First-generation linkage map of the warningly colored butterfly *Heliconius erato*

A Tobler[1], D Kapan[1], NS Flanagan[1], C Gonzalez[1], E Peterson[1], CD Jiggins[2], JS Johntson[3], DG Heckel[4,5] and WO McMillan[1]

[1]*Department of Biology, University of Puerto Rico, PO Box 23360, San Juan 00931, Puerto Rico;* [2]*Institute of Cell, Animal & Population Biology, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JT, Scotland, UK;* [3]*Department of Entomology, Texas A&M University, College Station, TX 77843, USA;* [4]*CESAR, Department of Genetics, The University of Melbourne, Parkville, Victoria 3010 Australia*

We report the first genetic linkage map of *Heliconius erato*, a species that shows remarkable variation in its warningly colored wing patterns. We use crosses between *H. erato* and its sister species, *H. himera*, to place two major color pattern genes, *D* and *Cr*, on a linkage map containing AFLP, allozyme, microsatellite and single-copy nuclear loci. We identified all 21 linkage groups in an initial genetic screen of 22 progeny from an F1 female × male *H. himera* family. Of the 229 markers, 87 used to identify linkage groups were also informative in 35 progeny from a sibling backcross (*H. himera* female × F1 male). With these, and an additional 33 markers informative in the second family, we constructed recombinational maps for 19 of the 21 linkage groups. These maps varied in length from 18.1 to 431.1 centimorgans (cM) and yielded an estimated total length of 2400 cM. The average distance between markers was 23 cM, and eight of the 19 linkage groups, including the sex chromosome (Z) and the chromosome containing the *Cr* locus, contained two or more codominant anchor loci. Of the three potential candidate genes mapped here, *Cubitus interruptus* (*Ci*), *Decapentaplegic* (*Dpp*) and *Wingless* (*Wg*), only *Ci* was linked, although loosely, to a known *Heliconius* color pattern locus. This work is an important first step for constructing a denser genetic map of the *H. erato* color pattern radiation and for a comparative genomic study of the architecture of mimicry in *Heliconius* butterflies.

Introduction

Heliconius butterflies exhibit a remarkable geographic mosaic in wing color patterns that warn predators of the butterflies' unpalatability (Punnett, 1915; Emsley, 1964; Turner, 1976; Brown, 1981). Several noxious *Heliconius* species display nearly identical warning color patterns: a phenomenon known as Müllerian mimicry (Müller, 1879). Within any given area, the wing patterns of these aposematic butterfly species converge on one of a handful of clearly differentiated types (usually six or less), called mimicry rings (Mallet and Gilbert, 1995). This local color pattern convergence is overlaid on a pattern of geographic divergence, with the color patterns of species within local mimicry rings often changing dramatically across their Neotropical distributions (Bates, 1862; Turner, 1983; Sheppard *et al*, 1985). Geographic differentiation is particularly marked in *H. erato*, a member of the 'red/orange' mimicry ring and a common inhabitant of gaps and edge habitat in wet forests of Central and South America. The wing patterns of *H. erato*, which are comprised of large patches of red, orange, yellow and/or white on a black or iridescent blue/green background, show spectacular divergence into nearly 30 geographic races.

The myriad of racial variation within *H. erato* and its relatives, coupled with the apparently simple genetic control of wing patterns and varying degrees of genetic differentiation among species and geographic races, make these butterflies excellent model systems to study adaptation, morphological differentiation and speciation (Sheppard *et al*, 1985; Mallet *et al*, 1990; Brower, 1994; Jiggins and McMillan, 1997; McMillan *et al*, 1997; Joron *et al*, 1999 (2001); Jiggins *et al*, 2001; Kapan, 2001; Beltrán *et al*, 2002; Gilbert, 2003; Naisbit *et al*, 2003; Flanagan *et al*, 2004). Yet, in spite of the long history of ecological and evolutionary work, the *Heliconius* genome remains poorly characterized. However, now that advances in molecular techniques make it easier to isolate molecular markers in nonmodel organisms, we can begin to take a genome-wide approach to addressing some long-standing questions in the ecology and evolutionary biology of *Heliconius*.

Here, using an efficient interspecific crossing design, we present the first genetic linkage analysis of *H. erato*. The principal goal of this study is to develop a long-term mapping strategy and genetic toolbox for detailed analysis of the major gene loci implicated in the *H. erato* color pattern radiation. Linkage analysis of *H. erato* is crucial to understanding the genetic architecture behind adaptive change in wing patterns. Wing pattern

Correspondence: *WO McMillan, Department of Biology, University of Puerto Rico, PO Box 23360, San Juan 00931, Puerto Rico.*
E-mail: wmcmilla@rrpac.upr.clu.edu
[5]*Current address: Max Planck Institute of Chemical Ecology, Hans-Knöll-Str. 8, Jena D-07745, Germany.*

evolution in *H. erato* is a classic example of adaptation proceeding via changes at a few 'major' genetic loci and serves as a prime example for current theories for the origin of mimicry (Turner, 1983; Nijhout, 1991; Mallet and Joron, 1999). Over total of 30 years of interracial crosses have identified about five loci of major effect that control the differences between races of *H. erato* and over 20 more that have a more minor effect on wing patterning (Sheppard *et al*, 1985; Jiggins and McMillan, 1997). Furthermore, a genetic map of the *H. erato* radiation will allow the quick determination of the linkage relationship between an expanding number of potential candidate loci (McMillan *et al*, 2002) and the genes known to cause color pattern change in this species. Finally, linkage analysis of *H. erato* will enable comparative study of the genetics of the parallel mimetic color pattern radiation of *H. erato* and its distant relative *H. melpomene* (Brower, 1996). These two species are relatively distant relatives of one another within the genus, yet show a nearly identical intraspecific color pattern radiation (Punnett, 1915; Emsley, 1964; Turner, 1976; Brown, 1981).

In this study, we exploited crosses between *H. erato* and the closely related *H. himera* to provide a common genetic background on which to study the genes responsible for the color pattern radiation of *H. erato*. *H. himera* is parapatric to *H. erato* and is found in mid-elevation dry scrub forests of southern Ecuador and northern Peru (Jiggins *et al*, 1997). Apart from wing patterns, *H. himera* shows only slight morphological differentiation from *H. erato* and was initially considered another geographic race of the latter (Brown, 1981). However, more recent behavioral and genetic work demonstrates its species level status and suggests that *H. himera* diverged from the main *H. erato* color pattern radiation within the last 1.5–2.0 million years through adaptation to its dry forest habitat (Jiggins *et al*, 1997; McMillan *et al*, 1997; Beltrán *et al*, 2002; Flanagan *et al*, 2004). Our initial linkage analysis of *H. erato* uses two sibling backcross broods of *H. e. cyrbia* and *H. himera* (Figure 1). The two taxa are completely interfertile and alleles for two major color pattern loci, *D* and *Cr*, segregate in our backcross progeny and in the field (Jiggins and McMillan, 1997). The *D* locus, likely homologous to the D^{Ry} complex (see Sheppard *et al*, 1985), controls both red pattern elements of the fore- and hindwing in *H. erato* and *H. himera* (Figure 1). The *Cr* locus controls yellow and white pattern elements, including the yellow forewing band of *H. himera* and the white fringe on *H. e. cyrbia* (Figure 1).

