Convergent Evolution in the Genetic Basis of Müllerian Mimicry in Heliconius Butterflies

Simon W. Baxter,*† Riccardo Papa,† Nicola Chamberlain,† Sean J. Humphray,§ Mathieu Joron,** Clay Morrison,†† Richard H. ffrench-Constant,†
W. Owen McMillan†† and Chris D. Jiggins*

*Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, United Kingdom; †University of California, Irvine, California 92697; ‡University of Exeter, Cornwall TR10 9EZ, United Kingdom; §The Wellcome Trust Sanger Institute, Cambridge CB10 ISA, United Kingdom; **Institute for Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT; ††Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695

Manuscript received October 5, 2007
Accepted for publication August 31, 2008

ABSTRACT

The neotropical butterflies Heliconius melpomene and H. erato are Müllerian mimics that display the same warningly colored wing patterns in local populations, yet pattern diversity between geographic regions. Linkage mapping has previously shown convergent red wing phenotypes in these species are controlled by loci on homologous chromosomes. Here, AFLP bulk segregant analysis using H. melpomene crosses identified genetic markers tightly linked to two red wing-patterning loci. These markers were used to screen a H. melpomene BAC library and a tile path was assembled spanning one locus completely and part of the second. Concurrently, a similar strategy was used to identify a BAC clone tightly linked to the locus controlling the mimetic red wing phenotypes in H. erato. A methionine rich storage protein (MRSP) gene was identified within this BAC clone, and comparative genetic mapping shows red wing color loci are in homologous regions of the genome of H. erato and H. melpomene. Subtle differences in these convergent phenotypes imply they evolved independently using somewhat different developmental routes, but are nonetheless regulated by the same switch locus. Genetic mapping of MRSP in a third related species, the “tiger” patterned H. numata, has no association with wing patterning and shows no evidence for genomic translocation of wing-patterning loci.

O NLY recently has it become feasible to study the molecular basis of phenotypic adaptation in higher organisms, which has generated considerable interest in identifying genes controlling adaptive traits (Nachman et al. 2003; Mundy 2005; Protas et al. 2006). Commonly, positional cloning of qualitative traits involves performing crosses for trait segregation, creating linkage maps using the progeny, identification of the genomic region responsible for variation, and then genetic mapping of candidate genes to test for association by linkage (Heckel et al. 1998). This approach has been successful in associating candidate genes to the causal locus in several cases, including insecticide resistance in Heliothis virescens (Gahan et al. 2001), pelvic reduction in threespine sticklebacks (Shapiro et al. 2004), wing color patterning in butterflies (Kronforst et al. 2006b), albinism in cave population Astyanax fish (Protas et al. 2006), and adaptive variation in coat color of mice (Steiner et al. 2007). Identifying the adaptive trait becomes more difficult where no suitable candidate genes exist or all known candidates can be ruled out by linkage mapping. In this case, sequencing of genomic bacterial artificial chromosome (BAC) clones spanning the genome region is required. Recently, (Colosimo et al. 2005) positionally mapped and sequenced the locus controlling armor plating in threespine sticklebacks from a genomic BAC library and identified a strong candidate for determination of lateral plate phenotypes. Identifying genetic markers tightly linked to the plate morph locus was achieved by screening stickleback crosses with amplified fragment length polymorphism (AFLP) markers, enabling markers linked to the region of interest to be identified without any prior knowledge of the genome sequence. Markers developed in this way can also facilitate comparative studies to determine whether the same regions of the genome are involved in adaptation in different species.

A comparative mapping approach has already been applied to the three butterfly species, Heliconius melpomene, H. erato, and H. numata (Joron et al. 2006). The first two species are comimics that share the same warningly
colored wing patterns in local populations and different phenotypes between geographic regions of Central and South America, while \( H. \) numata is a species with polymorphic populations in which a single genetic locus controls multiple morphs. A positional cloning approach in \( H. \) melpomene led to identification of markers linked to the \( Yb \) patterning locus, which controls a yellow hind-wing bar. These were subsequently mapped in the other two species and found to be tightly linked to the \( H. \) erato locus \( Cx \), which has similar phenotypic effects to \( Yb \), and the \( H. \) numata locus \( P \) which controls whole-wing phenotypic polymorphism. This surprising result indicates that a single genetic locus is controlling both convergent and divergent phenotypes in different lineages (Joron et al. 2006).

Wing color and pattern variation among the ~25 mimetic races of \( H. \) melpomene and \( H. \) erato is controlled by several unlinked loci with major phenotypic effect (Sheppard et al. 1985). Extensive crossing experiments have shown that in \( H. \) melpomene, genetic factors on at least 4 of the 21 chromosomes are involved in wing color patterning (Sheppard et al. 1985; Joron et al. 2006). \( H. \) erato has a similar architecture, with the involvement of at least 4 different chromosomes. Recent work has shown preliminary evidence for further homology between the two species in addition to that of \( Yb \) and \( Cx \). The locus \( Ac \), which controls a dumbbell-shaped element in \( H. \) melpomene, is on the homologous chromosome to \( Sl \), which controls forewing band shape in \( H. \) erato (Kronforst et al. 2006a). Loci controlling red wing pattern elements genetically map to the same chromosomal region as \( Cubitus \) interruptus (\( C \)) in \( H. \) erato (D locus), \( H. \) melpomene (B and D loci), and sister species \( H. \) cydnus (G locus) (Joron et al. 2006; Kronforst et al. 2006a).

