

# The maintenance of species differences across a *Heliconius* hybrid zone

C. D. JIGGINS\*, W. O. McMILLAN, P. KING & J. MALLET

Department of Biology, University College London, The Galton Laboratory, 4 Stephenson Way,  
London NW1 2HE, U.K.

A contact zone between *Heliconius erato* and *H. himera* in southern Ecuador provides an opportunity to study the transition from races to species in Heliconiine butterflies. Genetic differentiation at 30 allozyme loci ( $D = 0.28$ ) is five times greater between *himera* and *erato* than among races of *erato* ( $D < 0.062$ ). Analysis of restriction fragment length polymorphisms in an 800-bp region of the mitochondrial genome shows fixed differences between the species. Despite 5–10 per cent hybridization per generation, these mtDNA, allozyme and colour pattern differences are in almost complete linkage disequilibrium throughout the contact zone. In mixed populations, there was no consistent evidence for convergence of *himera* and *erato* allozyme allele frequencies, and only four individuals out of 383 examined showed evidence for interspecific mtDNA gene flow. Linkage analysis of backcross broods showed that the 11 allozyme and two colour pattern loci, which are markedly divergent between the species, map to eight of 21 chromosomal linkage groups. Therefore, barriers to gene flow are not restricted to just a few strongly selected loci. Although analysis of population structure shows little evidence of interspecific gene flow, strong differences between allozyme loci in levels of divergence suggest that selection and gene flow may affect loci in different ways. Hybrid zones such as this, in which divergent genotypes coexist, should provide good model systems for the study of speciation.

**Keywords:** allozymes, interspecific hybridization, mitochondrial DNA, parapatry, population structure, speciation.

## Introduction

The claim that hybrid zones have provided 'natural laboratories' for the study of speciation (Barton & Hewitt, 1989) may be overly optimistic. Hybrid zones in organisms as diverse as mice, grasshoppers, crickets, sunfish, toads and butterflies separate highly differentiated taxa, but the morphological, ecological and behavioural differences between hybridizing forms are broken down in contact zone populations, which consist mainly of hybrids (Hunt & Selander, 1973; Avise & Smith, 1974; Butlin & Hewitt, 1985; Mallet, 1986; Szymura & Barton, 1986, 1991; Moore, 1987; Mallet & Barton, 1989; Rand & Harrison, 1989). In other words, speciation has not taken place and, given that many of these zones appear to have been stable over thousands of generations, it is not clear that speciation is imminent. Although such studies permit the estimation of (1) dispersal; (2) the strength and nature of selection; and (3) even the number of loci under selection (Szymura & Barton, 1991; Barton & Gale, 1993),

they tell us little about the coexistence of divergent genotypes.

Contact zones in which multilocus genotypic differences are maintained in mixed populations may be better natural laboratories for the study of speciation (e.g. Chu *et al.*, 1995). Needless to say, many of the selection pressures identified in traditional hybrid zone studies are likely to be important in speciation. For example,  $F_1$  males in crosses between *Chorthippus parallelus* races are completely sterile, i.e. 'Haldane's rule' (Haldane, 1922; Hewitt *et al.*, 1987). It is now important to make the link between the processes operating in hybrid zones and those involved in the coexistence of species.

*Heliconius* butterflies provide an opportunity to study the changes associated with the transition from randomly mating racial varieties to nonhybridizing species (Jiggins *et al.*, 1996; McMillan *et al.*, 1997). *Heliconius* spp. show striking divergence in warning colour both within and between species. For example, there are nearly 30 parapatric colour pattern races of *H. erato* across the New World tropics, which coexist with races of the Müllerian

\*Correspondence. E-mail: c.jiggins@ucl.ac.uk

co-mimic, *H. melpomene* (Brown, 1979). Boundaries between these races are abrupt clines at just a few loci, maintained by strong frequency-dependent selection (in some cases  $s \approx 0.5$ , Mallet & Barton, 1989; Mallet *et al.*, 1990). This strong selection on colour pattern loci generates only a weak barrier to gene flow across most of the genome. There is no evidence that races differ at allozyme loci, and geographical differentiation in mitochondrial haplotypes is not concordant with colour pattern boundaries (Turner *et al.*, 1979; Brower, 1996).

*Heliconius himera* is a sister taxon to *H. erato*, which appears to have evolved sufficient barriers to gene flow to allow stable coexistence throughout areas of overlap. *Heliconius himera* replaces *erato* in the dry forests of south-western Ecuador and northern Peru. The taxa meet in narrow zones where mixed populations consist of only 5–10 per cent phenotypically distinguishable hybrids (Jiggins *et al.*, 1996). This deficit of hybrids is primarily the result of strong mating preferences (McMillan *et al.*, 1997). Colour pattern differences between *himera* and *erato* are defined by only a few major loci, which are likely to be under strong stabilizing selection (Benson, 1972; Mallet & Barton, 1989; Jiggins & McMillan, 1997). It is conceivable, therefore, that the deficit of colour pattern hybrids observed in the *himera/erato* hybrid zone does not reflect a genome-wide pattern. Crosses between the two parental types, as well as first- and second-generation backcrosses and  $F_2$ s, produce fertile and viable offspring (McMillan *et al.*, 1997). A combination of persistent hybridization and weak postmating isolation may permit most of the genome to recombine freely, as occurs across racial boundaries.

Here, allozyme and mtDNA markers are used to investigate genetic differentiation between *himera* and *erato*. Although nothing was known about levels of differentiation at allozyme loci, a previous study suggested marked differences ( $\approx 1.5$  per cent sequence divergence) between mitochondrial cytochrome oxidase sequences of the two species (Brower, 1994). However, this conclusion was based on only two *himera* individuals collected from northern Peru, and it is unclear whether the observation reflects fixed species differences. Extensive sampling across the range of both species allows us to address the following questions.

- 1 Are there genetic differences between *himera* and *erato* in addition to colour pattern differences?
- 2 If there are genetic differences, is there strong linkage disequilibrium between allozyme, mtDNA and colour pattern loci across the contact zone?

3 From the patterns observed, what deductions can be made about the strength and nature of the barrier to gene flow?

## Methods

Butterflies were collected in Ecuador between January 1994 and April 1995. Wings were removed and the bodies frozen in liquid nitrogen and transported to the U.K. Collection sites correspond to those of Jiggins *et al.* (1996). Additional collections were made in allopatric populations 390 km (*cyrbia*; Alluriquín, Pichincha province) and 60 km (*himera*; Vilcabamba, Loja province) from the contact zone centre. Throughout this paper, 'allopatric populations' are defined as all populations that contain no individuals with colour pattern markers characteristic of the other species; 'sympatric populations' are all populations that contain colour pattern alleles of both species.

