenotypically hybrid individuals having 'pure' microsatellite genomes, and a few individuals indistinguishable phenotypically from parental forms with highly mixed microsatellite genomes. Nonetheless, in general there was considerable concordance between the assignment of individuals to the two forms and their colour pattern phenotype (Table S1, Fig. S2).

Bimodality represents a strong deficit of heterozygote individuals and LD between loci diagnostic for parental forms. The hybrid zone between *H. erato venus* and *H. erato chestertonii* shows several characteristics of a bimodal zone, with a marked deficit of heterozygotes at the loci controlling colour pattern, and a similar pattern at seven microsatellite loci. This contrasts with other hybrid zones within *H. erato* in which colour pattern loci generally do not deviate from Hardy–Weinberg equilibrium (Mallet & Barton 1989), but appears similar to the *H. erato/H. himera* hybrid zone in which they do (Jiggins *et al*. 1997). Moreover, the majority of the individuals were phenotypically similar to one or other parental form (74.4%) in the contact zone.

The distribution of hybrid classes is also informative regarding the form of reproductive isolation between these forms. All phenotypic hybrid individuals (25.6%) were classified as genotypic backcrosses or pure types based on microsatellite genotypes, suggesting either a low rate of F_1 formation via a strong parental assortative mating or a low survival rate of F_1 individuals due to strong selection. Even though 4% of individuals were classified as phenotypic F_1s, many of these are likely to be backcrosses with F_1-like phenotypes as they show microsatellite genotypes with a higher proportion of one or other parent (Table S1). A low frequency of F_1 individuals is known from other hybrid zones and is generally associated with strong assortative mating (Arnold & Bennett 1993; Goodman *et al*. 1999; Rieseberg *et al*. 1999; Ross & Harrison 2002; Gay *et al*. 2007). For instance, in a hybrid zone between Lousiana iris species, F_1 genotypes are extremely rare but backcrosses are relatively abundant (Arnold & Bennett 1993). In this case, there is very strong prezygotic isolation and a strongly bimodal adaptive landscape, in which backcross genotypes persist as they largely recover parental fitness in the different habitats (Arnold & Bennett 1993). Similarly, it seems therefore likely that there is strong assortative mating between *H. erato venus* and *H. erato chestertonii*. As in other examples, this is likely associated with selection against hybrids which reduces the success of the hybrids (Barton & Hewitt 1985; Mallet & Barton 1989; Bert & Arnold 1995; Crow *et al*. 2007). In this case, the association between the two races and

particular types of habitat, the extreme narrowness of the hybrid zone (≈ 4 km) and the strong differentiation observed between races suggest that both assortative mating and selection against hybrids are operating.

Furthermore, strong LD was observed between colour pattern and some of the microsatellite loci. Interestingly, these patterns were not predicted by patterns of chromosomal linkage between the loci (Table 4). Thus, Cr is in significant LD with four microsatellite loci all located on different chromosomes, while the linked loci Hel13 and Hel14 are not in LD with this colour pattern locus, despite being found on the same chromosome. However, none of the markers analysed here can be considered in tight linkage to colour patterns (Kapan *et al*. 2006), which strongly suggests that LD is generated by recent selection. On the other hand, unlike the colour pattern alleles, the microsatellite loci studied here are assumed to be neutral and are therefore expected to diffuse from one parental population to the other (Hewitt 1988). However, this gene flow of neutral alleles between parental populations is impeded by association with selected loci. The strength of this genetic barrier depends on intensity of selection (e.g. the fitness of hybrids) and the recombination rate between neutral and selected loci (Barton & Hewitt 1985). Given the lack of physical linkage, the strong LD between selected colour pattern loci and neutral microsatellite loci implies relatively strong reproductive isolation in the hybrid zone.

