PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research

Cite this article: Jay P, Leroy M, Le Poul Y, Whibley A, Arias M, Chouteau M, Joron M. 2022 Association mapping of colour variation in a butterfly provides evidence that a supergene locks together a cluster of adaptive loci. *Phil. Trans. R. Soc. B* **377**: 20210193.
https://doi.org/10.1098/rstb.2021.0193

Received: 6 December 2021
Accepted: 21 March 2022

One contribution of 15 to a theme issue 'Genomic architecture of supergenes: causes and evolutionary consequences'.

Subject Areas:
evolution, genomics, genetics, ecology

Keywords:
inversion, association study, multivariate association, wing colour pattern, cluster of adaptive loci, divergence hitchhiking

Authors for correspondence:
Paul Jay
e-mail: paul.yann.jay@gmail.com
Mathieu Joron
e-mail: mathieu.joron@cefe.cnrs.fr

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5983511.

THE ROYAL SOCIETY PUBLISHING

Association mapping of colour variation in a butterfly provides evidence that a supergene locks together a cluster of adaptive loci

Paul Jay[1], Manon Leroy[1], Yann Le Poul[1], Annabel Whibley[2], Mónica Arias[3,4], Mathieu Chouteau[1,5] and Mathieu Joron[1]

[1]CEFE, Université de Montpellier, CNRS, EPHE, IRD, 34293 Montpellier cedex 5, France
[2]School of Biological Sciences, University of Auckland, Auckland 1010, New Zealand
[3]CIRAD, UMR PHIM, F-34398 Montpellier, France
[4]PHIM, Univ Montpellier, CIRAD, INRAE, Institut Agro, IRD, CEDEX 5, 34398 Montpellier, France
[5]LEEISA, USR 63456, Université de Guyane, CNRS, IFREMER, 275 route de Montabo, 797334 Cayenne, French Guiana

PJ, 0000-0001-5979-1263; MA, 0000-0003-1331-2604; MJ, 0000-0003-1043-4147

Supergenes are genetic architectures associated with discrete and concerted variation in multiple traits. It has long been suggested that supergenes control these complex polymorphisms by suppressing recombination between sets of coadapted genes. However, because recombination suppression hinders the dissociation of the individual effects of genes within supergenes, there is still little evidence that supergenes evolve by tightening linkage between coadapted genes. Here, combining a landmark-free phenotyping algorithm with multivariate genome-wide association studies, we dissected the genetic basis of wing pattern variation in the butterfly *Heliconius numata*. We show that the supergene controlling the striking wing pattern polymorphism displayed by this species contains several independent loci associated with different features of wing patterns. The three chromosomal inversions of this supergene suppress recombination between these loci, supporting the hypothesis that they may have evolved because they captured beneficial combinations of alleles. Some of these loci are, however, associated with colour variations only in a subset of morphs where the phenotype is controlled by derived inversion forms, indicating that they were recruited after the formation of the inversions. Our study shows that supergenes and clusters of adaptive loci in general may form via the evolution of chromosomal rearrangements suppressing recombination between co-adapted loci but also via the subsequent recruitment of linked adaptive mutations.

This article is part of the theme issue 'Genomic architecture of supergenes: causes and evolutionary consequences'.

1. Introduction

Recombination is a central force in evolution, allowing the shuffling of genetic diversity and continually exposing new combinations of alleles to natural selection. However, recombination also has a homogenizing effect on diversity, breaking apart beneficial combinations of alleles and preventing alternative combinations to persist through time. This is illustrated in heterogeneous environments, where recombination shuffles combinations of alleles evolving under divergent selection in populations connected by gene flow. When local adaptation involves changes at multiple loci, spatial heterogeneity generates selective pressure favouring either lower or higher recombination rates among these loci, depending on the strength and direction of selection that these loci experience [1]. This is expected to lead to genomic variation in recombination rate between and within chromosomes, and may lead to the formation of clusters of locally adaptive loci [1,2].