### Protein electrophoresis

A total of 259 individuals from allopatric populations and 411 individuals from mixed populations were analysed for electrophoretic variation. The thorax and abdomen of each individual was homogenized on ice in 80  $\mu$ L of grinding buffer (Mallet *et al.*, 1993; Emelianov *et al.*, 1995). The homogenate was centrifuged to remove cellular debris. Electrophoresis was performed on Helena cellulose acetate plates using two buffer systems: phosphate (0.36 per cent  $\text{NaH}_2\text{PO}_4$ , 0.22 per cent  $\text{Na}_2\text{HPO}_4$ , pH 6.3) and Tris-Gly (0.3 per cent Trizma, 1.4 per cent glycine, pH 8.6). Plates were run with samples of both species in order to ensure accurate scoring of alleles between runs. Plates were stained according to recipes described elsewhere (Mallet *et al.*, 1993; Emelianov *et al.*, 1995).

### mtDNA RFLP analysis

Total cellular DNA was extracted from either 50  $\mu$ L of protein homogenate in 250  $\mu$ L of STE buffer (0.05 M Tris buffer, 0.45 M NaCl, 0.1 M EDTA) or a small amount of frozen tissue (approximately half a thorax) homogenized in 300  $\mu$ L of STE buffer. Sodium dodecyl sulphate was added to make a final concentration of 0.5 per cent. The solution was then digested overnight at 55°C with 0.6 mg of proteinase K. After three extractions, first with phenol, then with phenol-chloroform and finally with chloroform, the DNA was precipitated with ethanol (Sambrook *et al.*, 1989).

An 800-bp region of mitochondrial DNA spanning the 3' end of the cytochrome oxidase subunit I (COI) gene, the leucine tRNA gene and most of the cytochrome oxidase subunit II (COII) was amplified using primer sequences 5'-CCATAATAAATTC-CAGGYCGGTAA-3' and 5'-GAAATATTATTT-CCTCTTTTGGATCC-3'. The PCR protocol used was 40 cycles with a cycling profile of 94°C for 30 s, 50°C for 30 s and 73°C for 1.5 min. In published sequences, there were 19 basepair substitutions between *himera* and *erato cyrbia* in this region (Brower, 1994, 1996). Two of these fell within the recognition sequences of the restriction enzymes *MunI* and *RcaI* (positions 689 and 455, respectively, of Appendix fig. 1, Brower, 1996). This allowed the simple diagnosis of individual haplotypes through analysis of differences in restriction fragment lengths (RFLP). The PCR product was digested overnight with the two restriction enzymes (0.1 unit  $\mu\text{L}^{-1}$ ) at 37°C. Digestion products were separated by electrophoresis on 1.5 per cent agarose gel and detected by staining with ethidium bromide.

#### Analysis of data

Calculation of genetic distance and tests for deviation from Hardy-Weinberg equilibrium were performed using BIOSYS-1 (Swofford & Selander, 1989). For sex-linked loci,  $\chi^2$ -tests for deviation from Hardy-Weinberg equilibrium were carried out excluding all hemizygous (i.e. female) individuals. Estimates of  $F_{ST}$  were calculated using GENEPOP (Raymond & Rousset, 1995) at all loci for which data were available in all populations. This program employs the method of Weir & Cockerham (1984) to calculate  $\theta$  as an estimate of  $F_{ST}$  and tests whether these estimates differ significantly from zero using a Fisher exact test.

Two further races of *H. erato*, *H. e. etylus* (Zamora, Zamora-Chinchipe province) and *H. e. lativitta* (Tena, Napo province), were sampled on the eastern slopes of the Andes. These were not included in the population study because of a lack of data at several loci, but were used to generate a neighbour-joining tree (Saitou & Nei, 1987) of the three races of *erato* (*etylus*, *cyrbia* and *lativitta*) and *H. himera*. The tree was drawn according to Swofford & Olsen (1990) using Nei's (1978) genetic distance, calculated for the 19 loci at which data were available (*GPI*, *GOT-s*, *GOT-f*, *PGM*, *MPI*, *MDH-s*, *3-P*, *G6-s*, *PK*, *AK*, *ENO*, *ACON-f*, *HBDH*, *ME*, *GPT*  $\alpha$ -*GPD*, *6-PGD*, *LA* and *SDH*). Genetic distance and allele frequency estimates were calculated between allopatric populations of *H. himera* and *H. erato*.

Linkage relationships between diagnostic allozyme markers were established using hybrid broods. It is relatively simple to produce crude linkage maps in *Heliconius* as there is no crossing-over in females (Turner & Sheppard, 1975). Thus, only a small number of progeny are necessary to determine whether two loci are syntenic, provided the female parent is a double heterozygote. In such crosses, any recombination implies that markers lie on different chromosomes. Protein electrophoresis for 11 differentiated loci was carried out on eight backcross broods, four  $F_1 \times \textit{himera}$  and four  $F_1 \times \textit{erato}$ , with a minimum of 24 progeny from each. Support for synteny or otherwise is deduced from at least one of these broods. The parents of each brood were run on the same plate as the offspring. The colour patterns of *Heliconius* butterflies are controlled by relatively few genes of large effect (Sheppard *et al.*, 1985; Mallet, 1989). In the case of *himera* and *erato*, two major loci have been identified (Jiggins *et al.*, 1997). These were also scored in hybrid broods to investigate linkage with allozyme loci.

#### Species index scores

None of the allozyme loci was completely fixed for different alleles between the species, so introgressed alleles were difficult to identify with any certainty. 'Species index' scores were devised to give populations or individuals a numerical score, which measures their genetic similarity with allopatric *himera* or *erato*. This was carried out by weighting each allele according to the species in which it was more frequent. Only alleles that showed either a greater than 0.1 frequency difference between the two species or those only found in one species were used. Alleles most common in *himera* were given a score of 0 and those in *erato* 1 (Table 1). Mean values were then calculated over all alleles to give a score between 0 and 1. These mean scores were calculated both for individuals (over all loci) and for each locus (summed over individuals within a population). For example, a three-locus genotype [*aa*, *ab*, *ab*], where *as* are *himera* alleles and *bs* are *cyrbia* alleles, would give that individual a species index score of  $(0 + 0 + 0 + 1 + 0 + 1)/6 = 0.33$ . Correlations between individual locus index scores and distance from the centre of the contact zone were investigated within each species, for all populations within 30 km of the contact zone.

