Evolution of a mimicry supergene from a multilocus architecture

Robert T. Jones\textsuperscript{1,2,3,}, Patricio A. Salazar\textsuperscript{2}, Richard H. ffrench-Constant\textsuperscript{1}, Chris D. Jiggins\textsuperscript{2} and Mathieu Joron\textsuperscript{3}

\textsuperscript{1}School of Biosciences, University of Exeter, Penryn TR10 9EZ, UK
\textsuperscript{2}Department of Zoology, University of Cambridge, Cambridge CB2 3EF, UK
\textsuperscript{3}CNRS UMR7205, Département Systématique et Evolution, Muséum National d’Histoire Naturelle, 75005 Paris, France

The origin and evolution of supergenes have long fascinated evolutionary biologists. In the polymorphic butterfly \textit{Heliconius numata}, a supergene controls the switch between multiple different forms, and results in near-perfect mimicry of model species. Here, we use a morphometric analysis to quantify the variation in wing pattern observed in two broods of \textit{H. numata} with different alleles at the supergene locus, ‘\textit{P}’.

Further, we genotype the broods to associate the variation we capture with genetic differences. This allows us to begin mapping the quantitative trait loci that have minor effects on wing pattern. In addition to finding loci on novel chromosomes, our data, to our knowledge, suggest for the first time that ancestral colour-pattern loci, known to have major effects in closely related species, may contribute to the wing patterns displayed by \textit{H. numata}, despite the large transfer of effects to the supergene.

Keywords: morphometric analysis; \textit{Heliconius numata}; colour pattern; supergene

1. INTRODUCTION

A variety of complex adaptive traits are controlled by supergenes. These are tightly linked genetic loci that are inherited as a single unit, and are associated with the switch between different polymorphic features. Supergenes were first described for the heterostylus system of plants such as buckwheat and \textit{Primula} species, in which they control the development of ‘pin’ or ‘thrum’ flower morphs. In these cases, different elements of the supergene control the different and developmentally separate features, such as stamen level or stigma height [1,2].

In addition to the heterostylus system in primroses, supergenes have been studied in a number of different polymorphisms including the Rhesus blood groups of humans and shell morphology in snails [3,4], but perhaps the best-known examples come from butterflies. Supergenes in the palatable Batesian mimics \textit{Papilio memnon} and \textit{Papilio dardanus} contain multiple loci that control different elements of the mimetic pattern on both the forewing and hindwing, as well as body colour and the presence of hindwing tails in \textit{P. memnon} [5,6]. Because recombination could result in the expression of wing patterns that are neither cryptic in pattern nor good mimics, selection acts to maintain the linkage of co-adapted haplotypes [7], and ensure that crosses between individuals of different genotypes will not produce maladaptive recombinant progeny [8].

Although earlier evolutionary geneticists predicted that tightly linked genetic loci would control polymorphisms, the mechanisms that led to the unusual genetic architecture of supergenes provided a puzzle. It was initially suggested that different genes which contribute to a single phenotype might be brought together from around the genome by translocations, and selection for reduced recombination would tighten their linkage [3,9,10]. However, theoretical work has shown that this route is unlikely, and the alternative ‘sieve’ hypothesis has gained favour. For mimicry polymorphisms, this hypothesis suggests that different genetic mechanisms may be capable of producing similar pattern elements, but only mutations in regions with multiple loci that are already closely linked, and which are together capable of producing major phenotypic effects, will be recruited [11,12]. This leads to a condition in which one region of a particular chromosome accumulates all the variants that affect the mimetic pattern, and these loci can become integrated into the supergene through selection for tighter linkage [7]. Alongside evolution of the supergene itself, the rest of the genome is expected to accumulate epistatic ‘modifier’ loci that improve the adaptive phenotype of one or more of the ‘supergene’ alleles [13,14]. In this study, we focus on a supergene that controls morph determination in the butterfly \textit{Heliconius numata}, and map the genetic loci that contribute to wing pattern variation.