Supergenes are clusters of adaptive loci associated with major phenotypic variation in many species, often coordinating changes in life-history, colour and behaviour in animals (e.g. in mammals, birds, fishes or insects [3–8]) or pollination traits in plants [9–11]. Supergenes are often formed by polymorphic chromosomal rearrangements such as inversions, which suppress recombination between standard and rearranged segments. Recombination suppression is thought to facilitate the maintenance of coadapted alleles in close linkage within a single population. Yet there is still little evidence that supergenes include multiple coadapted loci, and how and why such clusters of co-adapted loci may form is poorly understood [12–14]. The idea that rearrangements evolve because they 'capture' sets of beneficial alleles requires the existence of polymorphisms at the selected loci prior to rearrangement formation. An alternative route to supergene formation is the serial recruitment of co-adapted mutations within a non-recombining region initially containing a single locus under selection (a process related to the divergence hitch-hiking). In this case, the formation of coadapted haplotypes involving multiple loci follows the arrest of recombination.

Testing these hypotheses and gaining a better understanding of the evolution of supergenes requires finely dissecting both the phenotypic effects of the loci forming the supergenes and the origin of their linkage. This has proven difficult, first because of the complexity of untangling the genetic architecture of complex multi-dimensional traits, and second because recombination-mapping approaches are often inefficient to decipher the individual effects and evolutionary history of the loci in non-recombining regions [13]. As a consequence, except for the specific cases of self-incompatibility loci of plants [15–18] and mating-type chromosomes in fungi [19,20], the individual contributions of loci maintained in linkage disequilibrium by supergenes remain largely unknown [13]. This constitutes a major obstacle to our understanding of the evolution of supergenes and of genomes in general.

Neotropical butterflies in the genus *Heliconius* have been extensively studied over the past decades both ecologically and genetically. Most *Heliconius* species display a geographic mosaic of wing patterns, matching in every locality the patterns of coexisting local butterfly species (Müllerian mimicry). Four major chromosomal regions are known to underlie these variations [21–26]. Nevertheless, the genetic architecture of wing patterning varies between species and often involves a subset of these four regions. Crosses between different mimetic forms from distinct geographic regions often result in the formation of recombinant, non-mimetic phenotypes. This is observed naturally in the transition zones between the geographic ranges of mimetic forms [24,27], and the recurrent formation of these poorly protected individuals is expected to favour the evolution of clusters of wing pattern loci [2,28]. Consistent with this prediction, two independent inversions have evolved around the wing-patterning gene *cortex*, one in *H. numata* and *H. pardalinus* [29] and one in *H. sara*, *H. demeter*, *H. hecalesia and H. telesiphe* [30]. Those inversions represent major adaptive alleles and have been shown to flow between species [30,31].

Among these taxa, *H. numata* appears as an outlier. Indeed, besides its geographical wing pattern variation, *H. numata* also displays a striking polymorphism of colour patterns within populations (figure 1*b*). Variation in wing patterning in *H. numata* involves variation in the presence of certain wing pattern elements, such as a broad yellow band on the forewings, as well as more quantitative variation, such as the positional shift of certain colour patches or the spread of black patterning on hindwings (figure 1*b* and [33]). Up to seven morphs of *H. numata* can be observed within a single population, each one engaged in mimicry relationships with distinct toxic and non-toxic species. Non-mimetic morphs in *H. numata* are strongly selected against by bird predation [34,35], which should translate into selection on mechanisms limiting the formation of recombinant forms. Previous studies have shown that wing pattern diversity in *H. numata* is associated with three polymorphic inversions forming a supergene, called P, on chromosome 15 (as shown in figure 1*a*; [29,36]). In addition to the inversion capturing the gene *cortex* mentioned above, called P_1, this supergene also includes two other polymorphic inversions, P_2 and P_3, all in adjacent positions. Together the three inversions suppress recombination over a 3 Mb region encompassing 107 predicted genes [29]. These inversions were formed between *ca* 1.8 and 3.0 million years ago [29,31]. In natural populations, three chromosomal arrangements may be found (figure 1*a*): Hn0, the standard arrangement without any inversion; Hn1, the first derived arrangement carrying the P_1 inversion only; and Hn123, the second derived arrangement with the three adjacent inversions P_1, P_2 and P_3. Previous studies have found that derived, inverted haplotypes are dominant over standard, non-inverted haplotypes, i.e. individuals heterozygous for the rearranged chromosomes have similar phenotype to homozygotes for the same chromosomal arrangements (figure 1*a*; [32,33]). Because the *H. numata* P supergene spans a region repeatedly found to be associated with wing pattern variation in other Lepidoptera, including recombining wing patterning loci in *H. melpomene* [21,37,38], it has been hypothesized that the P inversions have evolved because they reduce recombination between several linked wing patterning loci and hamper the formation of maladapted recombinant phenotypes [36,38], but this remains to be demonstrated.