Furthermore, a phylogenetic analysis of mtDNA sequences also shows *H. erato venus* and *H. erato chestertonii* to be strongly differentiated, implying a lack of historical gene flow of mtDNA haplotypes at the contact zone (Fig. 3). Although our sample size within the hybrid zone is small, there was no evidence for the introgression of mtDNA haplotypes between the forms that might have been expected if recent introgression were frequent. These results parallel the situation in the *H. himera* × *H. erato cyrbia* hybrid zone where there was no evidence for mtDNA gene flow in a much larger sample size (Jiggins *et al*. 1997), and might indicate selection acting on the mtDNA perhaps related to altitudinal adaptation, reduced female F_1 fitness (Naisbit *et al*. 2002) or maternal effects on female F_1 mate choice. The two hybrid individuals sampled are also informative, as they possessed one haplotype each from the two parental forms. This implies that both reciprocal directions of hybridization occur naturally (male *venus* × female *chestertonii* and vice versa).

Implications for speciation

Here we aimed to test the prediction made previously that parapatric adaptation to dry forest habitats is correlated with speciation in the *H. erato* species complex. This prediction was derived from studies of a bimodal hybrid zone between the two species *H. himera* and *H. erato cyrbia* in southern Ecuador. The latter showed a lower frequency of hybrids as compared to the zone studied here (10% as compared to 25%), with correspondingly greater genetic differentiation and LD at nuclear loci, mtDNA, and colour pattern genes (Jiggins *et al*. 1996, 1997). More importantly, the ecological similarities between the two cases support our hypothesis that habitat adaptation plays a role in speciation. First, both *H. erato chestertonii* and *H. himera* are found at higher altitudes and in generally dryer habitats than most other *H. erato* populations. Second, virtually all populations of *H. erato* mimic *H. melpomene*, but *H. erato chestertonii* and *H. himera* are unusual in that neither mimics *H. melpomene*. *H. himera* is nonmimetic, while *H. erato chestertonii* mimics a race of *H. cydno*. This is presumably due to the absence of *H. melpomene* in the habitats of these two forms, but the switch in mimetic relationship might also have contributed to the evolution of reproductive isolation. In particular, in the case of *H. erato chestertonii*, this is the only form of *H. erato* that lacks any red or orange pattern elements. Given the role of colour pattern in mate finding in other *Heliconius* species, and in particular the attractiveness of red to most races of *H. erato*, it seems likely that this mimetic shift has contributed to the pre-mating isolation of *H. erato chestertonii* from *H. erato venus* in the Calima hybrid zone (Estrada & Jiggins 2008).

Thus, this hybrid zone fits well into the sequence of divergence from geographical races to species proposed by Mallet *et al*. (1998). The bimodality of the *H. erato venus* × *H. erato chestertonii* hybrid zone implies some degree of both pre- and postmating isolation, and laboratory experiments support this with both assortative mating and hybrid incompatibility found in crosses between these forms (A. Muñoz, in preparation). This contrasts with the *H. himera* and *H. erato cyrbia* hybrid zone where there is no evidence for intrinsic postmating isolation. Thus, the exact details of how reproductive isolation acts are likely to be different in the two cases, but in both cases divergence is associated with the same kinds of ecological differences, consistent with a causative role for ecological adaptation. This study therefore adds to a growing number of studies of ecotonal divergence (Schneider *et al*. 1999; Grill *et al*. 2007; Blum 2008). For example, in a group of Australian skinks morphological divergence was more likely between populations in adjacent divergent habitats than between populations separated in historical refugia (Schneider *et al*. 1999). A similar study of an African passerine, the little greenbul, reached the same conclusions regarding parapatric divergence in the face of gene flow (Smith *et al*. 2005). These studies support the long-established theoretical result that natural selection can promote divergence between adjacent populations despite ongoing gene flow (Endler 1977). Given the widespread observation of local adaptation to geographical variation in ecological conditions, it seems likely that parapatric divergence is a common cause of speciation.

Acknowledgements

We are very grateful to the Smithsonian Tropical Research Institute where part of the molecular work was carried out and for financial support through the internship program. We thank Maribel Gonzalez for help in the laboratory. We are grateful to four anonymous reviewers and the editor for comments on the manuscript. CA was funded by Facultad de Ciencias and Departamento de Ciencias Biológicas, Universidad de los Andes, and COLCIENCIAS. ML was partially funded from COLCIENCIAS, Banco de la República, and private donations: Continautos S.A., Proficol El Carmen S.A. and Didacol S.A. CJ was funded by a Royal Society University Research Fellowship and the BBSRC and Leverhulme Trust.

References

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, **160**, 1217–1229.