While \( H. \) melpomene B/D (hereafter \( Hmb/Hmd \)) and \( H. \) erato D (\( HdD \)) control broadly similar, and adaptively convergent mimetic pattern elements, there are nonetheless some fundamental differences both in the action of the switch loci in the two species and in the phenotypic expression of their pattern elements. First, \( HdD \) controls both red/orange and yellow color in the medial forewing band. In contrast the \( H. \) melpomene yellow forewing is controlled by a completely unlinked locus, \( N \), and is possibly influenced by an additional locus \( M \) (Mallet 1989). Second, the hind-wing ray elements in \( H. \) erato are determined by vein position, with all pattern elements lying in intervein compartiments. The \( H. \) melpomene rays involve an orange band that bisects vein boundaries and is apparently homologous in position to the yellow band controlled by the locus \( Yb \) on LG15. Thus, these genes control adaptively convergent patterns but their resultant phenotypes and mode of action are sufficiently distinct, leading to speculation whether the same genes are homologous (Njihout 1991) or not (Mallet 1989). We here carry out fine-scale mapping to determine whether red patterning elements of \( Hmb/Hmd \) and \( HdD \) are indeed homologous.

Genetic analysis using AFLPs provides a method for rapidly generating large numbers of markers that can be used to analyze poorly characterized genomes (Heckel et al. 1999; Brugmans et al. 2002; Kazachkova et al. 2004; Jiggins et al. 2005). Previous genetic work on \( H. \) melpomene wing patterning relied on the discovery of an AFLP marker tightly linked to the \( Yb \) locus on linkage group 15 (Jiggins et al. 2005). Here we apply an AFLP method for targeted identification of molecular markers linked to phenotypic traits of interest in Lepidoptera. Markers flanking \( H. \) melpomene \( Hmb \) and \( Hmd \) were developed and, unexpectedly, show these loci to be tightly linked. BAC library screening using AFLP markers linked to red patterning loci identified candidate clones for sequencing that will ultimately enable positional cloning of these phenotypes. Finally, comparative mapping shows a gene linked to \( HdD \) in \( H. \) erato is in a homologous chromosomal region to \( Hmb/Hmd \) of \( H. \) melpomene, however is unlinked to wing patterning in the related species \( H. \) numata.

**MATERIALS AND METHODS**

**Wing pattern phenotypes:** Here, we use the terminology \( Hmb \) for the locus controlling presence or absence (\( Hmb \)) of the \( H. \) melpomene red medial forewing band and \( Hmd \) for the locus controlling the presence or absence (\( Hmd \)) of the red forewing base, hind-wing bar, and hind-wing rayed phenotype. A single locus, \( D \), controls this convergent pattern in \( H. \) erato, referred to here as \( Hed \). Note that \( Hed \) was originally termed \( D^e \) by Sheppard et al. (1985) and subsequently simplified to \( D \) by Jiggins and McMillan (1997) (Figure 1).

**Insect crosses:** All \( H. \) melpomene crosses were performed within sealed, outdoor cages at the Smithsonian Institute, Panama. Single-pair matings were set up between \( H. \) m. melpomene collected from French Guiana (4°54’ N, 52°21’ W), homozygous for the \( Hmb \) phenotype and \( H. \) m. malleti collected from Ecuador (0°10’ S, 77°41’ W) homozygous for the \( Hmd \) phenotype. Crosses were then performed between unrelated \( F_1 \) heterozygotes, to generate three large \( F_2 \) broods that were reared to adulthood (brood 44 with 164 progeny; brood 48 with 109 progeny; and brood 52 with 99 progeny). The combined 372 \( F_2 \) progeny segregated for the presence or absence of \( Hmb \) and \( Hmd \) phenotypes in Mendelian ratios of 1:2:1 (100 \( bd/B ; 197 B/D 75 Bd/d \)), not differing significantly from expected \( \chi^2 = 4.66, P > 0.1 \) [2 degrees of freedom (d.f.)]. A similar crossing strategy was performed on \( H. \) cydnus at the University of Puerto Rico. For this cross, outbred stocks of \( H. \) erato \( etylus \), collected from the Zamora River (3°35’ S, 78°50’ W) and \( H. \) himera, collected from Vikabamba (4°16’ S, 79°13’ W), both in Loja Province in southern Ecuador were crossed to produce \( F_1 \) hybrid lines. Unrelated \( F_1 \) individuals were crossed to produce a family containing 92 progeny that segregated for the \( H. \) e. \( etylus \) \( Hed^e \) phenotype and \( H. \) himera \( Hed^r \) phenotypes in the expected 1:2:1 Mendelian ratio (23 \( Hed^e Hed^e, 44 Hed^r Hed^e, 25 Hed^r Hed^r \)) not differing significantly from expected \( \chi^2 = 0.26, P > 0.8 \) (2 d.f.) (Kapan et al. 2006). For all crosses, wings were removed from butterflies and stored in paper envelopes, and bodies preserved in a salt-saturated solution of 10% DMSO and 0.2 M EDTA at −20°C.