\textit{Heliconius numata} is an unusual case of a polymorphic Müllerian mimic. It is found in South America and joins a distinct coexisting mimicry rings with various unpalatable butterflies and day-flying moths. Their wing patterns contain yellow, brown, orange and black colour elements (figure 1 and electronic supplementary material, figure S1) that are controlled by variation at a supergene locus named \textit{PushmiPullyu}, abbreviated to \textit{P} [15]. \textit{P} is orthologous in its genomic location with major mimicry loci (\textit{Yb}, \textit{Sb}, \textit{N}) that primarily control the presence of yellow pattern elements in three other \textit{Heliconius} species: \textit{Heliconius melpomene}, \textit{Heliconius cydno} and \textit{Heliconius erato} [16]. However, in other \textit{Heliconius} species loci on several
different linkage groups (LG) are implicated in the control of major variation in wing pattern [17]. For instance, B and D on LG18 in *H. melpomene* control the presence of red pattern elements, while K on LG01 controls the yellow versus white switch of the forewing band colour in *H. cydno* [17]. Each of the known colour-pattern loci in the *Heliconius* toolkit controls different components of the wing pattern, and makes different contributions across species and races [18,19]. This contrasts with the single supergene locus on LG15 in *H. numata* that controls multiple element whole-wing polymorphism, corresponding to mimicry of *Melitaea* and *Mechanitis* species [15]. The B–D region and other loci identified in other species have not previously been found to be involved in colour-pattern determination in *H. numata* [20]. The P supergene, therefore, appears to have taken control of the entire *Heliconius* colour-pattern ‘toolkit’ in response to selection against colour-pattern recombinants in polymorphic populations; switch loci in other parts of the genome have presumably played less and less part in pattern variation.

The expression of P alleles has a precise dominance hierarchy. In the Tarapoto region of Peru, P has at least seven distinct alleles; the ‘bicoloratus’ phenotype is at the top of the dominance series and ‘silvana’ is the most recessive [15,21]. This hierarchy ensures that heterozygous offspring do not display intermediate and non-mimetic phenotypes, and results in multiple mimetic forms [22]. Within each mimetic morph there is some variation in pattern displayed by different individuals, but the genetics of this quantitative variation has not been investigated. Indeed, little is known about the influence of minor-effect alleles in the rest of the genome because *Heliconius* phenotypes have largely been described in a qualitative manner. We were therefore interested in quantifying variation in *H. numata* wing patterns, and were particularly interested in the possibility that variation could be controlled by loci in the *Heliconius* toolkit that regulate major aspects of colour pattern in other species. To address this question and to quantify variation associated with the mimicry supergene, we measured colour-pattern variation across the entire wing surface. We then adopted a quantitative trait locus (QTL) analysis to determine which chromosomes contain the minor effect loci. Additional genetic markers on selected chromosomes were then used to further map several of the QTL identified more finely.

### 2. MATERIAL AND METHODS

#### (a) Heliconius numata broods

Two broods were used: brood 502 (B502) and brood 472 (B472), both already described elsewhere [16]. Briefly, B502 derives from a heterozygous *elegans* male parent (425, *p*ele*psil*) and a heterozygous *aurora* female parent (502, *p*aurl*p*aurl), sired, respectively, by *silvana* male × *elegans* female, and *tarapotensis* male × *aurora* female. Offspring of the brood therefore had *p*aurl*p*aurl, *p*ele*psil*, *p*ele*psil* and *p*aurl*p*aurl* genotypes. B472 derives from a heterozygous *elegans* male parent (421, *p*ele*psil*) and a homozygous *silvana* female parent (472, *psil* *psil*). Brood progeny were of the genotype *p*ele*psil* or *psil* *psil*. B502 had 88 offspring and B472 had 80 offspring (electronic supplementary material, table S1).

#### (b) Chromosome prints

There is no recombination during female meiosis in Lepidoptera [23]. This leads to inheritance of whole, unrecombined maternal chromosomes that can be detected by following ‘female informative’ genetic markers for each LG. The patterns of inheritance of these polymorphic female-informative markers among the offspring for each LG are known as chromosome ‘prints’ [20]. We used microsatellite and gene markers to generate full chromosome prints for our broods (electronic supplementary material, methods). The same technique was also used for ‘male-informative’ markers, which allow for mapping within an LG.