#### Results

A total of 30 enzyme loci were scored (Table 1). Of these, eight loci were monomorphic, and at six more

**Table 1** Allele frequencies in combined allopatric populations of *Heliconius himera* and *H. erato cyrbia*, and in the two east Andean races, *H. e. lativitta* and *H. e. etylus*

Locus	Mobility	Index score	<i>H. himera</i>	<i>H. e. cyrbia</i>	<i>H. e. etylus</i>	<i>H. e. lativitta</i>	Locus	Mobility	Index score	<i>H. himera</i>	<i>H. e. cyrbia</i>	<i>H. e. etylus</i>	<i>H. e. lativitta</i>	Locus	Mobility	Index score	<i>H. himera</i>	<i>H. e. cyrbia</i>	<i>H. e. etylus</i>	<i>H. e. lativitta</i>
<b>Glucose-6-phosphate isomerase (<i>GPI</i>)</b>							<b>3-Phosphoglycerate dehydrogenase (<i>3-P</i>)</b>							<b>Superoxide dismutase (<i>SOD</i>)</b>						
( <i>n</i> )		NA	128	131	10	8	( <i>n</i> )			86	93	7	6	( <i>n</i> )			88	103	—	8
	255		—	0.061	—	—		150	0	0.413	0.016	0.286	0.583		—100	1	0.006	1.000	—	—
	144		0.035	0.137	—	0.063		100	1	0.587	0.984	0.714	0.417		—10	0	0.994	—	—	1.000
	120		—	—	—	0.063	<b>Glucose-6-phosphate dehydrogenase (<i>G6-s</i>)</b>							<b>Isocitrate dehydrogenase (NADP) (<i>IDH</i>)</b>						
	100		0.820	0.737	0.900	0.750	( <i>n</i> )			127	130	7	8	( <i>n</i> )		NA	64	76	—	8
	50		0.145	0.038	—	—		125	1	—	0.019	—	—		128		0.016	0.105	—	0.188
	33		—	0.023	0.100	0.125		100	1	0.134	0.969	1.000	1.000		117		0.328	0.046	—	—
	13		—	0.004	—	—		88	0	0.004	—	—	—		100		0.656	0.783	—	0.813
<b>Glutamic oxaloacetic transaminase (<i>GOT-f</i>)</b>								60	0	0.862	0.012	—	—		69		—	0.066	—	—
( <i>n</i> )		NA	128	131	10	8	<b>Pyruvate kinase (<i>PK</i>)</b>							<b>Malic enzyme (<i>ME</i>)</b>						
	120		—	0.031	—	—	( <i>n</i> )			128	131	3	8	( <i>n</i> )		NA	52	38	10	8
	110		—	—	—	0.188		155	—	—	—	0.167	0.563				—	0.039	—	—
	100		1.000	0.958	1.000	0.813		144	0	1.000	0.050	0.667	—		117		—	—	—	—
	72		—	0.011	—	—		100	1	—	0.950	0.167	0.438		104		0.010	0.118	0.050	—
	55		—	—	—	—	<b>Adenylate kinase (<i>AK</i>)</b>								100		0.981	0.803	0.950	1.000
<b>Glutamic oxaloacetic transaminase (<i>GOT-s</i>)</b>							( <i>n</i> )			126	130	10	8		89		0.010	0.039	—	—
( <i>n</i> )			128	131	10	8		115	0	0.381	0.012	—	—	<b>Malate dehydrogenase (<i>MDH-f</i>)</b>						
	—220	0	0.961	0.004	—	—		100	1	0.619	0.973	1.000	1.000	( <i>n</i> )		NA	45	23	—	8
	—200	1	—	0.023	—	0.188	<b>Enolase (<i>ENO</i>)</b>								144		—	0.022	—	0.188
	—180	1	—	0.042	—	0.063	( <i>n</i> )		NA	128	131	10	8		113		0.100	0.043	—	0.188
	—160	—	—	—	—	0.063		85	1	—	0.015	—	—		100		0.433	0.913	—	0.625
	—100	1	0.039	0.931	1.000	0.688	<b>Aconitase (<i>ACON-f</i>)</b>								91		0.467	0.022	—	—
<b>Phosphoglucumutase (<i>PGM</i>)</b>							( <i>n</i> )			128	131	10	8	<b>Glutamic pyruvic transaminase (<i>GPT</i>)</b>						
( <i>n</i> )			127	131	10	8		100		0.996	0.893	0.800	0.688	( <i>n</i> )		NA	60	52	10	8
	138	—	—	—	0.100	—	<b>Aconitase (<i>ACON-s</i>)</b>								130		—	0.010	—	—
	129	0	0.161	0.004	—	—	( <i>n</i> )			127	127	10	7		100		1.000	0.952	0.950	0.938
	118	1	0.004	0.118	0.400	0.313		113	1	—	0.008	—	—		77		—	0.029	—	—
	100	1	0.331	0.634	0.400	0.500		108	1	—	0.004	—	—		58		—	0.010	0.050	0.063
	92	—	—	—	0.050	—		104	1	0.012	0.260	—	—	<b><math>\alpha</math>-Glycerophosphate dehydrogenase (<math>\alpha</math>-<i>GPD</i>)</b>						
	82	—	0.244	0.229	0.050	0.188		100	0	0.984	0.661	1.000	0.929	( <i>n</i> )		NA	60	52	10	8
	68	0	0.079	0.015	—	—		94	1	0.004	0.054	—	0.071		113		0.008	—	—	—
	50	0	0.181	—	—	—		88	1	—	0.016	—	—		100		0.992	1.000	1.000	1.000
<b>Mannose-6-phosphate isomerase (<i>MPI</i>)</b>							<b><math>\beta</math>-Hydroxy-butyrate dehydrogenase (<i>HBDH</i>)</b>							<b>6-Phosphogluconate dehydrogenase (<i>6-PGD</i>)</b>						
( <i>n</i> )			128	130	10	8	( <i>n</i> )			124	131	10	8	( <i>n</i> )		NA	45	45	7	8
	119	1	—	0.004	—	—		200	0	0.044	—	—	—		120		0.022	—	—	—
	112	0	0.750	0.035	—	0.063		143	0	0.956	0.057	—	—		100		0.978	1.000	0.643	0.875
	108	1	0.016	0.392	0.150	0.125		100	1	—	0.924	0.900	0.750		83		—	—	0.357	0.125
	100	—	0.234	0.342	0.500	0.500	<b>Leu-Ala peptidase (<i>LA</i>)</b>							( <i>n</i> )		NA	21	16	—	—
	96	1	—	0.088	0.050	0.125				127	129	—	8		100		0.929	1.000	—	—
	92	1	—	0.046	0.150	0.125				124	131	10	8		90		0.071	—	—	—
	86	1	—	0.088	0.150	0.063	<b>Malate dehydrogenase (<i>MDH-s</i>)</b>							<b>Glucose-6-phosphate isomerase (<i>GPI</i>)</b>						
( <i>n</i> )		NA	127	128	10	8				127	129	—	8	( <i>n</i> )		NA	128	131	10	8
	180		—	0.008	—	—				127	129	—	8				—	0.061	—	—
	157		—	0.063	—	—				127	129	—	8				0.035	0.137	—	0.063
	100		1.000	0.926	1.000	1.000				127	129	—	8				0.820	0.737	0.900	0.750
	20		—	0.004	—	—				127	129	—	8				0.145	0.038	—	—

Loci in bold are those used to calculate estimated  $F_{ST}$  values (i.e. all polymorphic loci with a complete data set) and those in italics were used to calculate species index scores (i.e. all loci with a complete data set and marked species differences — see Methods for details). In addition, the following loci were scored but were found to be monomorphic across both species: fructose 1-6 diphosphatase; sorbitol dehydrogenase; isocitrate dehydrogenase (NAD); Leu-Gly-Gly peptidase; hexokinase; glyceraldehyde-3-phosphate dehydrogenase; glucose-6-phosphate dehydrogenase; acid phosphatase.