Our crossdesign took advantage of the fact that oogenesis is achiasmatic in Lepidoptera (ie there is no crossing over during meiosis in females; Suomalainen *et al*, 1973; Turner and Smiley, 1975). Thus, female-informative markers on the same chromosome are inherited in complete linkage and can be used to positively identify linkage groups (Yasukochi, 1998). We used dominant AFLP markers to generate the many loci required to identify all 21 *Heliconius* linkage groups expected based on karyotyping in *Heliconius* ($n = 21$; Suomalainen *et al*, 1973; Turner and Smiley, 1975; Brown *et al*, 1992). The AFLP technique is technically robust and is now commonly employed in linkage and QTL analysis of a wide range of organisms (Yasukochi, 1998; Hawthorne, 2001; Tan *et al*, 2001; Parsons and Shaw,
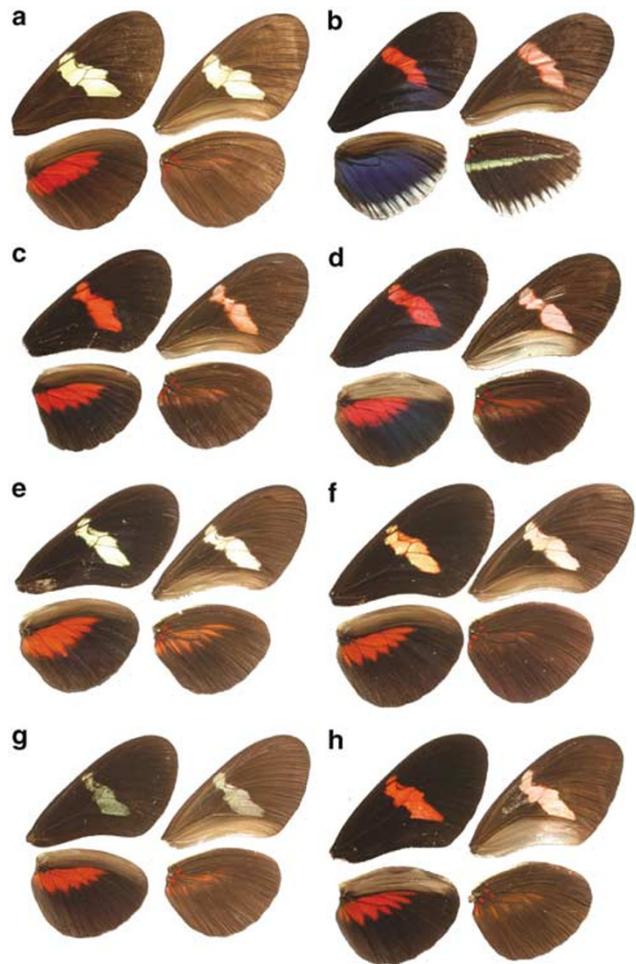
Figure 1 We crossed an *H. himera* female (**a**) with an *H. e. cyrbia* male (**b**). The resulting full sib F1 parents male (**c**) and female (**d**) were backcrossed to *H. himera* pure types female and male, respectively (not shown). Offspring segregate into phenotypic classes ranging from homozygous *H. himera* types (**e**) to individuals heterozygous for both alleles at the *Cr* locus and the *D* locus (h, see also Jiggins and McMillan, 1997). The forewing band coloration is affected by epistasis (see Discussion) and ranges from partially red and partially yellow (rarely red and white) (**f**) to nearly *H. himera* like with a smattering of black scales over the forewing yellow bar (**g**). Presumed genotypes as scored here are as follows: $Cr^{hi} Cr^{hi}D^{hi} D^{hi}$ (**e**), $Cr^{hi} Cr^{hi}D^{hi} D^{cyr}$ (**f**), $Cr^{hi} Cr^{cyr}D^{hi} D^{hi}$ (**g**) and $Cr^{hi} Cr^{cyr}D^{hi} D^{cyr}$ (**h**).

2002; Sandal *et al*, 2002). In addition to AFLPs, we used a suite of codominant anchor loci, which include allozymes, microsatellites and single-copy nuclear loci (SCNL) developed for this project. Codominant markers are essential to our goal of producing a fully integrated map of *H. erato* containing all of the major and minor color pattern loci underlying the color pattern radiation within *H. erato* and for comparative linkage analysis of mimicry in butterflies.

Methods

Genetic stocks, larval rearing, and crossing strategy
Stocks of *H. erato* and *H. himera* were collected in Loja Province in southern Ecuador under a permit from the

Ecuadorian government. Our focal F1 brood was produced by mating a newly eclosed female *H. himera* to a male *H. e. cyrbia* (Figure 1). The newly mated female was housed in a 2 m³ cage (1 m (l) × 1 m (w) × 2 m (h)) with ample larval host plants and adult feeding plants. Eggs were collected daily and larvae were raised in individual plastic pots according to the procedure outlined in (McMillan *et al*, 1997). Two F1 siblings were backcrossed to either a male or female *H. himera* stock (Broods 63 and 67, respectively). The offspring of these crosses were raised until emergence, when their bodies were frozen at −70°C until needed and wings were preserved in glassine envelopes and later digitally photographed for wing pattern analysis. Our sibling cross design maximized the number of AFLP markers shared between the two backcross families to facilitate the integration of linkage groups and recombination distances into a single map (see Linkage analysis and map construction below).

Genomic DNA extraction and AFLPs

Genomic DNA was extracted from individuals using a simple organic extraction method outlined in Beltrán *et al* (2002) based on a method first described by Harrison *et al* (1987). Approximately 150 mg of thorax ($\frac{1}{3}$ of the total thorax) was placed into 1.5 ml centrifuge tubes and homogenized in 500 µl of Insect Grinding Buffer (10 mM Tris-HCl (pH 8.0); 100 mM EDTA (pH 8.0), 0.5% SDS). The homogenate was digested for 3 h at 55°C with 0.5 mg of proteinase K. The homogenate was then extracted three times, (1) with equal volumes of buffered phenol, (2) with equal volumes of buffered phenol/chloroform/isoamyl alcohol (25:24:1) and (3) with equal volumes of chloroform/isoamyl alcohol (24:1). The DNA was then precipitated with 2× volume of 95% ethanol and $\frac{1}{10}$ volume of 3.0 M NaOAC (pH 5.2). Genomic DNA was resuspended in 100 µl 0.1 × TE buffer.