#### (c) Image analysis

Image analysis of *H. numata* wings was performed with MATLAB (v. 7.0, The MathWorks), using the Active Appearance Model (AAM) Toolbox (A. Hana, University of East Anglia). This program analyses the shape of the wing as well as the pattern of pixel colour and intensity across its surface, and accurately captures variation between images [24]. A principal component (PC) analysis is used by the AAMToolbox to find patterns in the data, and break the uncorrelated variables into different ‘components’ (electronic supplementary material, methods). For the foregoing analysis, we used a template with 13 ‘primary’ points, placed at clear landmarks around the
wing such as at the termini of veins, and 10 'secondary' points positioned between these points. For the hindwings, we used eight primary and 10 secondary points (figure 1d) around the ventral side of the wing.

(d) Statistical analysis
A general linear model (GLM) analysis of variance was used to associate PC data with genotype data using GenStat (v. 15). Eigenvalues were produced in the AAMToolbox upon generation of the statistics model, and only PCs with an eigenvalue greater than 1 were used in the analyses. We repeated the analysis for both broods of *H. numata*, using whole-brood, within-sex and within-morph analyses, and used a Bonferroni adjustment when interpreting p-values to correct for multiple tests (electronic supplementary material, methods).

3. RESULTS

(a) Chromosome prints
We used molecular markers on the 21 LGs of *H. numata* to generate chromosome 'prints' for each individual in two broods, B502 and B472 (electronic supplementary material, table S2). This allowed direct identification of LGs that carry QTL without the need for a dense array of markers. Having identified an LG, male-informative markers could be used for mapping at a finer scale.

(b) Analysis of brood 502
(i) PC analysis of forewings
A photograph of the forewing of each individual in B502 was analysed by the AAMToolbox, which reduced all the pattern traits into uncorrelated variables using PC analysis. For the whole brood, 17 PCs had an eigenvalue greater than 1 and together accounted for 99 per cent of variation (electronic supplementary material, figure S5). This allowed direct identification of whole-brood variation, we found a significant association between PC scores and genotype data using MInitAB (v. 15). Eigenvalues were produced in the AAMToolbox distinguishing between the different alleles of markers. Having identified an LG, male-informative markers could be used for mapping at a finer scale.

(ii) Statistical analysis of forewings
The most significant associations between PC scores and genotype data are summarized in table 2. When analysing whole-brood variation, we found a significant association of PC1 with LG15, which is consistent with the AAMToolbox distinguishing between the different alleles of markers. Having identified an LG, male-informative markers could be used for mapping at a finer scale.

Figure 2. (a) Distribution of pattern variation in PCs 1 and 2. *p*~**sil**~*p*~**sil**~ individuals (diamond) were given low scores by AAMToolbox in PC1, separating them in morphospace (left arrow) from *p*~**aur**~*p*~**sil**~ (cross), *p*~**ele**~*p*~**sil**~ (triangle) and *p*~**aur**~*p*~**ele**~ (circle) individuals. PC2 separated *p*~**aur**~*p*~**sil**~ wings (vertical arrow). (b,c) Pattern variation identified in PC1: male *p*~**aur**~*p*~**sil**~ has no hindmarginal black bar and no cubital spot. Arrows show extension of subapical spots and migration of yellow median band towards *p*~**aur**~*p*~**ele**~ phenotype. (c) Male *p*~**aur**~*p*~**ele**~ is dark orange with extensive melanization in basal and discal regions, defining the hindmarginal bar (h) and cubital spot (c) (s ~ subapical spot, d ~ dagger). (d) Size of effects of LGs contributing to variation in B502 whole-brood analysis: forewings (light grey) and hindwings (dark grey). Asterisk indicates LG with known colour-pattern locus.

P. PC2 also showed a significant association with LG15, while the third component associated with LG21, the sex-determining chromosome. The contribution to overall wing morphology of the remaining PCs was small, but significant associations were found between PC4, PC5 and PC6 with LG10 (electronic supplementary material, table S3).

Although there was some structure in the distribution of our GLM residuals (electronic supplementary material, figure S5), we note that this relates to interaction between the maternally and paternally inherited P alleles and is variation in wing pattern that cannot be described...
Table 1. Variation captured in PCs 1-10 in B502 whole-brood analyses. Arrows and shading are used to highlight those regions that vary along the PC axis. LGs with significant associations are indicated.