*himera* and *erato* shared a common allele with a frequency of  $>0.9$  in both species. The remaining 16 polymorphic loci showed some frequency or allelic differences between the species. Six loci were virtually fixed for different alleles in *himera* and *erato* (*SOD*, *PK*, *GOT-s*, *G6-s*, *ACON-s* and *HBDH*), whereas the remainder showed frequency differences between the species and in most cases rare alleles characteristic of one or other species. Overall levels of heterozygosity were not significantly different in the two species [ $0.139 \pm 0.040$  SE in *himera* and  $0.143 \pm 0.035$  SE in *erato*; Nei's (1978) unbiased estimates]. However, the mean number of alleles per locus was significantly lower in *himera* ( $1.9 \pm 0.2$ ) compared with *erato* ( $2.7 \pm 0.3$ ). The overall genetic distance between allopatric populations of *H. e. cyrbia* and *H. himera* was 0.278 [Nei's (1978) unbiased estimate] over all 30 loci. This estimate was nearly five times greater than the maximum distance ( $D = 0.062$ ) between geographical races of *H. erato* (Fig. 1).

Deviation from Hardy-Weinberg equilibrium was examined within each species, at each of the 14 loci used for population structure analysis (Table 1), excluding colour pattern hybrids. Eleven out of 153 tests showed significant deviations from Hardy-Weinberg expectations in *H. e. cyrbia*, and six out of 109 tests in *H. himera*. These deviations were almost invariably caused by occasional homozygotes for rare alleles with very low expected values, which may have been mis-scored individuals. When  $\chi^2$ -values were summed within loci or populations, there were no significant deviations from Hardy-Weinberg equilibrium [using the 'pooled genotype test' where appropriate (Swofford & Selander, 1989)]. Most allozyme data sets show a few significant deviations from Hardy-Weinberg equilibrium (e.g. Mallet *et al.*, 1993; Emelianov *et al.*, 1995) but, if there is no consistent pattern across loci or populations, these

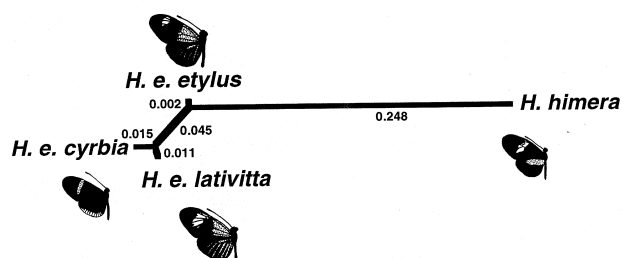
deviations are unlikely to result from pervasive effects, such as selection, nonrandom mating or hybridization.

There was no evidence for population structuring within either species. Estimates of  $F_{ST}$  ( $\theta$ ) within each species were, without exception, very low and not significantly different from zero. Overall values of  $\theta$  were 0.0060 in *H. erato cyrbia* and 0.0003 in *H. himera*. Although most of the samples were from within 30 km of the contact zone, this estimate also includes samples 390 km (*H. e. cyrbia*) and 60 km (*H. himera*) from the contact zone.

In contrast, there is a sharp genetic break between *himera* and *erato* with no consistent evidence for convergence of species gene frequencies in the contact zone (Fig. 2). Correlation analysis of species index score against distance from the centre of the zone showed only two significant associations out of 22 tests (11 loci in each species). These were *3-P* in *H. himera* ( $r = 0.698$ ,  $P = 0.025$ ) and *ACON-s* in *H. e. cyrbia* ( $r = 0.578$ ,  $P = 0.05$ ). Both of these were in the right direction to suggest introgression, i.e. gene frequencies became more similar to the other species in populations near the contact zone. Although the correlation coefficient was significant in *ACON-s*, the slope of the regression line was very shallow (Fig. 2j). In contrast, the pattern at *3-P* does suggest that there was consistent introgression at this locus, although it only occurred in one direction (from *erato* into *himera*). There was, therefore, only evidence in one test out of 22 that hybridization has significantly affected the geographical distribution of allele frequencies. Furthermore, the combined data showed no significant effects (*cyrbia*:  $r = 0.296$ , NS; *himera*:  $r = 0.366$ , NS).

Indeed, the distribution of individual species index values in sympatric populations (Fig. 3) forms a bimodal distribution. The phenotypically 'pure species' form two distinct genetic clusters with the colour pattern hybrids lying between the two. If these hybrids were to be removed from the plot, there would be little difference between the allopatric and sympatric genotypic distributions of *himera* and *erato* (Fig. 3).

Linkage analysis of *himera*  $\times$  *erato* hybrid broods has identified three linkage groups among the 11 loci which show marked species differences [*AK-G6-s-GOT-s*]; [*PK-ACON-A*]; [*3-P-SOD*] (Table 2). The latter group is X-linked. In addition, *ACON-s* is linked to a colour pattern locus, *Cr*. The lack of recombination in female *Heliconius* (Turner & Sheppard, 1975) means that these linkage groups, and other unlinked loci, can be confidently allocated to different chromosomes with the brood sizes



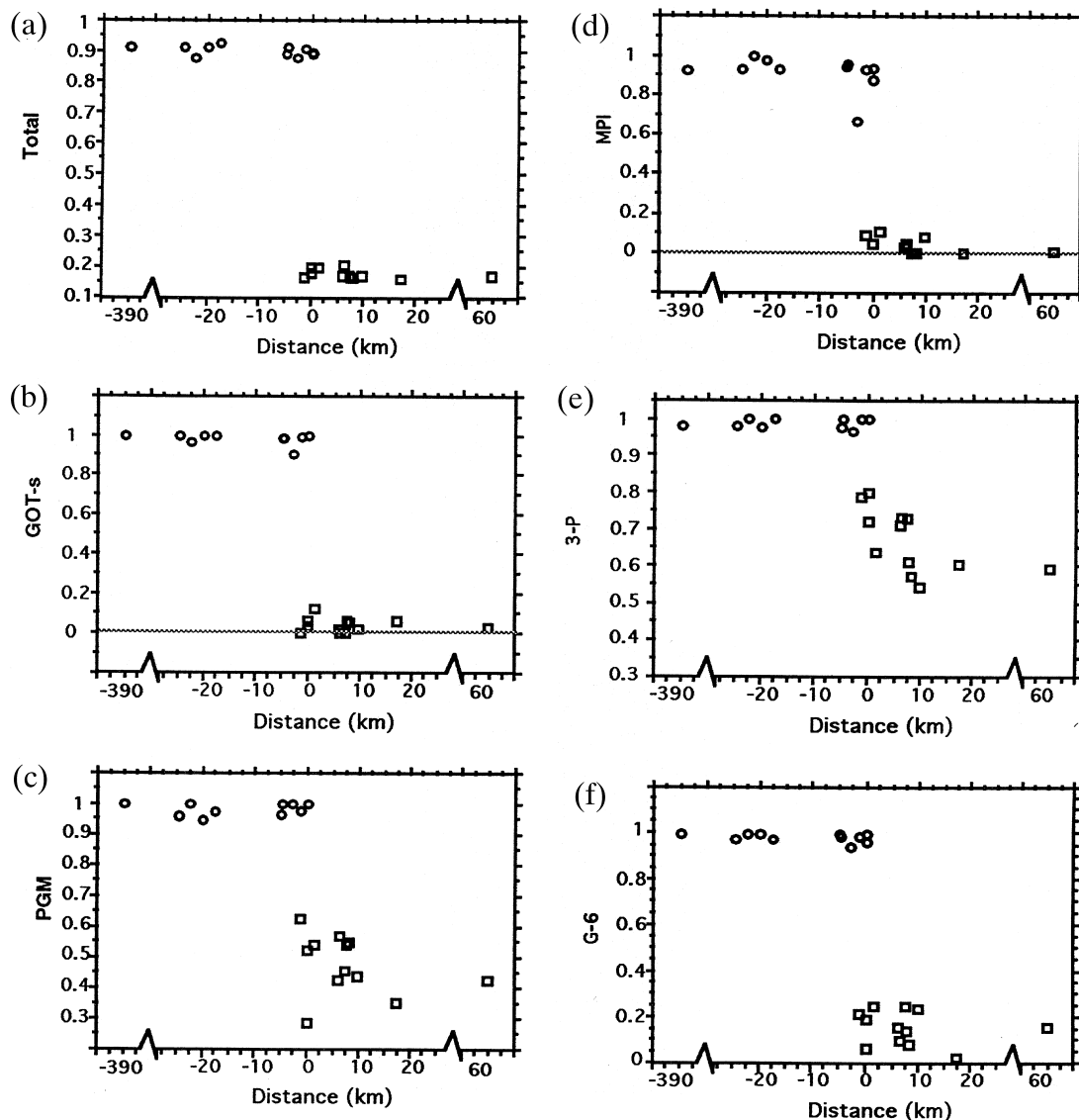
**Fig. 1** Neighbour-joining tree of *Heliconius himera* and three races of *H. erato*. *Heliconius erato cyrbia* is from the western Andes, whereas *H. e. lativitta* and *etylus* are from the eastern slopes. Nei's (1978) genetic distances were calculated over 19 loci (see Methods).