In general, we followed the procedures described in Vos *et al* (1995) for the generation of AFLP fingerprints. We digested 60–100 ng of genomic DNA with the restriction enzymes *Eco*RI and *Mse*I or *Taq*I to yield genomic DNA fragments in the 100–600 bp size range. Adapter sequences were ligated to these DNA fragments to provide primer sites for PCR reactions (see Table 1). To reduce the complexity of our AFLP fingerprints, we performed a two-step PCR procedure. In the first, or

preselective step, we used a 1:10 dilution of the restriction–ligation reaction as template DNA for a PCR reaction with primers that contained a single-base extension. The preselective PCR cycle was 94°C for 30 s, 56°C for 1 min, and 72°C for 1 min and was repeated for 30 cycles. In the second or selective step, we used a 1:20 dilution of the preselective reaction as template for a PCR reaction with primers that contained either a two- or a three-base extension.

Wild-caught *H. erato* and *H. himera* were initially screened with 48 different AFLP primer combinations. This screen allowed us to estimate the genetic differences between the two species at AFLP marker loci and identify the most informative subset of primer combinations. Based on these patterns, we determined the AFLP fingerprint for grandparents, parents and backcross offspring using several radioactive and nonradioactive labeling methods. Depending on visualization system (see below), the *Eco*RI primer was either end-labeled with [γ-³²P]ATP as described in Vos *et al* (1995), with one of three fluorescent tags (Hex, Tet, Fam, Applied Biosystems) or with IRDye™ 700 near-infrared markers (LI-COR Inc.). The selective PCR amplification was carried out in a touchdown PCR reaction. The first phase consisted of 12 cycles of a step-down PCR: 94°C for 30 s, 65°C for 30 s (−0.7°C each cycle) and 72°C for 1 min. These cycles were followed by an additional 23 cycles with a 56°C annealing temperature. All adaptors and core primers sequences are listed in Table 1.

For radioactively labeled AFLP reactions, we loaded 1.5 µl of each PCR product onto a 6% denaturing acrylamide gel. Samples were electrophoresed at 100 W, 50 mA, 50°C for 60–90 min. AFLP banding patterns were revealed by exposure of X-ray films (Kodak X-OMAT, Sigma Chemical Co.). Fluorescent-tagged bands were visualized on the ABI PRISM® 377 DNA sequencer (Applied Biosystems). We loaded 1.5–2.0 µl of each PCR product on a 6% acrylamide gel with TAMRA 500 size standards (Applied Biosystems). Samples were electrophoresed at 100 W, 50 mA, 50°C for 2 h and 45 min. AFLP banding patterns were size corrected in each lane trace using Genescan (version 3.1, Applied Biosystems). Size corrected data were imported into a modified version of Genographer, an open source java program for AFLP analysis (http://hordeum.oscs.montana.edu/genographer/) to visualize all individuals for one

Table 1 Oligonucleotide adaptors and primers used for AFLP analysis

	Core	Enzyme	Extension
*Eco*RI Adaptor-f	5′-CTCGTAGACTGCGTACC-3′		
*Eco*RI Adaptor-r	5′-AATTGGTACGCAGTCTAC-3′		
*Mse*I Adaptor-f	5′-GACGATGAGTCCTGAG-3′		
*Mse*I Adaptor-r	5′-TACTCAGGACTCAT-3′		
*Taq*I Adaptor-f	5′-GACGATGAGTCCTGAG-3′		
*Taq*I Adaptor-r	5′-CGCTCAGGACTCAT-3′		
*Eco*RI Preselective	5′-GACTGCGTACC	AATTC	C-3′
*Mse*I Preselective	5′-GATGAGTCCTGAG	TAA	C-3′
*Taq*I Preselective	5′-GATGAGTCCTGAG	CGA	C-3′
*Eco*RI Selective	5′-GACTGCGTACC	AATTC	CN-3′
*Mse*I Selective	5′-GATGAGTCCTGAG	TAA	CNN-3′
*Taq*I Selective	5′-GATGAGTCCTGAG	CGA	CNN-3′

primer combination in a 'metagel' similar in format to the 96-lane radioactive gels. A similar procedure was carried out for additional primer combinations labeled with IRDye™ 700 near-infrared dye. In the latter case, we loaded 0.8 μl of a 1:10 dilution of each PCR product on a 10% acrylamide gel to achieve fine resolution. Samples were electrophoresed on an LI-COR NEN 4200 global DNA Analyzer (LI-COR Inc.) at 35 W, 35 mA, 50°C for 90–150 min with 50–700 sizing standards (LI-COR Inc.). LI-COR gels were analyzed using Saga MX 3.0 AFLP software (LI-COR Inc.) and two gels for each primer combination were combined for final analysis using the display function. In all analyses of radioactive-, fluorescent- or IRDye-labeled AFLPs, only loci showing a band present in one of the parents and absent in the other parent were used.

For radioactive gels, AFLP markers were named according to primer and band number in the gel, for example, CG/CCC-3 is a fragment generated with the primer combination *Eco*RI-CG/*Mse*I-CCC that corresponded to the third fragment analyzed. All primer combinations had an *Eco*RI primer and either *Mse*I or *Taq*I primer. On nonradioactive gels, band names included the primer combination (as above) and were numbered with the band size in base pairs (see Supplementary Appendix). This facilitated direct com-

parison between data generated with the same primer combination on the ABI 377 and the Licor machine (not shown).

Codominant loci

We followed the segregation pattern of a variety of codominant markers (Table 2). Primers for two house-keeping enzymes (*Tpi* and *Mpi*) were developed as described in Beltrán *et al* (2002), and the three additional SCNL are – *Cubitus interruptus* (*Ci*), *Dopa decarboxylase* (*Ddc*), and *Decapentaplegic* (*Dpp*). The gene *Wingless* (*Wg*) was also scored using primers published in Brower and Egan (1997). SCNL were amplified from the parents and backcross offspring of our two target broods. Amplification conditions on genomic DNA were carried out as outlined in Table 2. The amplified fragments for the parental individuals were sequenced and analyzed for allele-specific polymorphism. The fragments amplified from the genomic DNA of each backcross progeny were digested with diagnostic restriction enzymes, separated by electrophoresis in a 3% agarose/1 × TBE gel and analyzed for allele segregation within the progeny. In addition, we examined the inheritance patterns of 11 microsatellite loci developed for this study (Flanagan *et al*, 2002) and an additional locus first reported here. Microsatellites were amplified from genomic DNA for all

Table 2 PCR primers and amplification conditions for codominant loci used in this study