\( H. \) numata crosses have been described previously (Joron et al. 2006). Here, two polymorphic \( F_2 \) broods were analyzed that segregated for three alleles of the supergene \( P \) from the
**Genetic markers and linkage map assembly:** As crossing over between sister chromatids does not occur during oogenesis in Lepidoptera, all markers on the same chromosome inherited from the mother are in complete linkage. Three *H. melpomene* F₂ broods (44, 48, and 52) segregating for *HmbB* and *HmdD* were genotyped for two linkage group 18 (LG18) molecular markers, nuclear gene *Cubitus interruptus* and microsatellite *HmH4*. Alleles inherited from the F₁ mother were used to assign chromosome prints to progeny. This provided a molecular diagnostic to determine whether *HmbB/D* heterozygous progeny inherited a *Bd* or *bd* chromosome from the F₁ mother. To calculate recombinational distances between *HmbB* and *HmdD* and known markers on LG18, chromosomes inherited from the F₁ parents were analyzed (crossing over does occur during spermatogenesis in Lepidoptera). Direct estimation of recombination between *HmbB* and *HmdD* was not possible as these dominant cosegregating phenotypes are inherited in repulsion (Figure 2).

**Linkage map assembly:** The *H. erato* linkage map was adapted from KAPAN et al. (2006). MapMaker Macintosh 2.0 was used to assemble maps for *H. melpomene* and *H. numata*. *H. numata* brood 472 contained 80 F₂ individuals and was genotyped for *Rpb30* and *MRSP* (there was no variation at *Ci*), and brood 502 contained 88 individuals and was genotyped for *Rpb30* and *Ci* (there was no variation at *MRSP*).

**H. melpomene BAC library preparation, screening, and end sequencing:** A *H. melpomene* BAC library was constructed using vector pECBAC1 and contains 18,816 clones with an average insert size of 123 kb, providing a 7.9× genome coverage (Ambíx Press) (Joron et al. 2006). A bacterial colony physical map of the genome was then constructed using *HindIII* restriction enzyme fingerprinting (Humphray et al. 2001; Schein et al. 2004). BAC library plates were directly cultured into 170 μl 2× T/Y in 384-well plates. After overnight growth the BAC DNA was extracted by alkaline lysis on a Packard MiniTrack and digested with *HindIII* in the 384-well plates. Following electrophoresis on 121 lane 1% agarose gels, the data were collected using a Typhoon 8600 fluorimager and raw images were entered into the fingerprint database using the software IMAGE (http://www.sanger.ac.uk/Software/Image). The output of normalized band values, sizes, and gel traces were analyzed in FPC (Soderlund et al. 2000), which bins and orders clones on the basis of shared bands, taking marker data into account when available. Fingerprinting was successful for 17,494 clones, of which 13,010 assembled into 2086 contigs (the remainder were singletons). Each contig averaged 6.2 clones and 791 contigs contained 5 or more clones. AWebFPC database displaying all contigs is available at http://www.sanger.ac.uk/Projects/H_melpomene/WebFPC/helmin/small.shtml.

The *H. melpomene* BAC library clones were printed onto nylon membranes and hybridization performed using eight probes linked to the *HmbB* and *HmdD* loci including three AFLP markers, *Hm_eCCGmCTA* (primers CG-CTA_F2 X CG-CTA_R2), *Hm_eCGmATG213*, and *Hm_eACTmGTA195*; four BAC end sequences, *bHm40A21_sp6*, *bHm36K2_sp6*, *bHm28L23_sp6*, and *bHm40C14_sp6*; and gene *MRSP* (primers *MRSP_F1* X *MRSP_R1*). PCR products were generated under the following conditions: 50 ng genomic DNA of *H. melpomene* 1× reaction buffer, 2.5 mM MgCl₂, 0.2 mM dNTP, 50 pmol each primer, Taq polymerase (Bio-Line). PCR products were then labeled with

---

**Figure 1.—** Heliconius melpomene and *H. erato* comimics. D phenotype: Dorsal and ventral wings from *H. melpomene* ecuadorensis and *H. erato* eurynome, displaying the D phenotype. Specimens were collected near Zamora, Ecuador. The yellow forewing patch is controlled by *HeD* in *H. erato* and by a second locus (called *N*) in *H. melpomene*. B phenotype: *H. melpomene* melpomene and *H. erato* hydara were collected near Cayenne, French Guiana (4°91’ N, 52°36’ W). All individuals are males.

**Nucleotide extraction and processing:** Genomic DNA was extracted from small sections of abdomen or thorax with the DNeasy tissue kit (QIAGEN) and resuspended in TE buffer (100:1, pH 8.0). AFLP templates were prepared for 35 *H. melpomene* brood 44 progeny, F₁ parents, and grandparents using half reactions of AFLP analysis system II (Invitrogen Life Technologies) according to manufacturer’s instructions, with *EcoR*I and *Msi*I restriction enzymes. AFLP PCR reactions were performed using fluorescently labeled Eco primers with two or three selective bases and paired with an *Mse*I primer containing three selective bases. AFLP PCR reactions were diluted 1/60 in water, and 1 μl added to 9 μl of HiDi formamide solution containing 0.01 μl of GeneScal-500 LIZ size standard (Applied Biosystems). Products were separated on a 3730 DNA Analyzer (Applied Biosystems) using GeneMarker. Data files were converted to GeneScan format and AFLP peaks sized using GeneScan Analysis software (version 3.1.2). Analyzed files were then imported into Genographer (version 2.0), which digitally reconstructed chromatogram peaks to form images similar to electrophoretic gels.