<table>
<thead>
<tr>
<th>PC</th>
<th>wing</th>
<th>forewing</th>
<th>hindwing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(LG15)</td>
<td>proximal-distal migration of yellow pigment. Discal and post-discal melanic patches emerge</td>
<td>white apical spots lost. Melanic scales cover costal and R1 cells. Yellow elements darken to orange</td>
</tr>
<tr>
<td>2</td>
<td>(LG15)</td>
<td>melanization less pronounced in subapex and discal regions. Subapical spot background colour darkens</td>
<td>melanin band across wing separates from margin. Apical and marginal spots become less bright</td>
</tr>
<tr>
<td>3</td>
<td>(LG21, 10 × 14)</td>
<td>melanization in subapex, dagger and hindmarginal bar becomes fainter. Post-discal patches extend proximally</td>
<td>marginal black patch fades and extends towards costa. Costal patch darkens. Yellow patches stretch apically</td>
</tr>
<tr>
<td>4</td>
<td>(LG10, 15 × 19)</td>
<td>subapical spots extend distally and merge with apical spots. Discal band and dagger become lighter</td>
<td>melanin patch across centre of hindwing loses definition. White marginal spots of apex reduce in size</td>
</tr>
<tr>
<td>5</td>
<td>(LG10)</td>
<td>vein separating cells A1 and Cu1 becomes pronounced. Discal patches condense and darken</td>
<td>melanin patch along tornus becomes smaller and fainter. Pale spots in R2 and M1 become less defined</td>
</tr>
<tr>
<td>6</td>
<td>(LG10)</td>
<td>subapical spots darken. Melanic scales that separate them lighten. Apical spots fade</td>
<td>orange colour of background becomes lighter. White marginal spots reduce in length</td>
</tr>
<tr>
<td>7</td>
<td>(LG06 × 11)</td>
<td>apical spots extend proximally. Discal patches fade and cubital spot reduces in size</td>
<td>melanization in the apex becomes paler. White marginal spots of dorsum extend towards costal cell</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>discal and post-basal patches become more defined. Subapical and apical spots separate</td>
<td>costal melanization becomes paler. Dorsal melanization extends towards anterior</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>subapical spots extend costally and begin to merge together. Discal melanin patches reduce in size</td>
<td>background orange colour becomes paler. Central melanin band becomes fainter and less defined</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>cubital spot and dagger reduce in size. Intensity of melanization reduces in discal and post-basal regions</td>
<td>white marginal spots of dorsum extend costally. Melanization around spots in R1 and R2 becomes faint</td>
</tr>
</tbody>
</table>
through female-informative chromosome prints. To test for a contribution of LG interactions in the captured variation, we incorporated LG-by-LG interactions into our GLM. Four interactions were found to have significant associations with PCs, and involved LG10 and LG15, as well as LGs that do not contain known Heliconius colour-pattern genes and by themselves did not have significant contributions (electronic supplementary material, table S4).

In our within-male analysis, PC1 showed a significant association with LG15, PC4 with LG10 and PC6 with LG18 (electronic supplementary material, table S5). Our within-female analysis found no correlations that fell within our alpha level of 0.05 after Bonferroni correction (electronic supplementary material, table S6).

Because GLM cannot be applied to datasets with fewer than two predictor variables, we excluded LG21 from the within-sex analyses and LG15 from the within-morph analyses.

Within $P^{au}P^{sil}$ individuals, our analysis found that PC1, capturing a proximodistal shift in pattern elements, significantly associated with LG21 (table 2). PC3 captured variation in melanin patch distribution in discal and post-discal areas of the wing, and associated with LG01 as well as LG19, which was identified as interacting with LG15 in the whole-brood interaction study (electronic supplementary material, tables S4 and S7). For the remaining morphs, there were too few individuals to perform a GLM, and we were unable to find associations between PC scores and genotype.