Locus	Primer name	Primers sequence	Annealing temp (°C)	Intron?
Mpi	*Mpi 14*	5′-TTTAAGGTGCTCTATATAACRAARGC-3′	50	Yes
	Mpi 15	5′-TTCTGGTTTGTGATTTGGATCYTTRTA-3′		
Tpi	*Tpi 1*	5′-GGTCACTCTGAAAGGAGAACCATCTT-3′	50	Yes
	Tpi 2	5′-CACAACATTTGCCCAGTTGTTGCCAA-3′		
Wg	*Wg F*	5′-CCCAGTTTTAGATCTGTCG-3′	50	No
	Wg R	5′-TCTCGGTCCGGTATCCGCG-3′		
Dpp	*Dpp F34*	5′-AGAGAACGTGGCGAGACACTG-3′	55	No
	Dpp R327	5′-GAGGAAAGTTGCGTAGGAACG-3′		
Ci	*Ci F190*	5′-ATGCGGAGACATACTGGTGAA-3′	53	Yes
	Ci R350	5′-GATGCTTAGCTCTGTCACTGG-3′		
Ddc	*Ddc F125*	5′-CAAGCTCATTCGTCTGTCGAG-3′	57	Yes
	Ddc R291	5′-CGAAAGCGCAAGACGATGTAG-3′		
He-CA-001	*A25F*	5′-CGGGACACAAAATGTGACTG-3′	55	N/A
	A25R	5′-A-ACTGGACGGTGGACGTTAG-3′		
Hel-01	*B6R*	5′-TCGTAGATATCCATTACTCTGGTCTG-3′	54	N/A
	B6F	5′-AGGGCGTCGTTAGTTTGTGT-3′		
Hel-08	*A2F*	5′-ACATCTCAGAACTGGTCGGC-3′	55	N/A
	A2R	5′-CTCGATCAGCCGGTGATTAT-3′		
Hel-13	*CA13F*	5′-ATTTCATAGTAACGCCCTCC-3′	52	N/A
	CA13R	5′-TGACTTATCGCTAAGGTCAA-3′		
Hel-14	*AAT4F*	5′-GCACATTTACTTACACTAACGC-3′	53	N/A
	AAT4R	5′-ATTTGTTTCGAACGACTGCC-3′		
Hel-04	*GA-8F*	5′-GGAACGGAGTGCCCTAAAAC-3′	55	N/A
	GA-8R	5′-CGTTGCCGCTTATACTTTCC-3′		
Hel-05	*GA-11F*	5′-TGCTGTCCATACCCAACTCA-3′	55	N/A
	GA-11R	5′-CGAACTCACAACCATCAGTCA-3′		
Hel-07	*GA-19R*	5′-CCCGAACCACTAAAGTCGAA-3′	55	N/A
	GA-19F	5′-GCGGGGACAACTACATAAGC-3′		
Hel-11	*CA-10F*	5′-TTTCTTTTGAGTCCCGATGG-3′	55	N/A
	CA-10R	5′-ATCTCAGAACTGGTCGGCAG-3′		
Hel-10	*CA-9F*	5′-TCTCACTTTCCCACACAGCA-3′	55	N/A
	CA-9R	5′-TGTGAAGAGACACATGGGGA-3′		
Hel-03	*GA-7F*	5′-CCAATTATGTCACATGGATCTGTT-3′	54	N/A
	GA-7R	5′-CTCTGTCGTCCTCTGCAGTC-3′		
Hel-06	*GA-16F*	5′-TAGCCTTCACTTTGAACCCG-3′	55	N/A
	GA-16R	5′-CCCACTCGAAGCAATGAAAT-3′		

N/A = not applicable.

individuals using fluorescent-labeled primers (Table 2), according to the PCR protocols outlined in Flanagan *et al* (2002). PCR amplification products were resolved together with the TAMRA-500 size standard in the ABI 377 Sequencer (Applied Biosystems), on a 5% Thermo-PAGE acrylamide denaturing gel (Amresco, K191) with 6 M urea at 3000 V for 3 h. Finally, electrophoretic mobility was examined in 14 enzyme systems on Helena acetate plates as described previously (Jiggins *et al*, 1997). The enzymes examined were as follows: aconitase (ACON), adenylate kinase (AK), B-hydroxy-butyrate dehydrogenase (HBDH), enolase (ENO), glucose-6-phosphate dehydrogenase (G-6), glucose-6-phosphate isomerase (GPI), glutamate oxalacetate transaminase (GOT-F and GOT-S), isocitrate dehydrogenase (IDH), Leu-Ala peptidase (LA), malate dehydrogenase (MDH), mannose-6-phosphate isomerase (MPI), phosphogluco-mutase (PGM) and pyruvate kinase (PK).

Linkage analysis and map construction

A Pascal program (written by DGH) was used to identify linkage groups in the female-informative Brood 63. The program, AFLP63, looks for groups of cosegregating, nonrecombinant AFLP bands, while allowing for a small number of genotyping errors, to ensure that such errors do not artificially inflate the number of independently segregating groups. After linkage groups were identified in Brood 63, any AFLPs that were shared with the male-informative Brood 67 were used to define discrete chromosomes in that family. These loci, as well as novel AFLPs segregating only in Brood 67, were analyzed using Mapmaker/Exp 3.0 (Lincoln *et al*, 1993) using the following strategy. First, anchor loci were assigned to chromosomes. Second, the other male-informative loci were assigned to chromosomes if linkage was detected at LOD 4.0 or greater. Finally, the most likely order and spacing of markers on a chromosome were determined. For eight or fewer markers, all possible orders were examined using the 'compare' command, and the most likely order chosen. For nine or more markers, a preliminary 'scaffold' of eight markers was first constructed, then the other markers were inserted into the scaffold such that the likelihood of the resulting map was maximized. Genetic distance was expressed in terms of Haldane centimorgans (cM). All codominant loci were recoded so that segregation of the informative codominant allele was converted to the corresponding segregation pattern of a dominant allele.

Scoring color pattern variation

Mendelian color pattern variation at the *D* and *Cr* loci was scored in accordance with the model of Jiggins and McMillan (1997), which specifies the nine genotypes among the F_2 progeny of the interspecific cross. The *D* locus controls the basal red bar on the dorsal hindwing of *H. himera* ($D^{hi} D^{hi}$, Figure 1a) and the narrow red bar on the dorsal forewing of *H. e. cyrbia* ($D^{cyr} D^{cyr}$, Figure 1b). The *Cr* locus specifies the yellow forewing of *H. himera* ($Cr^{hi}Cr^{hi}$) and the white trailing edge of the hindwing in *H. e. cyrbia* ($Cr^{cyr} Cr^{cyr}$). In the F_1 hybrid (Figure 1c and d), the *H. himera* hindwing characters are dominant, and thus provide no information to distinguish the four genotypes segregating in our backcrosses to *H. himera*. However, because of interaction between *D* and *Cr*, it

was possible to assign genotypes using the scale colors in the forewing bar as illustrated: Figure 1e, yellow like *H. himera*, ($Cr^{hi}Cr^{hi}D^{hi}D^{hi}$); Figure 1g, yellow and black, ($Cr^{hi}Cr^{cyr}D^{hi}D^{hi}$); Figure 1f, more yellow than red ($Cr^{hi}Cr^{hi}D^{hi}D^{cyr}$); and Figure 1h, mostly red like the F_1 ($Cr^{hi}Cr^{cyr}D^{hi}D^{cyr}$). The latter two were frequently difficult to distinguish. For each locus, the *H. e. cyrbia* allele was coded as dominant for the purposes of the mapping program.