**Sequencing AFLP markers:** AFLP PCR reactions containing bands linked to the color pattern loci were separated using polyacrylamide gel electrophoresis (acrylamide 6%, 1X TBE, 7 μTBE) at 1750 V, 80 W for 1.5–3 hr. Acrylamide gels were stained with Silver Sequence (Promega) to visualize DNA bands, which were then excised and resuspended in water. All products were reamplified and sequenced using BigDye terminator v3.1 (Applied Biosystems) using an ABI3730. Ten AFLP markers were sequenced (FI109935–FI109944) and used to design genotyping assays for genetic mapping in *H. melpomene* brood 44, and where possible, broods 48 and 52. Genotyping was performed using (i) size variation in PCR products when separated using agarose gel electrophoresis or (ii) sequencing F₁ parents from crosses to identify restriction enzymes that recognize single nucleotide polymorphisms segregating in F₂ progeny. CodonCode Aligner V1.5.2 software was used to analyze sequence chromatograms.
**RESULTS**

**Molecular markers for red wing pattern loci:** Genetic crosses have previously shown *Hmb* and *Hmd* are dominant loci on the same linkage group (Turner 1972), designated linkage group 18 (LG18) in *H. melpomene* (Jiggins et al. 2005; Joron et al. 2006). To generate molecular markers linked to these loci, AFLP bulk segregant analysis was performed on the F₁ parents of brood 44 and two F₂ progeny bulks. The first bulk contained 12 F₂ individuals, homozygous *Hmb* (no red forewing band, *HmbD/-d*), and the second bulk contained 12 F₂ individuals homozygous for *Hmd* (no red hind-wing rays, bar, and red forewing base, *Hmbd/-d*). As the F₂ progeny were siblings, the two bulks were expected to share many AFLP bands, but should be strongly differentiated in genome areas tightly linked to *Hmb* or *Hmd*. In Lepidoptera, chromosomal crossing over occurs during spermatogenesis and not oogenesis (Mayda 1939; Turner and Sheppard 1975; Traut 1977). Therefore, AFLP analysis focused on bands inherited from the F₁ father for recombinational linkage mapping. Two hundred fifty-six AFLP primer combinations produced ~5100 bands in both F₂ bulks, inherited from the F₁ father. In addition, 19 bands were present in only one bulk and represented candidate molecular markers for red pattern elements. AFLP primer combinations for these 19 bands were reanalyzed in F₁ parents and 35 individual progeny (including the 24 individuals used to assemble the bulks). Ten AFLP markers were confirmed linked to *Hmb* or *Hmd* and 9 of the bands were excluded as false positives due to banding patterns not consistent with linkage group 18.

Five AFLP markers (*Hm_eACAmCAT133*, *Hm_eCTmGGC148*, *Hm_eACTmGTA195*, *Hm_eCGmGA139*, and *Hm_eCCmACA177*) were present in 16/16 of the *Hmb* individuals tested and absent from most of the *Hmd* individuals, suggesting these markers were genetically linked to the *Hmb* phenotype. The remaining five markers (*Hm_eCGmCTA386*, *Hm_eCGmATG213*, *Hm_eACAmCAG224*, *Hm_eCTmCTC156*, and *Hm_eCamAGC355*)...
### TABLE 1
Primer sequences for linkage mapping and *H. melpomene* BAC library probes