analysed here ( $n > 24$ ). Thus, the 11 allozyme and two colour pattern loci described represent eight chromosomal linkage groups of the 21 in *Heliconius erato* (Brown *et al.*, 1992).

The mtDNA data show a similarly abrupt break between *himera* and *erato*. There was a fixed mtDNA haplotype difference between the species in allopatric populations, and only four *himera* and no *erato* individuals had introgressed haplotypes in sympatric populations (Table 3). The allozyme genotypes of these introgressed individuals suggest that

three of the four could be first-generation backcrosses to *himera* (species index scores of the four individuals are 0.143, 0.263, 0.286 and 0.318). The pattern in hybrids was also informative. Of 11  $F_1$  phenotypes, nine had an *erato* haplotype and two a *himera* haplotype, which implies that hybridization occurs in both directions.

There was no evidence for heteroplasmy in any individuals. Two allopatric *cyrbia* individuals showed patterns which implied the loss or gain of a restriction site, but retained a characteristic fragment



**Fig. 2** Population species index values plotted against distance from the centre of the hybrid zone. Each point represents a single population sampled. Circles are *Heliconius erato cyrbia* and squares are *H. himera*. (a) Combined values over all loci; the remaining plots show each locus separately. Only two loci, 3-P in *himera* and ACON-s in *cyrbia*, show significant correlations of gene frequency with distance, a pattern that may imply introgression.

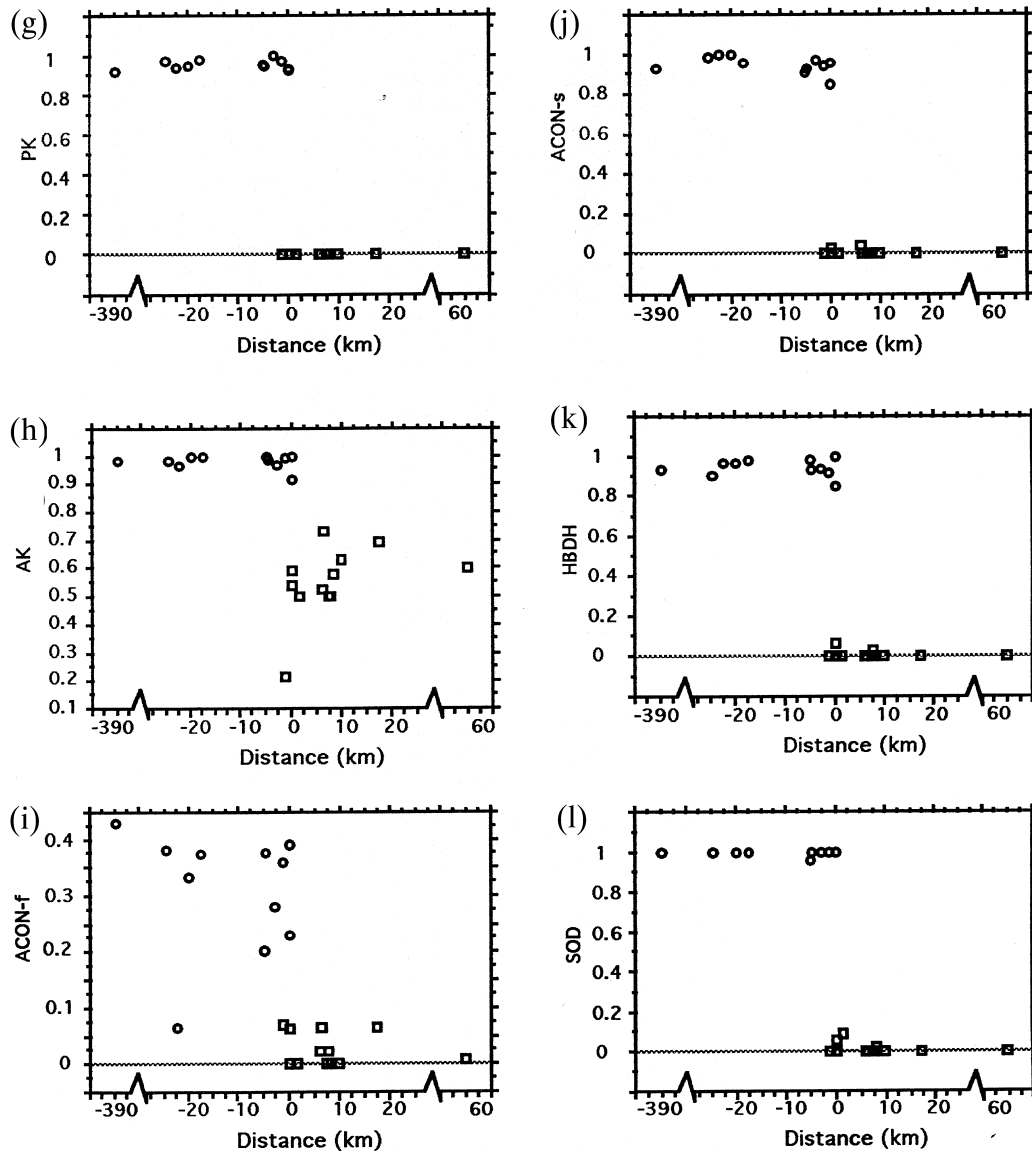


Fig. 2 continued

length and so could be confidently scored as *cyrbia* haplotypes. One allopatric *cyrbia* individual showed a banding pattern that could not be confidently assigned to either haplotype.