Genome size estimation

We estimated the size of the *H. erato* genome using flow cytometry using the methods outlined in Bennett *et al* (2003). Briefly, the brain of a single *H. e. petiverana* and the head of a single *D. melanogaster* Iso-1 female standard (1C = 175 megabases (Mb)) were placed into 1 ml of cold Galbraith buffer (Galbraith *et al*, 1983) in a 1.5 ml Kontes Dounce Tissue Grinder and stroked 15 times with an A pestle, and filtered through a 20-µm nylon filter. Propidium iodide was added to a final concentration of 50 ppm, and the mixture costained in the dark at 4°C for 20–40 min. The mean fluorescence of costained nuclei in five replicate samples of each sex was quantified using a Coulter Epics Elite (Coulter Electronics, Hialeah, FL, USA) with a laser tuned at 514 nm and 300 mW. Fluorescence was detected by a photomultiplier screened by a long-pass filter. DNA content was determined by comparing the ratio of the 2C mean of the sample with the 2C mean for *Drosophila* (1C = 175 Mb).

Results

An initial screening with 24 *Eco*RI-CN/*Mse*I-CNN and 24 *Eco*RI-CN/*Taq*I-CNN primer combinations on six wild-caught *H. erato* and *H. himera* generated an average of 80 AFLP bands per gel, of which 88% were polymorphic (data not shown). Of the polymorphic bands, 26% were fixed in either the *H. erato* or *H. himera* individuals and absent in the other species. These results are very similar to previous studies, which show strong differences between *H. erato* and *H. himera* using allozymes (Jiggins *et al*, 1997) and nuclear and mitochondrial sequence data loci (Beltrán *et al*, 2002; Flanagan *et al*, 2004).

Based on our initial screening, we examined AFLP fingerprints in our female-informative backcross family (Brood 63) using 11 *Eco*RI-CN/*Mse*I-CNN and six *Eco*RI-CN/*Taq*I-CNN primer combinations (Table 1, see Supplementary Appendix). These combinations generated 227 maternally inherited polymorphic markers of which 95% segregated in the expected 1:1 ratio. The number of markers showing segregation distortion was lower than that observed in the silkworm moth (Tan *et al*, 2001), but may be an artifact of our limited power to detect segregation distortion given our small brood size. Of the polymorphic markers, 202 could be assigned to 21 linkage groups corresponding to the 21 chromosomes expected based on previous cytogenetic work on *Heliconius* (Suomalainen *et al*, 1973; Turner and Smiley, 1975; Brown *et al*, 1992) and confirmed for both *H. e. cyrbia* and *H. himera* (F Marec and A Rego, unpublished data). The mean number of AFLP markers per linkage group was 9.6; however, there was high variance and linkage groups could be defined by as many as 20 and as few as two AFLP markers. Nonetheless, six or more

AFLP bands defined 18 of the 21 linkage groups, including the groups that contained both major color pattern loci segregating in this cross. The remaining 25 AFLP loci could not be assigned to any linkage group. This proportion of unlinked loci, approximately 11%, was similar to the proportion reported in other AFLP-based mapping studies (Tan *et al*, 2001; Parsons and Shaw, 2002). There were no obvious scoring difficulties with these loci and their anomalous behavior but may reflect genomic contamination or degraded DNA template. In addition to the AFLP loci, we also placed 11 allozymes, 11 microsatellites and five SCNL within the linkage groups.

We also examined the segregation of AFLP fragments within our male reciprocal sibling backcross (Brood 67) using 12 of the 15 primer combinations used to identify linkage groups. In total, 88 loci (68 AFLPs, seven

isozymes, 10 microsatellites and four SCNL) were informative in both backcrosses. In addition, there were another 30 AFLP loci, two isozymes and one microsatellite that only segregated in Brood 67. Most of these loci segregated in the expected 1:1 ratio, and based on their segregation patterns, we constructed recombinant maps for 20 of the 21 linkage groups identified (Figure 2). One of these linkage groups (group f), however, was defined by three AFLP markers that were completely linked. The final linkage map was composed of 98 AFLP, nine isozymes, 11 microsatellites and four SCNL. The average distance between markers was approximately 23 cM and linkage group size ranged from 18.1 to 431.1 cM. In total, two linkage groups were defined by four codominant markers, seven had two codominant markers and five had one codominant marker.

Figure 2 Genetic map for *H. erato*. Loci within parentheses at the bottom of linkage groups were linked based on segregation in the female-informative cross. Linkage groups were labeled starting with the sex chromosome Z, the remaining labeled with numbers in descending order by size first for linkage groups with at least one isozyme or SCNL, then for linkage groups with at least one microsatellite and finally for those groups only identified with only AFLPs (and hence not easily portable to future studies) by letter. Marker position along linkage group is given in cM. Full name corresponding to map abbreviation of loci in Supplementary Appendix.

The sex chromosomes in Lepidoptera are termed Z and W, with males being homogametic ZZ and females heterogametic ZW. The Z-chromosome contained five AFLPs and three codominant loci; two microsatellites (*Hel-07* and *Hel-06*) and the SCNL *Tpi*. *Hel-06* and *Tpi* behaved as expected of sex-linked loci: females were never heterozygous, always hemizygous. However, *Hel-07* provided evidence of either a sex-chromosome/autosome translocation or a pseudoautosomal region on the Z/W chromosomes. The locus clearly mapped to the Z-linkage groups based on the segregation patterns in the female-informative cross. However, female offspring in the male-informative cross were heterozygous. *Hel-07* mapped to one end of the Z-chromosome and was 27 cM from the nearest marker (Figure 2).

The backcross broods reported here generated phenotypes ranging from *H. himera* pure types to F1 types (Figure 1, compare e and h). Assigning genotypes could be difficult, because the relevant information was provided by the forewing band coloration, which is affected by both *D* and *Cr* (Figure 3). The absence of crossing over in females and the existence of marked linkage groups provided an opportunity to measure the accuracy of the scoring system. Based on segregation patterns in the female-informative cross, we placed the *Cr* and *D* locus on linkage groups 2 and 3, respectively. Subjectively, the *D* locus was easier to score, and this was borne out by a perfect cosegregation of the *D* scores and all other markers on linkage group 3. For the *Cr* locus, there were two inconsistencies between our phenotypic scores and the segregation patterns of the other markers on linkage group 2. These cannot be due to recombination, because of the absence of crossing over in females. Instead, these inconsistencies reflected scoring difficulties due to strong epistasis between Cr^{hi} and D^{cyr} alleles, with red forewing pigmentation largely epistatic to yellow (Figure 3). In the male-informative cross, the *Cr* locus mapped to one end of linkage group 2 and was flanked by two loosely linked AFLP markers. An equivalent number of scoring errors in that cross could inflate the genetic distances on that end of the linkage group. The large distance between the easier-to-score *D* locus and the other markers on linkage group 3 in the male-informative cross was probably a function of recombination and not scoring errors. The recombination map of this linkage group was defined by only two AFLPs and the SCNL *Ci* and the *D* locus was maximally distant from any other cosegregating marker (Figure 2).