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Annealing temp</th>
<th>Brood 44</th>
<th>Brood 48</th>
<th>Brood 52</th>
<th>BAC screen result</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFLP markers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACC_CAG_F</td>
<td>TTCTCAACACAAAACTACTCTTCA</td>
<td>58</td>
<td>LPA</td>
<td>—</td>
<td>—</td>
<td>NA</td>
<td>FI109944</td>
</tr>
<tr>
<td>ACC_CAG_R</td>
<td>ATTCAACCCACCCCAATACAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC-CTA_F1</td>
<td>TOCCGTCTAACACGAGC</td>
<td>68</td>
<td>LPA</td>
<td>LPA</td>
<td>LPA</td>
<td>Nonspecific</td>
<td>FI109959</td>
</tr>
<tr>
<td>CC-CTA_R1</td>
<td>GCTCTGCACTGGCTCTGGTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC-CTA_F2</td>
<td>TATCTCAAGAAGTCTCTGTA</td>
<td>60</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG-ATG_F1</td>
<td>AATTCGCGCAACATGAGA</td>
<td>58</td>
<td>LPA</td>
<td>—</td>
<td>—</td>
<td>Contig 446</td>
<td>FI109941</td>
</tr>
<tr>
<td>CG-ATG_R1</td>
<td>GTGAGGTCCAGCTGGTTGACAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACA-CAT_F1</td>
<td>TGCCGCGATAGAATAA</td>
<td>55</td>
<td>LPA</td>
<td>—</td>
<td>—</td>
<td>NA</td>
<td>FI109957</td>
</tr>
<tr>
<td>ACA-CAT_R1</td>
<td>GACATCTGGCTCAATTCGCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACT-GTA_F1</td>
<td>TATAGCTAATTTTACTTGGC</td>
<td>55</td>
<td>LPA</td>
<td>LPA</td>
<td>seq</td>
<td>Contig 1729</td>
<td>FI109956</td>
</tr>
<tr>
<td>ACT-GTA_R1</td>
<td>CGTCTAGTCAGGCAATGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT-CTC_F1</td>
<td>GATGCATGAGCTGGTGA</td>
<td>55</td>
<td>LPA</td>
<td>—</td>
<td>—</td>
<td>NA</td>
<td>FI109955</td>
</tr>
<tr>
<td>CT-CTC_R1</td>
<td>ACTCGATGAAATTTAAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRSP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRSP_F1</td>
<td>GAAACCCTGCTCTTCTTCAA</td>
<td>59</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>BanII, Sty1</td>
<td>Contig 16</td>
</tr>
<tr>
<td>MRSP_R1</td>
<td>AATGCTGAGGGAGACACGAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRSP_F2</td>
<td>GGTGCTTGAAGAGAAGACTGG</td>
<td>60</td>
<td>MseI</td>
<td>—</td>
<td>—</td>
<td>NA</td>
<td>EU711403</td>
</tr>
<tr>
<td>MRSP_R2</td>
<td>TGCCGCACTGCAAAGAGGAGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAC end primers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bHM5K10_t7F</td>
<td>AAAGACCTGGGAGAATTATT</td>
<td>58</td>
<td>LPA</td>
<td>LPA</td>
<td>LPA</td>
<td>NA</td>
<td>FI107292</td>
</tr>
<tr>
<td>bHM5K10_t7R</td>
<td>GGTGCTAGTGTAGTCAAGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bHM40C14_sp6F</td>
<td>TOCCATCAAGGGCTAGTG</td>
<td>58</td>
<td>NoII</td>
<td>SpI</td>
<td>NoII</td>
<td>Contig 1729</td>
<td>FI107289</td>
</tr>
<tr>
<td>bHM40C14_sp6R</td>
<td>TGGTAAAGTGGGCGACGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bHM7G5_sp6F</td>
<td>AGTCTGTGTTGGTGCTCAT</td>
<td>58</td>
<td>LPF</td>
<td>—</td>
<td>—</td>
<td>LPF</td>
<td>FI107293</td>
</tr>
<tr>
<td>bHM7G5_sp6R</td>
<td>GACCAGGGAGTGGACGAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bHM35j13_F</td>
<td>TTGGTGTTCAACAGGGTGA</td>
<td>58</td>
<td>—</td>
<td>LPA</td>
<td>LPA</td>
<td>NA</td>
<td>FI107286</td>
</tr>
<tr>
<td>bHM35j13_R</td>
<td>AAGAAGGCTGGCTGGTCTAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bHM7G5_1F</td>
<td>AGCGGCGAGAAAACCTAGT</td>
<td>58</td>
<td>LPF</td>
<td>LPF</td>
<td>LPF</td>
<td>NA</td>
<td>FI107294</td>
</tr>
<tr>
<td>bHM7G5_1R</td>
<td>ATGCCTGATCGTCTTTTCAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bHM40A21_sp6F</td>
<td>GAGGOCTTATGGGAGAAGAC</td>
<td>58</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Contig 936</td>
<td>Contig 1080</td>
</tr>
<tr>
<td>bHM40A21_sp6R</td>
<td>AACGAGGCTTGGGAGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bHM47N1_t7F</td>
<td>TGACGGGCTAATGCTAAA</td>
<td>58</td>
<td>LPF</td>
<td>LPF</td>
<td>LPF</td>
<td>Contig 1305</td>
<td>FI107290</td>
</tr>
<tr>
<td>bHM47N1_t7R</td>
<td>AAAGGAGACTGGCATGCAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bHM5D10_sp6F</td>
<td>GTGACCCTACTGGGGAGAG</td>
<td>58</td>
<td>LPF</td>
<td>LPF</td>
<td>LPF</td>
<td>NA</td>
<td>FI107291</td>
</tr>
<tr>
<td>bHM5D10_sp6R</td>
<td>AGCCAACAGTCAGAAATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bHM36K2_sp6F</td>
<td>GTGGTTATTTTCATGCTGGG</td>
<td>58</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>Contig 1080</td>
<td>Contig 1352</td>
</tr>
<tr>
<td>bHM36K2_sp6R</td>
<td>GCACCTCTCTTGAGGGTGTGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
were absent in all HmbD/b- individuals and present in most of the HmbD/-d individuals, suggesting these DNA fragments were linked to the Hmd phenotype. Recombination events between AFLP markers and Hmd were observed in two individuals that were not used in the bulked segregant analysis (Table 2).

**Hmb and Hmd are tightly linked loci:** The 10 Hmb- or Hmd-linked AFLP bands were sequenced, following separation with polyacrylamide gel electrophoresis, and ranged in size from 133 bp to 386 bp. Oligonucleotides were designed for each AFLP marker (Table 1) for recombinational linkage mapping in brood 44 to determine distances to Hmb and Hmd. A linkage map of LG18 was constructed using 164 progeny from brood 44 using microsatellite Hm14, recently identified LG18 nuclear genes Bm44, Ribosomal protein gene S30 (RpS30) (Pringle et al. 2007) as well as Ci and 6 of the 10 AFLP sequence-tagged sites (Hm_eCGmATG213, Hm_eACcmCAG224, Hm_eACAmCAT133, Hm_eACTmGA195, Hm_eCTmCTC156, and Hm_eCCmCTA386). The remaining four AFLP Hmb/D-linked markers either contained no sequence variation, which is required for linkage mapping, or amplified >2 alleles (Figure 3). All AFLPs that were successfully mapped were within 5.9 cM of Hmb and Hmd and closely linked to nuclear gene Bm44. There were no recombinants between AFLP marker Hm_eCGmATG213 and the Hmb/D locus.