Arnold ML, Bennett BD (1993) Natural hybridization in Louisiana irises: genetic variation and ecological determinants. In: *Hybrid Zones and the Evolutionary Process* (ed. Harrison RG), pp. 115–139. Oxford University Press, New York.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature*, **341**, 497–503.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2001) *Genetix 4.02, Logiciel sous Windows TM pour la Genetique des Populations*. Laboratoire Genome, Populations, Interactions: CNRS UMR 5000, Universite de Montpellier II, Montpellier, France.

Beltrán M, Jiggins CD, Bull V *et al*. (2002) Phylogenetic discordance at the species boundary: gene genealogies in *Heliconius* butterflies. *Molecular Biology and Evolution*, **19**, 2176–2190.

Bert TM, Arnold WS (1995) An empirical test of predictions of two competing models for the maintenance and fate of hybrid zones: both models are supported in a hard-clam hybrid zone. *Evolution*, **49**, 276–289.

Blum MJ (2002) Rapid movement of a *Heliconius* hybrid zone: evidence for phase III of Wright's shifting balance theory? *Evolution*, **56**, 1992–1998.

Blum MJ (2008) Ecological and genetic associations across a *Heliconius* hybrid zone. *Journal of Evolutionary Biology*, **21**, 330–341.

Bridle JR, Baird SJE, Butlin R (2001) Spatial structure and habitat variation in a grasshopper hybrid zone. *Evolution*, **55**, 1832–1843.

Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences, USA*, **91**, 6491–6495.

Brower AVZ (1996) Parallel race formation and the evolution of mimicry in *Heliconius* butterflies: a phylogenetic hypothesis from mitochondrial DNA sequences. *Evolution*, **50**, 195–221.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution*, **21**, 550–570.

Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–631.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, Massachusetts.

Crow KD, Munehara H, Kanamoto Z *et al*. (2007) Maintenance of species boundaries despite rampant hybridization between three species of reef fishes (Hexagrammidae): implications for the role of selection. *Biological Journal of the Linnean Society*, **91**, 135–147.

Dempster A, Laird N, Rubin D (1977) Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society (B)*, **39**, 1–38.

Endler JA (1977) *Geographic Variation, Speciation, and Clines*. Princeton University Press, Princeton, New Jersey.

Estrada C, Jiggins CD (2008) Interspecific sexual attraction because of convergence in warning colouration: is there a conflict between natural and sexual selection in mimetic species? *Journal of Evolutionary Biology*, **21**, 749–760.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Flanagan NS, Davison A, Alamo M *et al*. (2002) Characterisation of microsatellite loci in neotropical *Heliconius* butterflies. *Molecular Ecology Notes*, **2**, 398–401.

Gay L, Neubauer G, Zagalska-Neubauer M *et al*. (2007) Molecular and morphological patterns of introgression between two large white-headed gull species in a zone of recent secondary contact. *Molecular Ecology*, **16**, 3215–3227.

Goodman SJ, Barton NH, Swanson G, Abernethy K, Pemberton JM (1999) Introgression through rare hybridization: a genetic study of a hybrid zone between red and sika deer (Genus *Cervus*) in Argyll, Scotland. *Genetics*, **152**, 355–371.

Grant B, Grant PR (2003) What Darwin's finches can teach us about the evolutionary origin and regulation of biodiversity. *Bioscience*, **53**, 965–975.

Grill A, Raijmann LEL, Van Ginkel W, Gkioka E, Menken SBJ (2007) Genetic differentiation and natural hybridization between the Sardinian endemic *Maniola nurag* and the European *Maniola jurtina*. *Journal of Evolutionary Biology*, **20**, 1255–1270.

Guo S, Thompson E (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Harrison RG (1993) *Hybrid Zones and the Evolutionary Process*. Oxford University Press, New York.

Harrison RG, Bogdanowicz SM (1997) Patterns of variation and linkage disequilibrium in a field cricket hybrid zone. *Evolution*, **51**, 493–505.

Hewitt GM (1988) Hybrid zones – natural laboratories for evolutionary studies. *Trends in Ecology & Evolution*, **3**, 158–167.

Howard DJ, Waring GL (1991) Topographic diversity, zone width, and the strength of reproductive isolation in a zone of overlap and hybridization. *Evolution*, **45**, 1120–1135.