(iii) PC analysis of hindwings

The ventral side of H. numata hindwings varies considerably: $P^{sil}$ individuals have a largely orange hindwing with a black margin, and white patches in the inter-vein regions (figure 1). Morphometric analysis of B502 separated in PC1 the $P^{sil}$ hindwings from those of aurora and elegans individuals, which have a large black patch extending from basal to post-discal cells (figure 3). PC2 clearly separated a cluster of male $P^{ele}P^{sil}$ wings, which have less melanin pigmentation in the submarginal area of R1-M2 (figure 1e and table 1).

(iv) Statistical analysis of hindwings

Statistical analysis of hindwing data showed similar results to the forewings. The most significant correlations in the whole-brood analysis were between the major PCs and LGs 15 and 21 (electronic supplementary material, table S8). We also found interactions between LG15 and both LG06 and LG21 in our pairwise GLMs (electronic supplementary material, table S4). A G-test for the whole-brood analyses summarized in table 2 found a non-random distribution of PCs across the LGs ($G = 32.63, p = 0.037$). Furthermore, the distributions of PCs and of known Heliconius colour-pattern loci on LGs were found to be significantly correlated using contingency $\chi^2$ and Fisher's exact tests (electronic supplementary material, table S9).

Finally, for B502, we performed within-sex and within-morph GLM analyses. Within both sexes of the brood we found that the first PC significantly associated with LG15

Table 2. Summary of linkage groups showing significant associations with PCs. Only LGs with $^*p \leq 0.050$ following Bonferroni correction are shown. $^{**}p \leq 0.010$. Dash indicates LG removed from analysis. Known Heliconius colour-pattern loci are marked (QTL is locus identified in [26]).

<table>
<thead>
<tr>
<th>linkage group</th>
<th>brood 502 whole brood</th>
<th>forewing</th>
<th>hindwing</th>
<th>$p^{au}p^{sil}$</th>
<th>LG21 (sex)</th>
<th>PC3**</th>
<th>PC2**</th>
<th>PC1**</th>
</tr>
</thead>
<tbody>
<tr>
<td>LG01 (K)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LG10 (Ac)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LG13 (QTL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LG15 (Yb/Sb/N)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PC3**</td>
<td>PC2**</td>
<td>PC1**</td>
</tr>
<tr>
<td>LG18 (B/D)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LG19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. Variation captured in PC1-4, B502 whole-brood analysis of forewings (left) and hindwings (right). Images are aligned with lowest scores in the component to the left; highest to the right; means in centre. Images reconstructed by AAMToolbox.
as expected, while several correlations fell just outside of our significance level (electronic supplementary material, tables S10 and S11). Within \( P_{\text{ele}}P_{\text{sil}} \) individuals, LG15 was excluded and the only significant association was between PC1 and LG21.

(c) Analysis of brood 472

B472 has just two genotypes, \( P_{\text{ele}}P_{\text{sil}} \) and \( P_{\text{sil}}P_{\text{sil}} \). Whole-brood analysis of forewings found significant associations with LG10 and the sex chromosome (electronic supplementary material, table S12), and these were matched in the hindwing study (table 2). Although there was a clear separation of the \( P_{\text{ele}}P_{\text{sil}} \) phenotype from the \( P_{\text{sil}}P_{\text{sil}} \) phenotype in the first PC, there was no correlation with LG15. This is because brood genotypes were based on female-informative markers, and the mother is homozygous for the \( P_{\text{sil}} \) allele: the morph is dependent upon the paternally inherited LG15. We used male-informative markers along the chromosome and found the strongest associations with markers close to the location of \( P_{\text{ele}} \) (figure 4a), suggesting that the variation captured in PC1 associates with the supergene.

We found that the female-informative LG15 interacted with LG17 and LG08 in contributing to captured variation in the forewings and hindwings of B472, respectively. Interestingly, we also found contributions from other pairs of LGs, including LG10, LG13, LG18 and LG21 (electronic supplementary material, table S13), suggesting interactions involving loci on these LGs known to contain colour-pattern genes in other species.