Although Hmd and Hmb were inherited in repulsion, molecular markers show both phenotypes genetically map to the same chromosomal region in brood 44. To test the robustness of the association, two additional broods were genotyped for two Hmb/D-linked AFLP markers. From the combined 372 progeny of broods 44, 48, and 52, 199 inherited the B or b from the F1 male and 173 inherited D or d from the F1 male. The first marker, Hm_eCCmCTA386, showed 1/173 recombinants to the Hmd/d locus and 1/199 recombinants to Hmb/b. A marker flanking the two wing-patterning loci, Hm_eACTmGA195, showed 3/173 recombinants to Hmb/d and 4/199 recombinants to Hmb/b. The AFLP marker closest to Hmb/D in brood 44 (Hm_eCGmATG213) was not polymorphic in broods 48 or 52.

**BAC tile path surrounding Hmb and Hmd loci:** Radiolabeled probes designed from three Hmb/D-linked AFLP markers (Hm_eACTmGA195, Hm_eCCmCTA386, and Hm_eCGmATG213) were hybridized to H. melpomene BAC library filters and positive clones referenced against the BAC fingerprint database. Hm_eACTmGA195 hybridized to a single clone, in contig 1729 (Figure 4). Genetic markers designed from the sequenced BAC end of contig 1729 were mapped in 372 progeny from H. melpomene broods. Three recombination events were detected between contig ends, enabling orientation and an estimation of recombination frequency in physical distance (3/372 recombinants in ~140 kb).

Probe Hm_eCCmCTA386 hybridized to 19 of 32 BAC clones within the same fingerprint contig 196. Again,
recombination was used to orientate the clone, with one end containing no recombinants to the \textit{HmB} or \textit{HmD} loci. Probe Hm\textsubscript{e}CGmATG213 hybridized to all 6 clones from contig 446 and 5 of the 10 clones from contig 936. The BAC ends from contigs identified using AFLPs were used to rescreen the library and close the gap between contigs 446 and 196. A BAC tile path spanned the color-pattern locus \textit{HmB}, flanked by recombinant individuals (Figure 4). The \textit{HmD} locus is located in a larger region, overlapping with \textit{HmB}.

Convergent patterns map to homologous loci: The loci controlling red wing pattern elements in \textit{H. melpomene} and \textit{H. erato} are known to be on homologous chromosomes (Joron \textit{et al.} 2006; Kapan \textit{et al.} 2006; Kronforst \textit{et al.} 2006a). A previously identified \textit{H. erato} AFLP marker (He\textsubscript{CCAC}491), positioned 1.5 cM from the \textit{HeD} locus (Kapan \textit{et al.} 2006), was sequenced and used to screen a \textit{H. erato} BAC library. The probe hybridized to clone BBAM-25K4 (GenBank AC216670); however, recombinational linkage mapping ruled out the possibility that the \textit{HeD} locus was contained within this clone. To identify nuclear genes close to \textit{HeD}, BBAM-25K4 was sequenced and contained a gene encoding a Methionine Rich Storage Protein (MRSP) (EU711403). Comparative mapping of MRSP in \textit{H. melpomene} broods identified 2/173 recombinants to the \textit{HmD} locus and 2/199 recombinants to the \textit{HmB} locus and is positioned farther from contig 196 (Figure 4). \textit{H. melpomene} BAC library screening was performed with MRSP and identified contig 16, containing 63 clones that did not overlap with any previously detected contigs. For \textit{H. erato}, we observed one recombinant between the \textit{HeD} locus and MRSP from 88 F\textsubscript{2} individuals (in the \textit{H. erato etylus} × \textit{H. himera} cross).

Color diversity in a third Heliconius species, \textit{H. numata}, is controlled by a single locus called \textit{P} on linkage group 15 of this species. Theoretical models converge on the view that supergenes cannot form by bringing together previously unlinked loci through a graduation tightening of linkage or through translocation of genetic elements from different regions of the genome; instead, supergenes are believed to form via the

\begin{table}
\centering
\caption{AFLP markers identified using bulked segregant analysis that are linked to \textit{HmB/b} and \textit{HmD/d} wing-patterning loci}
\begin{tabular}{llll}
\hline
\textit{H. melpomene} AFLP marker & Accession no. & Size (bp) & Band present in F\textsubscript{2} progeny \\
\hline
Hm\textsubscript{e}CCmCTA386 & FI109939 & 386 & 0/16 & 19/19 \\
Hm\textsubscript{e}CCmATG213 & FI109941 & 213 & 0/16 & 19/19 \\
Hm\textsubscript{e}ACCmCAG224 & FI109944 & 224 & 0/16 & 19/19 \\
Hm\textsubscript{e}CAmAGC355 & FI109943 & 355 & 0/16 & 15/16 \\
Hm\textsubscript{e}CTmCTC156 & FI109935 & 156 & 0/16 & 17/19 \\
Hm\textsubscript{e}ACAmCAT133 & FI109937 & 133 & 16/16 & 1/18 \\
Hm\textsubscript{e}CTmGCCG148 & FI109938 & 148 & 16/16 & 1/19 \\
Hm\textsubscript{e}ACTmGTA195 & FI109936 & 195 & 16/16 & 1/18 \\
Hm\textsubscript{e}CGmGAA139 & FI109940 & 139 & 16/16 & 1/18 \\
Hm\textsubscript{e}CCmACA177 & FI109942 & 177 & 16/16 & 2/19 \\
\hline
\end{tabular}
\end{table}
recruitment of linked genetic elements (Charlesworth and Charlesworth 1975; Turner 1977; Le Therry d’Ennequin et al. 1999). To empirically evaluate those theoretical results, we wanted to determine whether there was evidence for translocation from the \(HmB/\)\(HmD\) region to the \(P\) locus in \(H.\) \(numata\). Three genes, \(MRSP\), \(RpS30\), and \(Ci\) were genetically mapped in crosses segregating for \(P\) phenotypes. Both \(MRSP\) and \(RpS30\) mapped to LG18 in \(H.\) \(numata\), separated by 2.9 cM, while \(RpS30\) and \(Ci\) were distantly linked at 59 cM, demonstrating the same gene order and similar distances between markers as in \(H.\) \(melpomene\) and \(H.\) \(erato\). This provides no evidence for large-scale rearrangements between LG18 and LG15 in \(H.\) \(numata\).