Huelsenbeck JP, Ronquist FR (2001) MrBayes: Bayesian inference of phylogeny. *Bioinformatics*, **17**, 754.

Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution*, **15**, 250–255.

Jiggins CD, McMillan WO (1997) The genetic basis of an adaptive radiation; mimicry in two *Heliconius* sibling species. *Proceedings of the Royal Society B: Biological Sciences*, **B 264**, 1167–1175.

Jiggins CD, McMillan WO, Neukirchen W, Mallet J (1996) What can hybrid zones tell us about speciation? The case of *Heliconius erato* and *H. himera* (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society*, **59**, 221–242.

Jiggins CD, King P, McMillan WO, Mallet J (1997) The maintenance of species differences across a *Heliconius* hybrid zone. *Heredity*, **79**, 495–505.

Jiggins CD, Mavarez J, Beltran M *et al*. (2005) A genetic linkage map of the mimetic butterfly *Heliconius melpomene*. *Genetics*, **171**, 557–570.

Kapan DD, Flanagan NS, Tobler A *et al*. (2006) Localization of müllerian mimicry genes on a dense linkage map of *Heliconius erato*. *Genetics*, **173**, 735–757.

Linares M (1996) The genetics of mimetic coloration in the butterfly *Heliconius cydno weymeri*. *Journal of Heredity*, **87**, 142–149.

Maddison WP, Maddison DR (2001) *MacClade version 4.02: Analysis of Phylogeny and Character Evolution*. Sinauer & Associates, Sunderland, Massachusetts.

Mallet J (1986) Hybrid zones in *Heliconius* butterflies in Panama, and the stability and movement of warning colour clines. *Heredity*, **56**, 191–202.

Mallet J, Barton NH (1989) Strong natural selection in a warning color hybrid zone. *Evolution*, **43**, 421–431.

Mallet J, Barton N, Lamas G *et al*. (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, McMillan WO, Jiggins CD (1998) Mimicry and warning colour at the boundary between races and species. In: *Endless Forms: Species and Speciation* (eds Howard DJ, Berlocher SH), pp. 390–403. Oxford University Press, New York.

Mallet J, Beltrán M, Neukirchen W, Linares M (2007) Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC Evolutionary Biology*, **7**, 28.

Mavarez J, Gonzalez M (2006) A set of microsatellite markers for *Heliconius melpomene* and closely related species. *Molecular Ecology Notes*, **6**, 20–23.

Meglecz E, Petenian F, Danchin E *et al*. (2004) High similarity between flanking regions of different microsatellites detected within each of two species of Lepidoptera: *Parnassius apollo* and *Euphydryas aurinia*. *Molecular Ecology*, **13**, 1693–1700.

Merchan A, Jiggins CD, Linares M (2005) A narrow *Heliconius cydno* (Nymphalidae; Heliconiini) hybrid zone with differences in morph sex ratios. *Biotropica*, **37**, 119–128.

Naisbit R, Jiggins CD, Linares M, Salazar C, Mallet J (2002) Hybrid sterility, Haldane's rule and speciation in *Heliconius cydno* and *H. melpomene*. *Genetics*, **161**, 1517–1526.

Nei M (1972) Genetic distance between populations. *American Naturalist*, **106**, 283–292.

Nylander JAA (2004) *MrModeltest 2.0*. Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.

van Oosterhout CV, Hutchinson WF, Wills PM, Shipley PF (2003) *Micro-Checker 2.2.3*. The University of Hull, Hull, UK.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Raymond M, Rousset F (1995) GenePop (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Redenbach Z, Taylor EB (2003) Evidence for bimodal hybrid zones between two species of char (Pisces: *Salvelinus*) in northwestern North America. *Journal of Evolutionary Biology*, **16**, 1135–1148.

Rieseberg LH, Whitton J, Gardner K (1999) Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics*, **152**, 713–727.

Rolan-Alvarez E, Johannesson K, Erlandsson J (1997) The maintenance of a cline in the marine snail *Littorina saxatilis*: the role of home site advantage and hybrid fitness. *Evolution*, **51**, 1838–1847.