### Discussion

An important characteristic of a species is the maintenance of genetic integrity when coexisting with related taxa (Templeton, 1989). Colour pattern differentiation between *himera* and *erato* is associated with a geographically concordant break at nuclear ( $D = 0.28$ ) and cytoplasmic loci. Furthermore, despite persistent hybridization and no hybrid

sterility or inviability (McMillan *et al.*, 1997), *Heliconius himera* and *erato* maintain their genetic differences in sympatry. Throughout the contact zone, there is almost complete linkage disequilibrium between colour pattern, mtDNA and allozyme markers. Linkage analysis shows that these multilocus genetic differences map to eight of the 21 chromosomes in *H. erato*. This implies that barriers to gene flow are dispersed across the genome and not just associated with a few strongly selected colour pattern loci.

The abrupt genetic break between *himera* and *erato* is far larger than any differentiation acquired

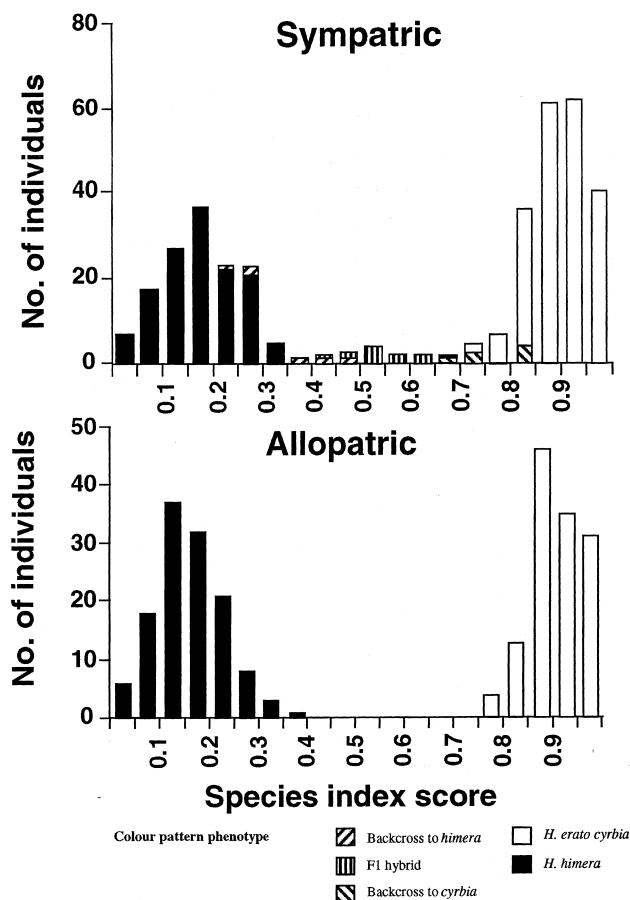
solely through isolation by distance. Across western Ecuador, no evidence for population structure within either species was found. This is largely in agreement with previous studies of *H. erato*, which showed no large frequency or fixed differences between populations from as far apart as Trinidad, Panama, western Ecuador and Peru (Turner *et al.*, 1979; Silva & Araújo, 1994; J. Mallet, unpublished observations). The allozyme cluster analysis shows that races of *erato* from both sides of the Andes cluster together with  $D \approx 0.06$  (Fig. 1). In contrast, the genetic break between *himera* and *erato* is five times greater and occurs over the space of just a few kilometres.

The barrier to gene flow is highlighted by the lack of mtDNA introgression. Because cytoplasmic markers are less likely to be linked to genes involved in reproductive, behavioural or ecological differ-

ences, they might be expected to introgress more readily than nuclear markers (Barton & Jones, 1983). This prediction is supported by a number of studies, which have shown introgression of mtDNA across strong barriers to nuclear gene flow, such as hybrid zones or even between reproductively isolated sympatric species (Takahata & Slatkin, 1984; Gyllenstein *et al.*, 1985). However, there is little evidence for introgression in *himera* and *erato*, in which only four introgressed mtDNA haplotypes were found among 383 individuals sampled from mixed populations.

Despite the abrupt transition at many unlinked loci, there is some evidence for differential effects of selection among allozyme loci and even gene flow across this species boundary. Allozyme loci fall into three distinct groups (Fig. 4), with six showing almost fixed differences between *himera* and *erato* ( $\theta > 0.8$ ), five showing intermediate differentiation ( $0.2 < \theta < 0.35$ ) and the remaining 11 little or no differentiation ( $\theta < 0.15$ ). Under neutral theory, the rate of evolution and, hence, divergence between populations is directly proportional to mutation rate (Kimura, 1983). When loci are compared within a population, thus eliminating variation in population size, levels of heterozygosity will reflect mutation rate. If the patterns observed at different loci were solely a result of neutral divergence, those loci with higher levels of heterozygosity would be expected to show the greatest differentiation. This is not the case and there is no relationship between divergence ( $\theta$ ) and overall heterozygosity ( $r = 0.02$ ; NS), which suggests that the differences between loci cannot be explained by neutral theory alone. Similar patterns have been observed in studies of sibling species, such as *Drosophila pseudoobscura* and *D. persimilis* (Prakash, 1976), and may indicate that loci are subject to different evolutionary pressures (Takahata & Slatkin, 1984). For example, faster rates of divergence might result from disruptive selection on protein differences (Feder *et al.*, 1997), whereas slower rates of divergence could be caused by either gene flow across the species barrier or stabilizing selection on polymorphisms (Watt, 1983; Watt *et al.*, 1983; Bancroft *et al.*, 1995).

Marked differences between loci are also evident in phylogenetic comparisons of *himera* with the races of *erato*. MtDNA data imply a similar level of divergence between west and east Andean *erato* as between *erato* and *himera* (1.5–2 per cent sequence divergence in the mitochondrial COI and COII genes; Brower, 1994). In contrast, allozyme data show *himera* to be strongly differentiated from all the races of *erato* (Fig. 1). The most likely reasons



**Fig. 3** Distribution of individual species index values in all hybrid zone populations combined (sympatry) and nonhybrid zone populations (allopatry). Black bars are *Heliconius himera*, unshaded bars are *H. erato cyrbia* and hatched bars are colour pattern hybrids.