Given these data, we estimated the total length of the *H. erato* map to be approximately 2400 cM. However, this estimation should be viewed with caution. Several linkage groups (5, e and f) were largely uncharacterized and contributed very little to the overall size estimation. In addition, the size of at least two groups (3 and a) appears to be inflated, due to poor marker coverage and/or possible AFLP typing errors. Estimates of the physical size of the *H. erato* genome were 1C = 395.5 Mb ($\pm$3.3) for females ($n = 5$) and slightly larger, 1C = 397.3 Mb ($\pm$3.9), for males. This is somewhat smaller than the size of the *Bombyx mori* genome, estimated to be around 495 Mb (Gage, 1974). Our estimates of genome size based on total DNA content and recombination distance for *H. erato* yields a relationship of approximately 165 kb/cM. This rough estimate falls within the range reported in insects and was similar to the silk moth (Yasukochi, 1998) but was less than the 575 kb/cM reported in *Drosophila* (Merriam *et al*, 1991) and greater than the estimate in honey bees (52 kb/cM; Hunt and Page 1995). Of course, this estimate must be viewed with caution since we have probably underestimated recombination distance and the relationship between recombination distance and physical map distance is nonlinear.

Discussion

Wing color pattern variation in *H. erato* presents biologists with a rare opportunity to link natural adaptation, morphological evolution and development. We present the first linkage map of *H. erato* and demonstrate the utility of crossing this species to its sister species, *H. himera,* to map the genes responsible for *H. erato*'s adaptive radiation. *H. erato* × *H. himera* crosses reveal consistently higher levels of polymorphism at marker loci than found between parapatric *H. erato* races traditionally used in genetic studies of this diversification (Brower, 1994; Jiggins *et al*, 1997; Beltrán *et al*, 2002; Flanagan *et al*, 2004). This high level of polymorphism allowed us to identify all 21 expected linkage groups and place two major color pattern loci on the map using only a small set of AFLP primer combinations and codominant loci. In spite of the clear genetic differentiation, crosses between *H. himera* and *H. erato* did not show genomic incompatibilities that might impede further fine scale linkage analysis. All crosses between parapatric *H. himera* × *H. e. cyrbia* showed no obvious hybrid inviability or sterility (McMillan *et al*, 1997), a conclusion that extends to crosses between *H. himera* and allopatric races of *H. erato* collected from very distant locations in South and Central America (unpublished data). In addition, with the exception of a possible autosome/sex chromosomal translocation identified in this study, marker loci showed no obvious segregation distortion that would indicate selective mortality of certain genotypes. A larger AFLP screen of an F2 *H. himera* × *H. e. etylus* cross yielded a similar result. In this case, of the approximately 380 loci segregating in a 1:1 fashion, only 8% showed segregation distortion, which is not significantly different than expected by chance alone (Kapan *et al*, in preparation).

The two loci segregating in our crosses, *Cr* and *D*, serve to underscore the complex action and interaction of genes underlying pattern variation. Both loci were originally described based on phenotypic characteristics

Figure 3 Epistatic interactions between the alleles at the *Cr* and *D* locus affecting pattern in the central region of the forewing in backcrosses to *H. himera*. Individuals are heterozygous ($D^{hi} D^{cy}$) at the *D* locus but were either homozygous ($Cr^{hi} Cr^{hi}$) or heterozygous ($Cr^{hi} Cr^{cy}$) at the *Cr* locus. The bar indicates presumed genotype at the *Cr* locus.

$D^{hi}D^{cy}Cr^{hi}Cr^{hi}$ $D^{hi}D^{cy}Cr^{hi}Cr^{cy}$

segregating in Mendelian fashion in early *Heliconius* crosses (Emsley, 1964; Sheppard *et al*, 1985). The *Cr* locus was initially identified for a small patch of recessive yellow 'cream rectangles' on the ventral outer hindwing margin *H. e. phyllis* from Eastern Brazil. However, subsequent crosses showed that the locus affected recessive yellow and white pattern elements much more broadly across both the dorsal and ventral fore- and hindwing surfaces (Sheppard *et al*, 1985; Mallet, 1989), including, as demonstrated with *H. himera* and *H. e. cyrbia*, the color of the forewing band, the hindwing margins (melanic or iridescent blue scales versus white) and the ventral hindwing line (melanic or iridescent blue scales versus yellow, Figure 1) (Jiggins and McMillan, 1997). The *D* locus, in contrast, affects primarily red and orange pattern elements and was initially described from the segregation of the red/orange patch, the so-called 'Dennis', at the base of the dorsal forewing (Sheppard *et al*, 1985). The phenotypic effects, hypothesized to be allelic by Sheppard *et al* (1985), include 'radiate' or 'rayed' patterns on the dorsal and ventral hindwing of many Amazonian races (shown to cosegregate by Mallet, 1989), red band in the forewing and red bars on the base of the dorsal hindwing of *H. himera* (Jiggins and McMillan, 1997). Alleles at this locus, or a group of very tightly linked loci, affect red/orange pattern more broadly than initially described leading the entire locus, or complex of very tightly linked loci, to be called D^{Ry} (Sheppard *et al*, 1985; Mallet, 1989). In this case, the '*y*' refers to the loci's affect on forewing band color (red/yellow) in earlier crosses (Sheppard *et al*, 1985; Mallet, 1989).

When alleles of the *Cr* and *D* loci interact across the same area of the wing, they show strong epistatic effects that correlate with scale color (and perhaps microstructure); red scales are largely epistatic to melanic scales, which in turn mask white or yellow scales (Gilbert *et al*, 1988; Nijhout, 1991). This epistatic pecking order can make unambiguous scoring of genotypes difficult in some crosses. For example, although extensive crosses of *H. e. cyrbia* and *H. himera* clearly demonstrate the presence of two major loci (Jiggins and McMillan, 1997), the *Cr* locus is harder to score in backcrosses to *H. himera* due to the dominant melanic effects of the Cr^{hi} allele along the margin of the hindwing and the interaction between *Cr* and *D* alleles in the forewing. For instance, in the backcross to *H. himera* reported here, red in the forewing shows the unambiguous presence of a *D* allele from *H. e. cyrbia*. However, in backcrosses to *H. himera* individuals heterozygous at the *D* locus show a nearly continuous distribution of red in the forewing band (Jiggins and McMillan, 1997; see Figure 3). We inferred the *Cr* genotypes in these *D* heterozygotes based on the relative amount of expression of red in the forewing using a '50%' rule: individuals showing <50% red were scored as $Cr^{hi} Cr^{hi}$, whereas those with >50% red were scored as $Cr^{hi} Cr^{cyr}$. In most cases, genotypic scoring was unambiguous; however, some individuals fell near the 50% cutoff blurring the line between *Cr* homozygotes and heterozygotes (Figure 3).