Within \( P_{\text{ele}}P_{\text{sil}} \) forewings, PC1 correlated most significantly with LG21, while PC7 associated with LG10 and LG18 (table 2 and electronic supplementary material, table S14), and captured subtle variation in subapical and postdiscal melanic scale distribution. The sample size of \( P_{\text{sil}}P_{\text{sil}} \) individuals was insufficient to perform ANOVA, and we were not able to compare the effect of different LGs between morphs in the same brood. However, we note interesting differences between the LGs of significance in the \( P_{\text{ele}}P_{\text{sil}} \) analysis from B472 and the \( P_{\text{ele}}P_{\text{sil}} \) analysis from B502: LG10 and LG18 have significant associations in the former, while LG01, LG19 and LG21 have significant associations in the latter (table 2).

(d) LG effect sizes

For the whole-brood analysis of B502 forewings, we found the greatest overall effects from LGs 10, 13, 15 and 21. A similar trend was seen for the corresponding hindwing analysis (figure 2d). These measures derive from the coefficient of determination, and reflect how much of the variation captured by the model can be attributed to each LG; differences in the size of the effect may reflect differences in the selection pressures on the hindwings and the forewings, or dorsal and ventral surfaces [27]. We also found interesting differences in

\[
\begin{array}{llll}
\text{R}_{\text{pP40}} & \text{<0.001} & \text{P} & \text{Hm08} \\
\text{R}_{\text{pP30}} & \text{<0.001} & \text{p} & \text{Hm01} \\
\text{Hm01} & \text{0.015} & \text{MRSP} & \text{Bm44} \\
\text{Hm08} & \text{0.539} & \text{Ci} & \text{Hm01} \\
\text{Hm08} & \text{0.539} & \text{Ci} & \text{Hm14} \\
\end{array}
\]
LG contributions between $P_{\text{sex}}$ and $P_{\text{sex}}$ (B502) and $P_{\text{sex}}$ (B472, electronic supplementary material, figure S6); while there was very little overall effect of LG19 on $P_{\text{sex}}$ wings, variation in this morph had a much greater contribution from LG16 and LG18. These differences were supported by hinting analyses.

(c) Mapping wing phenotype

Segregation of LG18 showed a significant association with PC7 in our analysis of $P_{\text{sex}}$ and PC6 of our within-male analysis of B502. We chose this LG for more detailed mapping because it contains the $B-D$ loci of *H. melpomene*, whose role has been characterized and genomic position has been mapped to within a centiMorgan [20]. Through male-informative markers, we found that the locus contributing to pattern variation in B472 appears to be in the same chromosomal location as $B-D$ (figure 4). We also used male-informative markers to begin mapping the location of QTL on LG01 and LG10 (figure 4c). For the former, we used B502 whole-brood forewing PC9, and found that the locus contributing to pattern variation was between the *wingless* and *dopa decarboxylase* markers. Using PC4 of the same analysis, our male-informative markers on LG10 suggest that this LG may contain two QTL; one close to *Hm17* and the other close to *Rpl13*. We note, however, that our marker order is only weakly supported and further mapping may support the alternative hypothesis that this LG contains only one QTL.

4. DISCUSSION

Butterfly wing patterns are now central to the study of natural and sexual selection. In *Heliconius* they are used to advertise unpalatability through aposematic colouring and as a cue for mate choice [28]. In *H. numata*, warning coloration is polymorphic; multiple discrete forms may be found in a single location and the different morph phenotypes are controlled by different alleles of the supergene $P$. To quantify the variation associated with the mimetic supergene and to examine the identity of loci interacting with it, we began by analysing wing pattern variation within a brood with four distinct morphs. We found that the first and second PCs of variation were significantly associated with LG15. The variation these components captured clearly relate to changes at $P$ [15], and therefore to mimicry determination, showing that PCs 1 and 2 recover expected chromosomal effects of the supergene. PC1 separated the recessive silvana wings from the *aurora* and *elegans* wings, while PC2 discriminated between the *aurora* and *elegans* phenotypes. The third PC, associated with the sex chromosome, captured a proximodistal migration of pattern elements. This probably reflects deformation of the wing caused by the presence of specialized androconial scales on the forewings and hindwings of males [29]. In the satyrid butterfly, *Bicyclus anynana*, differences in eye-spot size are consistent with one or more sex-linked loci [30], and the X-chromosomes are known to control several quantitative traits in other Lepidoptera [31], but none of the genetic loci previously associated with wing patterning in *Heliconius* is sex-linked [32]. Clearly, sex has an effect on the wing pattern of *H. numata*. This is an intriguing finding; it would be interesting to know if this variation reflects differences in mimetic resemblance, and, consequently, differences in the mimetic ecology of males and females [33].