**Table 2** Linkage relationships between species-diagnostic genetic markers used in this study

	1	2	3	4	5	6	7	8	9	10	11	12	13
1 <i>SOD</i>	x	—	—	—	—	—	—	+	—	—	—	—	—
2 <i>GOT-s</i>		x	—	—	+	—	+	—	—	—	—	—	—
3 <i>PGM</i>			x	—	—	—	—	—	—	—	—	—	—
4 <i>MPI</i>				x	—	—	(—)	—	—	—	—	(—)	(—)
5 <i>G6-s</i>					x	—	+	(—)	(—)	(—)	—	—	—
6 <i>PK</i>						x	—	—	+	—	—	—	—
7 <i>AK</i>							x	(—)	(—)	—	—	—	—
8 <i>3-P</i>								x	—	—	—	—	—
9 <i>ACON-f</i>									x	—	—	—	—
10 <i>ACON-s</i>										x	—	—	+
11 <i>HBDH</i>											x	—	—
12 <i>D</i>												x	—
13 <i>Cr</i>													x

+, Markers shown to be linked; —, markers shown to be unlinked. Markers 12 and 13 are colour pattern loci (Jiggins & McMillan, 1997). Data in brackets were not tested explicitly but are inferred from other linkage relationships. *SOD* and *3-P* are sex linked.

**Table 3** Concordance of colour pattern and mtDNA markers

mtDNA	Colour pattern phenotype				
	<i>himera</i>	Backcross to <i>himera</i>	F <sub>1</sub>	Backcross to <i>erato</i>	<i>erato</i>
Sympatric populations					
<i>himera</i> haplotype	147	6	2	0	0
<i>erato</i> haplotype	4	0	9	9	206
Allopatric populations					
<i>himera</i> haplotype	125	—	—	—	—
<i>erato</i> haplotype	—	—	—	—	110

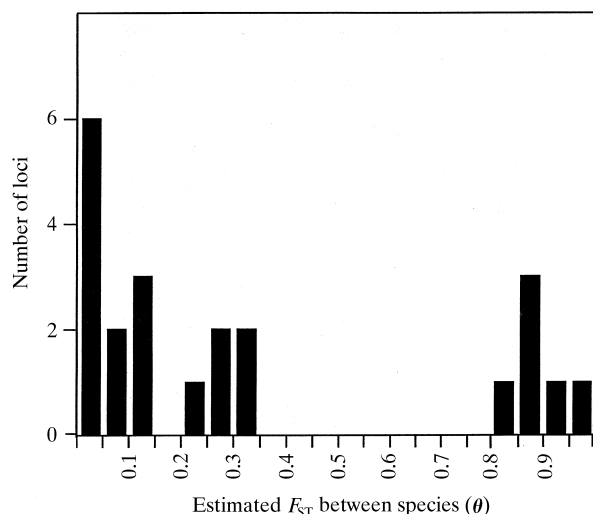
All individuals are classified into hybrid classes according to colour pattern phenotype (Jiggins *et al.*, 1996). The mtDNA haplotype of each individual was determined using an RFLP analysis.

for this discordance are, first, stabilizing selection on allozymes within *erato* (e.g. Karl & Avise, 1992), and secondly, disruptive selection on protein loci between *himera* and *erato*. Disruptive selection is likely to result from adaptation to novel environments, such as the shift from dry to wet forest which occurs across the *himera/erato* hybrid zone (Jiggins *et al.*, 1996). In the future, gene genealogies of protein-coding loci will provide a promising means of differentiating the effects of selection and gene flow among loci.

The apparent strength of the barrier to gene flow between *H. himera* and *H. erato* is perhaps surprising in the light of the observed level of hybridization. However, the lack of completely fixed species differ-

ences and the marked differences in rates of divergence between loci probably indicate that some interspecific gene flow occurs. The barrier to gene flow is sufficient to allow the accumulation of neutral or divergently adaptive differences, but may be permeable to the spread of globally advantageous alleles.

The pattern described provides a marked contrast, not only with racial hybrid zones in *H. erato*, but with most other hybrid zones described in the literature, in which even highly differentiated taxa mate randomly and the genomes recombine freely when in contact. For example, when the distribution of allozyme hybrid index scores is plotted for a *Bombina* hybrid zone, there is a unimodal distribu-



**Fig. 4** Estimated interspecific  $F_{ST}$  values at all polymorphic loci. These were calculated between combined allopatric populations of *Heliconius himera* and *H. erato cyrbia*. The trimodal pattern is not compatible with divergence through random drift (see Discussion) and suggests that at least some of these loci must be influenced by selection or gene flow across the species boundary.

tion (Szymura & Barton, 1986; Barton & Gale, 1993). Similarly, in the centre of hybrid zones between *Mus musculus musculus* and *M. m. domesticus*, there is no deviation from Hardy–Weinberg equilibrium at allozyme loci (Hunt & Selander, 1973). In contrast, across the *himera/erato* contact zone, parental genotypes coexist despite hybridization, which allows identification of the processes that prevent homogenization of gene pools in the early stages of species formation. In the case of *himera* and *erato*, the barrier to gene flow is most probably a result of divergence in mate preferences, warning colour and ecology without hybrid inviability or sterility (Jiggins *et al.*, 1996; McMillan *et al.*, 1997). Examples such as this can truly be considered ‘natural laboratories’ for the study of speciation.

### Acknowledgements

Many thanks to Pablo Andrade, Ashleigh Griffin and Angus Davison who helped collect butterflies and Yvonne Graneau for help in the laboratory. Also to INEFAN for granting permission to carry out research in Ecuador; the Museo de Ciencias Naturales in Quito and Fundación Arcoiris in Loja for their support. This project is funded by a BBSRC research grant and studentship.

### References

- AVISE, J. C. AND SMITH, M. H. 1974. Biochemical genetics of sunfish. I. Geographic variation and subspecific intergradation in the bluegill, *Lepomis macrochirus*. *Evolution*, **28**, 42–56.
- BANCROFT, D. R., PEMBERTON, J. M., ALBON, S. D., ROBERTSON, A., MACCOLL, A. D. C., SMITH, J. A. *ET AL.* 1995. Molecular genetic variation and individual survival during population crashes of an unmanaged ungulate population. *Phil. Trans. R. Soc. B*, **347**, 263–273.
- BARTON, N. H. AND GALE, K. S. 1993. Genetic analysis of hybrid zones. In: Harrison, R. G. (ed.) *Hybrid Zones and the Evolutionary Process*, pp. 13–45. Oxford University Press, New York.
- BARTON, N. H. AND HEWITT, G. M. 1989. Adaptation, speciation and hybrid zones. *Nature*, **341**, 497–503.
- BARTON, N. AND JONES, J. S. 1983. Mitochondrial DNA: new clues about evolution. *Nature*, **306**, 317–318.
- BENSON, W. W. 1972. Natural selection for Müllerian mimicry in *Heliconius erato* in Costa Rica. *Science*, **176**, 936–939.
- BROWER, A. V. Z. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 6491–6495.
- BROWER, A. V. Z. 1996. Parallel race formation and the evolution of mimicry in *Heliconius* butterflies: a phylogenetic hypothesis from mitochondrial DNA sequences. *Evolution*, **50**, 195–221.
- BROWN, K. S. 1979. *Ecologia Geográfica e Evolução nas Florestas Neotropicais*. 2 vols. (Livre de Docencia) Universidade Estadual de Campinas, Campinas, Brazil.
- BROWN, K. S., JR, EMMEL, T. C., ELIAZAR, P. J. AND SUOMALAINEN, E. 1992. Evolutionary patterns in chromosome numbers in neotropical Lepidoptera. I. Chromosomes of the Heliconiini (Family Nymphalidae. Subfamily Nymphalinae). *Heredity*, **117**, 109–125.
- BUTLIN, R. K. AND HEWITT, G. M. 1985. A hybrid zone between *Chorthippus parallelus parallelus* and *Chorthippus parallelus erythropus* (Orthoptera: Acrididae): morphological and electrophoretic characters. *Biol. J. Linn. Soc.*, **26**, 269–285.
- CHU, J., POWERS, E. AND HOWARD, D. J. 1995. Gene exchange in a ground cricket hybrid zone. *J. Hered.*, **86**, 17–21.
- EMELIANOV, I., MALLET, J. AND BALTENSWEILER, W. 1995. Genetic differentiation in *Zeiraphera diniana* (Lepidoptera: Tortricidae, the larch budmoth): polymorphism, host races or sibling species? *Heredity*, **75**, 416–424.
- FEDER, J. L., ROETHELE, J. B., WLAZLO, B. AND BERLOCHER, S. H. 1997. Reproductive isolation, fitness tradeoffs and temperature: sympatric race formation in the apple maggot fly. *Proc. Natl. Acad. Sci. U.S.A.*, (in press).
- GYLLENSTEN, U., WHARTON, D. AND WILSON, A. C. 1985. Maternal inheritance of mitochondrial DNA during backcrossing of two species of mice. *J. Hered.*, **76**, 321–324.
- HALDANE, J. B. S. 1922. Sex ratio and unisexual sterility in