The process of untangling the complex epistatic effects among color pattern loci will allow much better understanding of the developmental underpinnings of pattern formation. AFLP mapping and linkage analysis can speed this process in several important ways. First, the high level of AFLP polymorphism in *H. erato* × *H. himera* crosses will allow us to follow the segregation of thousand of markers with moderate effort. Thus, in crosses where a color pattern genotype can be scored without ambiguity (for instance, in $D^{him} D^{him}$ homozygotes or segregation of the Cr^{cyr} allele in backcrosses to *H. e. cyrbia*), we can identify markers tightly linked to particular color pattern genes (Kapan *et al*, in preparation; Jiggins *et al*, in press) to predict the color pattern genotype of individuals in crosses where dominance or epistasis makes phenotypic expression more variable. Furthermore, combining the large number of segregating markers with quantitative, rather than qualitative, measures of pattern may allow us to resolve the interactions among major loci by controlling for major effect loci as well as facilitating the identification of poorly described modifier loci of smaller effect important in refining the mimetic color patterns.

In spite of AFLP's utility for linkage analysis of color pattern variation segregating within a single cross, placing the major genes underlying the *H. erato* color pattern radiation on a single 'reference' linkage map requires the development of additional marker systems. Only a subset of over 20 described color loci segregate in a given cross between distinct color pattern races of *H. erato* (Sheppard *et al*, 1985); therefore, a full description of the genetic architecture of this radiation requires integrating many independent linkage maps. AFLP's high level of polymorphism makes it difficult to assign homology and integrate AFLP maps from different genetic crosses. This problem is particularly acute for mapping traits in outbred populations, as is the case here. For *H. erato*, increasing the number of shared AFLP markers between crosses by crossing different racial variants to the same *H. himera* line can help. However, even using our sibling backcross design, only 68 of the approximately 170 AFLP fragments examined were informative in both broods (Supplementary Appendix).

AFLPs are more useful when coupled with at least two codominant markers per linkage group. Codominant 'anchor loci' identify two or more identical reference points on homologous linkage groups (Shi *et al*, 1995), thus facilitating the ordering of AFLP maps from different genetic crosses. We achieved this level of codominant marker coverage for only eight of the 21 linkage groups (Figure 1). Of the three different codominant marker systems employed in this study, SCNL have three advantages over microsatellites and allozymes. First, SCNL were polymorphic: of six SCNL screened, only *Ddc* did not possess sufficient variation for mapping, a surprising fact since it contains two introns. In contrast, introns of several other SCNL (*Mpi*, *Tpi*, *Ci*) were exceptionally variable and even highly conserved developmental loci such as *Wg* and *Dpp* had distinctive alleles that segregated in our crosses.

Second, SCNL can be developed from an expanding list of potential developmental candidate genes. Butterfly wing pattern formation may be a two-step process where genes involved in cell signaling and signal transduction determine a prepattern in early development and this positional information is later interpreted by individual scale cells that produce pattern-specific color pigments (Nijhout, 1991). In the last decade, developmental biologists discovered that genes typically involved with axis specification and appendage development have

416

been redeployed and appear to be important in wing pattern development (reviewed in McMillan *et al*, 2002). For example, the gene network including distal-less, engrailed, cubitus interruptus, notch, hedgehog and spalt appears to regulate eyespots in several species (Carroll *et al*, 1994; Brakefield *et al*, 1996; Keys *et al*, 1999; Brunetti *et al*, 2001; Reed and Serfas, 2004). These genes from conserved developmental pathways, and those from the biochemical networks involved in pigment synthesis, provide many potential candidate loci for wing color patterns in *Heliconius* (see McMillan *et al*, 2002). In this study, we have excluded three developmental genes – key ligands (*Wg* and *Dpp*) or transcription factors (*Ci*) involved in three major *Drosophila* signaling pathways (Carroll *et al*, 2001) – as candidates for the *Cr* or *D* loci.

Third, unlike microsatellites, which proved surprisingly difficult to develop in *Heliconius* because of the presence of null alleles, SCNL developed from specific genes consistently amplified and showed sufficient polymorphism for mapping in other *Heliconius* species (Jiggins *et al*, in press). This is critical for our ultimate objective of understanding the genetic basis of convergent evolution between the distantly related comimics, *H. erato* and *H. melpomene*. Linkage maps of the two parallel color pattern radiations will allow us to assess the level of conserved synteny and determine if convergent evolution in the two species complexes involves homologous genetic regions or, alternatively, whether similar phenotypes have evolved from genetically distinct regions of the genome.

References

Bates HW (1862). Contributions to an insect fauna of the Amazon valley. Lepidoptera: Heliconidae. *Trans Linn Soc Lond* **23**: 495–566.

Beltrán M, Jiggins CD, Bull V, Linares M, Mallet J, McMillan WO *et al* (2002). Phylogenetic discordance at the species boundary: comparative gene genealogies among *Heliconius* butterflies. *Mol Biol Evol* **19**: 2176–2190.

Bennett MD, Leitch IJ, Price HJ, Johnston JS (2003). Comparisons with *Caenorhabditis* ($\sim$100 Mb) and *Drosophila* ($\sim$175 Mb) using flow cytometry show genome size in *Arabidopsis* to be $\sim$157 Mb and thus $\sim$25% larger than the *Arabidopsis* initiative estimate of $\sim$125 Mb. *Ann Botany* **91**: 547–557.

Brakefield PM, Gates J, Keys D, Kesbeke F, Wijngaarden PJ, Monteiro A *et al* (1996). Development, plasticity, and evolution of butterfly eyespot patterns. *Nature* **384**: 236–242.

Brower AVZ (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc Natl Acad Sci 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.

Brower AVZ, Egan MG (1997). Cladistic analysis of *Heliconius* butterflies and relatives (Nymphalidae: Heliconiiti): a revised phylogenetic position for *Eueides* based on sequences from mtDNA and a nuclear gene. *Proc Roy Soc Lond Ser B* **264**: 969–977.

Brown KS (1981). The biology of *Heliconius* and related genera. *Annu Rev Entomol* **26**: 427–456.

Brown KS, Emmel TC, Eliazar PJ, Soumalainen E (1992). Evolutionary patterns in chromosome numbers of neotropical Lepidoptera. I. Chromosomes of the Heliconiini (family Nymphalidae: subfamily Nymphalinae). *Hereditas* **117**: 109–125.

Brunetti CR, Selegue JE, Monteiro A, French V, Brakefield PM, Carroll SB (2001). The generation and diversification of butterfly eyespot color patterns. *Curr Biol* **11**: 1578–1585.

Carroll SB, Gates J, Keys DN, Paddock SW, Panganiban GEF, Seleque JE *et al* (1994). Pattern formation and eyespot determination in butterfly wings. *Science* **265**: 109–114.

Carroll SB, Grenier JK, Weatherbee SD (2001). *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design.* Blackwell Scientific: Malden.

Emsley MG (1964). The geographical distribution of the color-pattern components of *Heliconius erato* and *Heliconius melpomene* with genetical evidence for the systematic relationship between the two species. *Zool NY* **49**: 245–286.