(a) Evolution of the P supergene

Major loci that control patterning in other *Heliconius* species, such as $B$, $D$ and $K$, do not co-segregate with mimicry-ring association in *H. numata*. Instead, $P$ controls wing pattern through recruitment of a conserved set of loci, the putative orthologues of $Yb$, $Sb$ and $N$, which exist in close but incomplete linkage in the genomes of *H. melpomene* and *H. cydno*. Chromosomal rearrangements of the order of 200–400 kb around $P$ appear to have suppressed recombination to form the tightly linked supergene [34]. While the effects of other patterning loci are expected to have been downregulated as more pattern effects were transferred to the $P$ region, the drive towards inheritance of pattern by a single locus probably results from the atypical selection pressures associated with multiple mimicry [21]: different individuals live in different mimetic environments, with spatial variations of models at a small (5–10 km) scale producing spatially variable selection and promoting adaptive polymorphism [15]. Single-locus inheritance of pattern allows *H. numata* to maintain mimicry polymorphism in the variable environment this creates [35]. Although other examples of supergenes exist in *Heliconius*, such as $D$ in *H. melpomene* [36], control of almost the entire colour polymorphism by a single supergene appears to be unique to *H. numata*. Other examples of these features are found in snails [37], ladybird beetles [38] and vinegar flies [39], as well as in the more familiar cases of Papilio Batesian mimicry polymorphisms [40].

Where supergenes control mimicry, theoretical models suggest that additional modifier genes of small effect can be accumulated to improve the mimicry established by a mutation with large effect [5,7,41]. These classical unlinked modifier genes should be fixed in any one geographical location, where they improve the mimicry of a single model, and do not show polymorphism in natural populations. Their effects on mimetic resemblance can be revealed through crosses between individuals from different populations, disrupting co-adapted gene complexes [7,42]. The broods in our study were derived from parents collected from the same population and reveal within-population variation; we therefore do not expect to detect fixed modifiers. However, other classes of modifier genes have been described that give some degree of mimetic resemblance to a model or convert mimetic genotypes into mimics of other models [7]. Whether these loci are eliminated or remain segregated may depend on the model abundance being below a threshold level. As a result, where multiple mimicry alleles coexist in a polymorphic population, we may find genes of minor effect with morph-specific effects.

Natural populations of *H. numata* show pattern variation independent of allelic variation at $P$, suggesting that variable loci do affect the expression of each morph. This variation, which we began to map in this study, provides the raw material for the refinement of mimetic resemblance, and may be considered the ’small residual effect of background’ that has been described elsewhere for interracial crosses [36]. It would be interesting to directly test the hypothesis that these genes improve the mimicry established by $P$, but given the overall accuracy of mimicry covariation seen in other mimicry rings [32], and the overall strength of selection for mimicry [36], we might expect that subtle, quantitative effects on wing pattern will come under
selection by predators. Alternatively, it could be the case that quantitative variation is inevitable given the identity and interactions of major effect genes, and it is presently unknown whether the maintenance of multiple morphs within a population incurs a cost in terms of accuracy.

(b) Mapping minor-effect QTL
Our quantitative morphometric analysis suggests that multiple unlinked loci have a significant although minor effect on pattern variation in *H. numata*. Our findings also suggest that known colour-pattern loci may be among the QTL detected. For instance, using male-informative markers, we found that a QTL on LG18 maps close to the same region to which the $B-D$ loci have been mapped in *H. melpomene* [32]. Population genetic studies have recently provided evidence for strong selection at this site but showed no indication of a recent selective sweep [19], supporting the hypothesis that the colour-pattern locus is ancient.