Ross CL, Harrison RG (2002) A fine-scale spatial analysis of the mosaic hybrid zone between *gryllus firmus* and *gryllus pennsylvanicus*. *Evolution*, **56**, 2296–2312.

Schneider CJ, Smith TB, Larison B, Moritz C (1999) A test of alternative models of diversification in tropical rainforests: Ecological gradients vs. rainforest refugia. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 13869–13873.

Schneider S, Roessli D, Excoffier L (2000) *Arlequin Ver.2000: A Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Sheppard PM, Turner JRG, Brown KS, Benson WW, Singer MC (1985) Genetics and the evolution of muellerian mimicry in *Heliconius* butterflies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **308**, 433–613.

Smith TB, Calsbeek R, Wayne RK *et al*. (2005) Testing alternative mechanisms of evolutionary divergence in an African rain forest passerine bird. *Journal of Evolutionary Biology*, **18**, 257–268.

Swofford DL (2000) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer & Associates, Sunderland, Massachusetts.

Tobler A, Kapan D, Flanagan NS *et al*. (2005) First-generation linkage map of the warningly colored butterfly *Heliconius erato*. *Heredity*, **94**, 408–417.

Vedenina VY, Helversen Ov (2003) Complex courtship in a bimodal grasshopper hybrid zone. *Behavioral Ecology and Sociobiology*, **54**, 44–54.

Weir BS (1996) *Genetic Data Analysis II*. Sinauer & Associates, Sunderland, Massachusetts.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Carlos F. Arias is interested in the speciation and evolution of neotropical butterflies. Astrid G. Muñoz works on the phylogeography and evolution of *Heliconius erato*. Chris D. Jiggins is interested in the study of adaption and speciation in natural populations of tropical butterflies. Jesus Mavarez works on ecology, population genetics and evolution of tropical organisms. Eldredge Bermingham's research focuses on molecular population genetics as well as historical biogeography of neo vertebrates, butterflies and caribbean island birds. Mauricio Linares research focuses on the evolutionary genetics and phylogeography of *Heliconius* butterflies, with special interest on the potential contribution of homoploid hybrid evolution to the diversity of these insects.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1 Summary of the phenotypes in crosses between *Heliconius erato chestertonii* and *Heliconius erato venus*. The top left and the bottom right phenotypes are *H. erato chestertonii* and *H. erato venus*. *Heliconius erato venus* has a red band on the forewing against a black background, while *H. erato chestertonii* has an entirely melanic forewing. This difference is controlled by the codominant alleles at the red band locus, *D*. On the hindwing, *H. erato venus* has a broad

yellow bar on the underside, a white margin visible on both upper and underside and a narrow yellow band on the overlapping region between the forewing and hindwing on the underside, while *H. erato chestertonii* has a broad yellow bar on both the upper and undersides. The codominant yellow bar locus, *Cr*, control the presence or absence of this broad yellow bar on the upperside of the hindwing, and the white hindwing margin. Finally, the upperside background colour of both *H. erato venus* and *H. erato chestertonii* is iridescent, with a slight difference in colour, the latter being more greenish. The central phenotype is an F_1. The two boxes show the backcross to *H. erato chestertonii* (blue line) and backcros to *H. erato venus* (red line) phenotypes. (A. Muñoz, in preparation).

Fig. S2 Bayesian population assignment test. Bar plots showing Bayesian assignment probabilities from the software NewHybrid for six different genotypic classes (*Heliconius erato venus*, *Heliconius erato chestertonii*, F_1, F_2, backcross to *H. erato venus* and backcross to *H. erato chestertonii*). Each vertical bar corresponds to one individual. Horizontal bar on the figure shows the collection localities and the phenotypic classification from colour pattern loci *D* and *Cr*.

Table S1 Concordance of colour pattern and genotypic classes. All individuals were classified in to phenotypic classes according to colour pattern of the wings and genotype according to the assignment of Structure and NewHybrid. Structure results give the probability of assignment to *Heliconius erato chestertonii* cluster. NewHybrid results give the probability of assignment to six genotype classes (V, BV, F_1, F_2, BC and C). V, *H. erato venus*; BV, backcross to *H. erato venus*; F_1; F_2; BC, backcross to *H. erato chestertonii*; C, *H. erato chestertonii*. Highlighted individuals show discordance between the different assignment methods.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.