This provides no evidence for large-scale rearrangements between LG18 and LG15 in \(H.\) \(numata\). The distances between \(Ci\) and \(MRSP\) in these three species are \(H.\) \(erato\), 74.4 cM; \(H.\) \(melpomene\), 68.7 cM; and \(H.\) \(numata\), 61.9 cM (Figure 5).

**DISCUSSION**

The mimetic color patterns of Heliconius butterflies are striking phenotypic adaptations controlled by few loci of major effect. Through adapting traditional AFLP methods, we have identified molecular markers tightly linked to the \(HmB\) and \(HmD\) loci, controlling the presence or absence of red and orange pattern elements in \(H.\) \(melpomene\). Here we used the process of DNA-based bulked segregant analysis, which has been widely applied in plant studies to identify markers linked to a phenotypic locus (Michelmore et al. 1991). The principle involves dividing siblings of a cross into different pools corresponding to a segregating phenotype. The pools (or bulks) are largely heterozygous, except around the locus controlling the phenotype. When analyzed with PCR-based methods such as AFLPs, markers that are differentially amplified between the bulks should genetically map near the locus of interest. Although we have yet to identify the actual gene(s) controlling red wing pattern phenotypes, this work nonetheless demonstrates that positional cloning for phenotypic traits is a feasible goal for nonmodel organisms outside plant research fields.

Our comparative genetic mapping data in \(H.\) \(melpomene\) and \(H.\) \(erato\) have shown that homologous chromosomal regions are responsible for similar (but not
identical) red wing phenotypes in these two species. This provides further evidence of homologous genomic regions shared between *H. melpomene* and *H. erato*, which control similar pattern elements (Joron et al. 2006). There are however notable interspecific differences between the *D* locus of these species. First, the *H. erato* *HeD* locus controls not only the switch for presence or absence of red elements, but the switch between yellow and red in the forewing, while the yellow forewing band in *H. melpomene* is controlled by a separate unlinked locus, *N*. Second, the rayed phenotype of *H. melpomene* consists of a nail-shaped ray and a hind-wing bar that cuts across several wing veins and appears to be vein independent. In contrast the hind-wing rayed pattern of *H. erato* appears to be entirely vein dependent with all pattern elements lying in intervein regions. Differences such as these between the species have led some to speculate that the loci controlling pattern in the two species are unlikely to be homologous (Mallet 1989).

The results presented here and in a previous study (Joron et al. 2006) imply that common genetic switch mechanisms are responsible for convergent patterns in *H. melpomene* and *H. erato*, and these evolved independently in the two lineages. Despite findings consistent with this hypothesis, it may not prove true for all races of *H. melpomene*. Sheppard et al. (1985) reported recombination between *HmB* and *HmD* loci, on the basis of a series of *H. melpomene* progeny tests involving parental individuals collected from Suriname or Belém do Para, Brazil, (*HmD* homozygotes) and Apure in southwestern Venezuela or Trinidad (*HmB* homozygotes) (Turner and Crane 1962). Parental genotypes were calculated on the basis of phenotypes of progeny, and in 3 of 11 cases, parents must have undergone recombination events to produce the phenotypes observed in the progeny. From this, Sheppard et al. (1985) calculated a crossover rate of 3/11 = 27% (SE ± 13%). In contrast, here we show *HmB* and *HmD* to be tightly linked, more similar to that of the *HeD* locus in *H. erato* that controls the parallel red wing phenotype. Individuals used to genetically map the *HmB/D* loci in this study were collected from Ecuador (*HmD*), and French Guiana (*HmB*), and therefore it is possible that the genetic linkage relationships of these loci vary across the geographic range of *H. melpomene*, perhaps due to chromosomal rearrangement, which would explain the apparent discrepancy in our results. Alternatively, it has been suggested that two genes producing similar phenotypes to *HmB* may be present on LG18 (J. R. G. Turner, personal communication). There is evidence to support two independent loci controlling a different phenotype in *H. melpomene*, the yellow forewing band (Mallet 1989; Mallet et al. 1990). Generally, individuals with genotypes 11010 do not. In some cases however, field-caught individuals expected to be 11010 present a yellow band that may be due to a second locus mm (Mallet et al. 1990). Determining whether multiple loci control similar phenotypes in *H. melpomene* requires further crossing experiments and linkage mapping using markers linked to known color pattern loci.

There are several cases of convergent phenotypes that arose independently, either within the same gene or through different adaptive mechanisms. Protas et al. (2006) for example reported that albinism in two isolated populations of the Mexican cave fish Astyanax arose via independent mutations in the same pigmentation gene, Oca2. Hoekstra et al. (2006) investigated variable coat color between subpopulations of beach mice along the Florida Gulf Coast and identified a single nucleotide polymorphism (SNP) in the melanocortin-1 receptor pigmentation gene (*Mc1r*) with a major phenotypic effect. When compared to beach mice with similar coat variability along the Atlantic coast, the causal-derived SNP was absent, suggesting phenotypic adaptation evolved independently in this case. The
major-effect loci that control color pattern across the Heliconius species investigated so far demonstrate a common “toolbox” of genes controlling pattern polymorphism. Between the conomic species, homologous loci have broadly similar effects on color pattern, suggesting functional similarities. However, the common switch loci may affect downstream pathways in different ways, which would explain differences in mode of action of the major loci in different lineages.