Our study also found an association between captured variation and LG10, which is known to contain $Sd$ in *H. erato*, a locus that controls those parts of the forewing that are covered in melanic scales [36]. $Sd$ is syntenic with another colour-pattern locus, $Ac$, in *H. cydno* [43,44]. Consistently, LG10 was associated with PCs that captured variation in the distribution of melanic patches in the apical and discal parts of the *H. numata* wing. Similarly LG01, which contains $K$ in *H. cydno*, was found to control pattern variation in B502, and we used markers on these LGs to begin mapping the QTL we identified. Unfortunately, although the locations of $Sd$ and $K$ have been determined with amplified fragment length polymorphisms to within a few centiMorgans on the equivalent chromosome, the locations of other LGs to begin mapping the QTL we identified. Unfortu-

(c) Modifier effects of minor QTL
Analyses within morphs suggest that LG19 and other LGs contain QTL with morph-specific effects. LG01, for example, has a significant association with variation in the wings of *Paur* individuals but has no such associations in the major PCs when the whole of brood 502 was examined, suggesting there is epistasis between the QTL on this LG and the allele at $P$. Similarly, a locus (or loci) on LG18 contributes to variation in *P ele* but has no significant associations with variation in *Paur*. If known loci such as $B-D$ are responsible for the variation we have measured, their role is clearly not to control the presence/absence of major pattern elements as they do in other species. The contributions observed from LG18 are consistent with the type of pleiotropic and epistatic effects described previously for colour-pattern loci [17,36]. Our LG-by-LG analyses found several other interesting interactions, including effects from LG18 that had no significant associations when analysed individually. These represent previously undetected QTL with epistatic effects, which may be specific to a certain allele of the supergene or show population-specific interactions and act with multiple alleles of $P$.

Interactions between loci are clearly important to colour patterning in *Heliconius* [17]. In addition to epistasis, the dominance effects of many toolkit loci have been described, and relate to the type of scale cell that develops on the wing; the yellow hindwing bar of *H. melpomene* is caused by ‘type I’ scales that are recessive to black type II scales [17,32]. The red band on an otherwise black forewing, controlled by $B$, is produced by xanthommatin pigments expressed in a third scale type, which is dominant to types I and II [49]. These relations demonstrate the hierarchy that exists at the scale-level in *Heliconius*, and can apply to inter-locus epistasis where two different loci influence the same region of the wing [22,50]: an example exists in *H. melpomene*, where the red forewing band is moved distally through interactions with a dominant allele at an unlinked locus, $N$ [17]. Although there are different hierarchical relationships in *H. numata*, it is interesting to note that the QTL we have identified that influence the distribution of melanin and other scale types may have important additive or epistatic effects, just as major colour pattern loci do in other species [17,32]. LG10 is particularly interesting because it was found to associate with multiple PCs capturing variation in the distribution of melanic scales. Because $Sd$ affects melanin distribution in *H. erato*, we might expect to find this relationship between the LG and the variation with which it associates.

In *H. numata*, the mimicry supergene $P$ has evolved through the recruitment of a closely linked set of colour-pattern loci. There have been no major translocations or physical reorganizations of the chromosomes or patterning genes; instead, the genotype-to-phenotype map has been tuned such that it allows one locus in the multi-locus toolkit to take over. In this study, we have applied a novel morphometric approach to capture variation in the wing patterns displayed by *H. numata* and found that, with the exception of the minor-effect *H. melpomene* QTL locus [26], all of the chromosomes containing the other known *Heliconius* colour-pattern loci have retained subtle quantitative effects. In some cases, the effects appear to be specific to certain morphs and epistatic to the supergene, suggesting that they effectively act as modifier genes for wing pattern. We also found contributions from chromosomes not previously known to contain a colour-pattern locus, which are, therefore, good candidates for the identification of additional genes with modifier
effects. If any of the QTL we identified do include members of the *Heliconius* ‘toolkit’, this raises the interesting possibility that the same loci can be recruited again by natural selection for mimetic change at a later stage. This would provide a possible mechanism for the parallel evolution of mimicry using conserved sets of genes, as is the case in mimetic butterfly clades [9,20,43].

We thank Simon Baxter, Andrew Hanna, Enrico Coen, Annabel Whibley, Richard Merrill, Laura Ferguson and Christine Crossan. We are grateful to four anonymous reviewers for their key inputs to the development of this manuscript, and INRENA (Peru) for granting collecting and export permits (076-2007-INRENA-IFFS-DCB).

REFERENCES