- hybrid animals. *J. Genet.*, **12**, 101–109.
- HEWITT, G. M., BUTLIN, R. K. AND EAST, T. M. 1987. Testicular dysfunction in hybrids between parapatric subspecies of the grasshopper *Chorthippus parallelus*. *Biol. J. Linn. Soc.*, **31**, 25–34.
- HUNT, W. G. AND SELANDER, R. K. 1973. Biochemical genetics of hybridization in European house mice. *Heredity*, **31**, 11–33.
- JIGGINS, C. D. AND McMILLAN, W. O. 1997. The genetics of mimicry in two *Heliconius* sibling species. *Proc. R. Soc. B*, (in press).
- JIGGINS, C., McMILLAN, W. O., NEUKIRCHEN, W. AND MALLET, J. 1996. What can hybrid zones tell us about speciation? The case of *Heliconius erato* and *H. himera* (Lepidoptera: Nymphalidae). *Biol. J. Linn. Soc.*, **59**, 221–242.
- KARL, S. A. AND AVISE, J. C. 1992. Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. *Science*, **256**, 100–102.
- KIMURA, M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.
- McMILLAN, W. O., JIGGINS, C. D. AND MALLET, J. 1997. What initiates speciation in passion vine butterflies? *Proc. Natl. Acad. Sci., U.S.A.*, (in press).
- 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. 1989. The genetics of warning colour in Peruvian hybrid zones of *Heliconius erato* and *H. melpomene*. *Proc. R. Soc. B*, **236**, 163–185.
- MALLET, J. AND BARTON, N. H. 1989. Strong natural selection in a warning color hybrid zone. *Evolution*, **43**, 421–431.
- MALLET, J., BARTON, N., LAMAS, G., SANTISTEBAN, J., MUEDAS, M. AND EEEY, 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., KORMAN, A., HECKEL, D. G. AND KING, P. 1993. Biochemical genetics of *Heliothis* and *Helicoverpa* (Lepidoptera: Noctuidae) and evidence for a founder event in *Helicoverpa zea*. *Ann. Entomol. Soc. Am.*, **86**, 189–197.
- MOORE, W. S. 1987. Random mating in the northern flicker hybrid zone: implications for the evolution of bright and contrasting plumage patterns in birds. *Evolution*, **41**, 539–546.
- NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.
- PRAKASH, S. 1976. Genetic divergence in closely related sibling species, *Drosophila pseudoobscura*, *D. persimilis* and *D. miranda*. *Evolution*, **31**, 14–23.
- RAND, D. M. AND HARRISON, R. G. 1989. Ecological genetics of a mosaic hybrid zone: mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. *Evolution*, **43**, 432–449.
- RAYMOND, M. AND ROUSSET, F. 1995. GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.*, **86**, 248–249.
- SAITOU, N. AND NEI, M. 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.
- SAMBROOK, J., FRITSCH, E. F. AND MANIATIS, T. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, NY.
- SHEPPARD, P. M., TURNER, J. R. G., BROWN, K. S., BENSON, W. W. AND SINGER, M. C. 1985. Genetics and the evolution of Muellierian mimicry in *Heliconius* butterflies. *Phil. Trans. R. Soc. B*, **308**, 433–613.
- SILVA, L. M. AND ARAUJO, A. M. 1994. The genetic structure of *Heliconius erato* populations (Lepidoptera; Nymphalidae). *Rev. Bras. Genet.*, **17**, 19–24.
- SWOFFORD, D. L. AND OLSEN, G. J. 1990. Phylogeny reconstruction. In: Hillis, D. M. and Moritz, C. (eds) *Molecular Systematics*, 1st edn, pp. 411–501. Sinauer Associates, Sunderland, MA.
- SWOFFORD, D. L. AND SELANDER, R. B. 1989. *BIOSYS-1. A Computer Program for the Analysis of Allelic Variation in Population Genetics and Biochemical Systematics*, Release 1.7. Illinois Natural History Survey, Champaign, IL.
- SZYMURA, J. M. AND 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**, 1141–1159.
- SZYMURA, J. M. AND 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**, 237–261.
- TAKAHATA, N. AND SLATKIN, M. 1984. Mitochondrial gene flow. *Proc. Natl. Acad. Sci. U.S.A.*, **81**, 1764–1767.
- TEMPLETON, A. R. 1989. The meaning of species and speciation: a genetic perspective. In: Otte, D. and Endler, J. A. (eds) *Speciation and its Consequences*, pp. 3–27. Sinauer Associates, Sunderland, MA.
- TURNER, J. R. G. AND SHEPPARD, P. M. 1975. Absence of crossing-over in female butterflies (*Heliconius*). *Heredity*, **34**, 265–269.
- TURNER, J. R. G., JOHNSON, M. S. AND EANES, W. F. 1979. Contrasted modes of evolution in the same genome: allozymes and adaptive change in *Heliconius*. *Proc. Natl. Acad. Sci. U.S.A.*, **76**, 1924–1928.
- WATT, W. B. 1983. Adaptation at specific loci. 2. Demographic and biochemical elements in the maintenance of the *Colias* PGI polymorphism. *Genetics*, **103**, 691–724.
- WATT, W. B., CASSIN, R. C. AND SWAN, M. S. 1983. Adaptation at specific loci. III. Field behaviour and survivorship differences among *Colias* PGI genotypes are predictable from *in vitro* biochemistry. *Genetics*, **103**, 725–739.
- WEIR, B. S. AND COCKERHAM, C. C. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.