The _H. numata_ wing-patterning supergene is controlled at a single locus “P” on LG15, which enables this species to mimic specific Melinaea butterflies. According to theory, mimicry supergenes are unlikely to be formed by bringing together unlinked patterning genes to a single chromosomal region, _e.g._, via a translocation; instead, they are more likely to originate from a major gene mutation, followed by a series of linked modifiers to perfect the mimicry (Charlesworth and Charlesworth 1975) or by the selection of gene clusters through a sieve effect from redundant gene networks (Turner 1977; Le Thierry d’Ennequin et al. 1999). Our data represent the first explicit empirical evaluation of those predictions: in accordance with theory, we have found no genomic evidence, at this level of resolution, for the translocation of _H. numata’s_ HmB/HmD orthologous region from LG18 to LG15, to form part of the P supergene. Although the translocation hypothesis may be entirely ruled out only following definitive identification and cloning of the HmB/D genetic elements, the formation of the P supergene in _H. numata_ does not appear to have involved a rearrangement of chromosomal regions unlinked in related species. Nevertheless, it is interesting to note the differences in recombination rates between markers Rps30 and MRSP flanking HmB and HmD loci. In _H. melpomene_, 7.7 cM separates these two genes, yet in _H. numata_, the region is only 2.9 cM. The centimorgan differences may simply be an artifact of brood size, reduced recombination, or physical genome size differences. This lends support to the hypothesis that the genes involved in switching red patterns in various Heliconius species are positionally conserved in _H. numata_, but play little or no role in the gene networks underlying pattern polymorphism in this species. Whatever the explanation, these observations clearly imply that there is not a simple mapping of phenotype to genotype. Instead, a conserved developmental switch has been adopted to generate both divergent and convergent patterns in different ecological and genetic contexts.

In _H. melpomene_, we have demonstrated the feasibility of cloning color pattern genes using AFLP bulk segregant analysis, coupled with a BAC screening walk. Recombinants either side of the HmB locus were determined, within a physical distance of ~700 kb, and a recombinant bordering one side of the HmD locus was identified, producing a larger region to search for candidate genes. It is striking that multiple overlapping fingerprinted contigs, each containing many clones, were identified in the proximity of HmD/HmB. That these clones clustered into different contigs when fingerprint data were analyzed implies that this region either must contain highly divergent alleles sampled in our BAC library or possibly contains duplicated genomic segments. The _H. melpomene_ BAC library was constructed using six pupae of unknown wing pattern from a polymorphic laboratory population segregating for different alleles of HmB and HmD. Consequently, it is not unexpected that allelic variability has been sampled, which might reflect different color patterns. Sequencing multiple, overlapping clones will therefore be a useful tool for identifying allelic variants for future studies of field populations.

In our crosses, the components forming HmD and HmD (proximal red patch at the base of the forewing, the red hind-wing bar plus red hind-wing rays) are always observed together, in both species studied (Figure 1). Nonetheless, these phenotypes are not always associated. Rare recombinants (or mutants) are known in both species, in which the forewing patch and hind-wing bar are separated from the hind-wing rays (Mallet 1989). Furthermore, the races _H. melpomene_ meriana and _H. erato_ amalfreda display the red forewing proximal patch and hind-wing bar, but not the ray phenotype. These races and occasional recombinants suggest the HmD and HmD loci are controlling these red pattern elements by separate, tightly linked sites that may or may not be part of the same gene.

The data presented here represent an important step toward cloning the HmB/D gene(s) in _H. melpomene_ and HmD in _H. erato_. For HmB, we have now assembled a continuous BAC tile path encompassing the locus and are sequencing the BAC clones to identify candidate genes (supplemental Figure 1). The recombination fraction in _H. melpomene_ is estimated to be 180 kb/1 cM, and just outside the HmB/D tile path (contig 1729, Figure 4) we have found 3/372 recombinants within ~140 kb, which does not differ from expected (χ² = 0.25, P > 0.6). Of course, factors such as inversions, duplications, or recombinational “deserts” could potentially complicate cloning of the locus, but we are nonetheless optimistic that this study represents a significant step toward identification of the HmB/D gene(s). This will offer novel insights into the molecular evolution underlying these dramatic color patterns and the role of single loci in regulating developmental pathways in a recent evolutionary radiation. The study also demonstrates the feasibility of this method for positional cloning that can be applied to other non-model organisms with major-effect genes of economic or evolutionary interest.

We are grateful for the helpful comments and discussion provided by two anonymous reviewers. Technical support was received from Sequencing and Genotyping Facility at Universidad de Puerto Rico and The Wellcome Trust Sanger Institute, United Kingdom. The
Ministerio del Ambiente in Ecuador granted permission to collect butterflies and Autoridad Nacional del Medio Ambiente granted permission to work in Panama. This work was supported by the Biotecnology and Biological Sciences Research Council, the National Science Foundation (DEB-0715096 and IBN-0544705), the Royal Society, and the Smithsonian Tropical Research Institute.

LITERATURE CITED


Communicating editor: D. Rand