Flanagan N, Tobler A, Kapan D, Davison A, Pybus O, Planas S *et al* (2004). Historical demography of Müllerian mimicry in the Neo-tropical *Heliconius* butterflies. *Proc Natl Acad Sci USA* **101**: 9704–9709.

Flanagan NS, Blum MJ, Davison A, Alamo M, Albarrán R, Faulhaber K *et al* (2002). Characterization of microsatellite loci in neotropical *Heliconius* butterflies. *Mol Ecol Notes* **2**: 398–401.

Gage LP (1974). The genome: analysis by DNA reassociation kinetics. *Chromosoma* **45**: 27–42.

Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoosabady E *et al* (1983). Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* **220**: 1049–1051.

Gilbert LE (2003). Adaptive novelty through introgression in *Heliconius* wing patterns: evidence for shared genetic 'toolbox' from synthetic hybrid zones and a theory of diversification. In: Boggs CL, Watt WB, Ehrlich PR (eds) *Ecology and Evolution Taking Flight: Butterflies as Model Systems*. University of Chicago Press: Chicago.

Gilbert LE, Forrest HS, Schultz TD, Harvey DJ (1988). Correlations of ultrastructural and pigmentation suggest how genes control development of wing scales on *Heliconius* butterflies. *J Res Lepidoptera* **26**: 141–160.

Harrison RG, Rand DM, Wheeler WC (1987). Mitochondrial DNA variation in field crickets across a narrow hybrid zone. *Mol Biol Evol* **4**: 144–158.

Hawthorne DJ (2001). AFLP-based genetic linkage map of the Colorado potato beetle *Leptinotarsa decemlineata*: sex chromosomes and a pyrethroid-resistance candidate gene. *Genetics* **158**: 695–700.

Hunt GJ, Page Jr RE (1995). Linkage map of the honey bee, *Apis mellifera*, based on RAPD markers. *Genetics* **139**: 1371–1382.

Jiggins CD, McMillan WO (1997). The genetic basis of an adaptive radiation: warning colour in two *Heliconius* species. *Proc Roy Soc Lond Ser B* **246**: 1167–1175.

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

Jiggins CD, Naisbit RE, Coe RL, Mallet J (2001). Reproductive isolation caused by colour pattern mimicry. *Nature* **411**: 302–305.

Joron M, Wynne IR, Lamas G, Mallet J (1999 2001). Variable selection and the coexistence of multiple mimetic forms in *Heliconius numata*. *Evol Ecol* **13**: 721–754.

Kapan DD (2001). Three-butterfly system provides a field test of Müllerian mimicry. *Nature* **409**: 338–340.

Keys DN, Lewis DL, Selegue JE, Pearson BJ, Goodrich LV, Johnson RL *et al* (1999). Recruitment of a hedgehog regulatory circuit in butterfly eyespot evolution. *Science* **283**: 532–534.

Lincoln SE, Daly MJ, Lander ES (1993). *MAPMAKER/EXP*. Whitehead Institute for Biomedical Research: Boston.

Mallet J (1989). The genetics of warning colour in Peruvian hybrid zones of *Heliconius erato* and *H. melpomene*. *Proc Roy Soc Lond Ser B* **236**: 163–185.

Mallet J, Barton NH, Lamus G, Santisteban J, Muedas M, Eeley H (1990). Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in *Heliconius* hybrid zones. *Genetics* **124**: 921–936.

Mallet J, Gilbert LE (1995). Why are there so many mimicry rings? Correlations between habitat, behaviour, and mimicry in *Heliconius* butterflies. *Biol J Linn Soc* **55**: 159–180.

Mallet J, Joron M (1999). Evolution of diversity in warning color and mimicry: polymorphisms, shifting balance, and speciation. *Annu Rev Ecol Evol* **30**: 201–233.

McMillan WO, Jiggins CD, Mallet J (1997). What initiates speciation in passion-vine butterflies? *Proc Natl Acad Sci USA* **94**: 8628–8633.

McMillan WO, Monteiro A, Kapan DD (2002). Development and evolution on the wing. *Trends Ecol Evol* **17**: 125–133.

Merriam J, Ashuburner M, Hartl DL, Kafatos FC (1991). Toward cloning and mapping the genome of *Drosophila*. *Science* **254**: 221–225.

Müller F (1879). *Ituna* and *Thyridia*; a remarkable case of mimicry in butterflies. *Trans Entomol Soc Lond* **1879**: 20–29.

Naisbit RE, Jiggins CD, Mallet J (2003). Mimicry: developmental genes that contribute to speciation. *Evol Dev* **5**: 269–280.

Nijhout HF (1991). *The Development and Evolution of Butterfly Wing Patterns*. Smithsonian Institution Press: Washington.

Parsons YM, Shaw KL (2002). Mapping unexplored genomes: a genetic linkage map of the Hawaiian cricket *Laupala*. *Genetics* **162**: 1275–1282.

Punnett RC (1915). *Mimicry in Butterflies*. Cambridge University Press: Cambridge.

Reed RD, Serfas MS (2004). Butterfly wing pattern evolution is associated with changes in a notch/distal-less temporal pattern formation process. *Curr Biol* **14**: 1159–1166.

Sandal N, Krusell L, Radutoiu S, Olbryt M, Pedrosa A, Stracke S *et al* (2002). A genetic linkage map of the model legume *Lotus japonicus* and strategies for fast mapping of new loci. *Genetics* **161**: 1673–1683.

Sheppard PM, Turner JRG, Brown KS, Benson WW, Singer MC (1985). Genetics and the evolution of Müllerian mimicry in *Heliconius* butterflies. *Philos Trans Roy Soc Lond Ser B* **308**: 433–613.

Shi J, Heckel DG, Goldsmith MR (1995). A genetic linkage map for the domesticated silkworm, *Bombyx mori*, based on restriction fragment length polymorphisms. *Genet Res Camb* **66**: 109–126.

Suomalainen E, Cook LM, Turner JRG (1973). Achiasmatic oogenesis in the heliconiine butterflies. *Hereditas* **74**: 302–304.

Tan Y-D, Wan C, Zhu Y, Lu C, Xiang Z, Deng H-W *et al* (2001). An amplified fragment length polymorphism map of the silkworm. *Genetics* **157**: 1277–1284.

Turner JRG (1976). Adaptive radiation and convergence in subdivisions of the butterfly genus *Heliconius* (Lepidoptera: Nymphalidae). *Biol J Linn Soc* **58**: 297–308.

Turner JRG (1983). Mimetic butterflies and punctuated equilibria: some old light on a new paradigm. *Biol J Linn Soc* **20**: 277–300.

Turner JRG, Smiley J (1975). Absence of crossing-over in female butterflies (*Heliconius*). *Heredity* **34**: 265–269.

Vos P, Hogers R, Bleeker M, Reijans M, Lee Tvd, Hornes M (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* **23**: 4407–4414.

Yasukochi Y (1998). A dense genetic map of the silkworm, *Bombyx mori*, covering all chromosomes based on 1018 molecular markers. *Genetics* **150**: 1513–1525.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)