In *H. numata*, chromosomal arrangements Hn0 and Hn123 are associated with a variety of mimetic forms across the range as well as in sympatry [29,36], and haplotypes with the same arrangement should recombine normally in homozygous individuals (i.e. in Hn0/Hn0 or Hn123/Hn123 individuals). This should facilitate the identification of specific loci underlying wing pattern variation in *H. numata*. We therefore took advantage of the multiple morphs of *H. numata* sharing the same chromosomal arrangements at the colour pattern supergene to locate the loci associated with wing pattern variation. We re-sequenced the entire genomes of 131 specimens, used an unsupervised landmark-free algorithm to dissect their multi-dimensional wing pattern variation and performed genome-wide association studies to associate phenotypic and genetic variation. We show that multiple genomic intervals are associated with colour variation in *H. numata* and that all these regions are situated within the P supergene, supporting the hypothesis that the P inversions were recruited because of their role in maintaining beneficial combinations of wing pattern alleles. Several of these regions seem, however, to be involved in wing pattern variation only among a subset of forms harbouring inversions and not among forms without inversions, suggesting

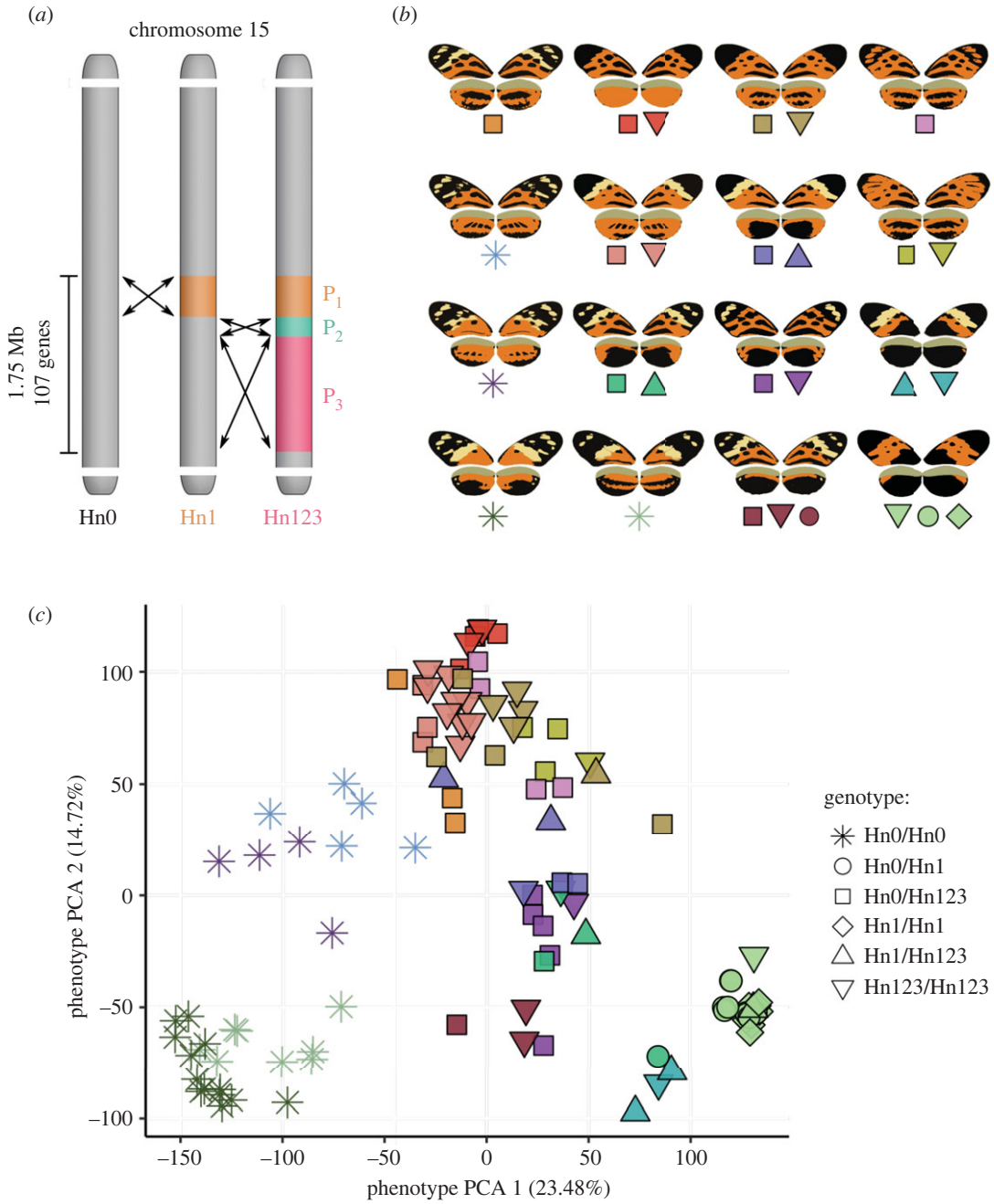

Figure 1. Genetic architecture and wing pattern diversity in *H. numata*. (*a*) Genetic architecture of the *H. numata* mimicry supergene P characterized by three polymorphic inversions of respective size 400 kb, 200 kb and 1150 kb. (*b*) Schematic diversity of wing patterns of *H. numata* in our dataset. (*c*) Two-dimensional approximation of the morphological space representing the phenotype diversity observed in *H. numata*. The dotplot displays results from a principal component analysis (PCA) (the first two components are displayed here) computed on wing pattern variations as obtained using colour pattern modelling (CPM) [32]. For display purposes, butterflies were manually classified into mimetic forms based on the literature [33]; different forms are depicted by different colours. The butterflies sampled for this study represent the commonest forms observed in *H. numata*. Different supergene genotypes are depicted by different symbol shapes. Results for PC 3 and PC 4 are presented in the electronic supplementary material, figure S8. PCAs computed on specimens with the same supergene arrangement and on specific parts of wing pattern are presented in the electronic supplementary material, figure S9. (Online version in colour.)

that their involvement in *H. numata* colour variation evolved after the formation of inversions. Our study therefore suggests that the P supergene has formed via the evolution of chromosomal rearrangements suppressing recombination between co-adapted loci but also via the subsequent recruitment of linked adaptive mutations.

2. Results

In order to decipher the evolutionary stages of the formation of the supergene P, we re-sequenced with a *ca*. 30x coverage 131 *H. numata* individuals classified into 16 mimetic forms according to the literature (figure 1, [33]). Reads were mapped against the *H. melpomene* reference genome (Hmel 2). Following previous studies [29,39–41], we used principal component analyses (PCAs) to identify individual genotypes at the supergene (electronic supplementary material figures S1–S4 and table S1) and found 39 specimens homozygous for the standard chromosomal arrangement (Hn0/Hn0), 20 homozygous or heterozygous for the first derived chromosomal arrangement (8 Hn1/Hn1 and 11 Hn1/Hn0) and 72 homozygous or heterozygous for the second derived chromosomal arrangement (27 Hn123/Hn123, 37 Hn0/Hn123 and 8 Hn1/Hn123).

For each chromosomal rearrangement, the occurrence of homozygous individuals in substantial proportions in natural populations [42] suggests that haplotypes with the same chromosomal arrangements should recombine normally. Patterns of linkage disequilibrium along the supergene supported this hypothesis: we found that the level of linkage disequilibrium within groups of individuals homozygous for each of the three chromosomal arrangements (Hn0/Hn0, Hn1/Hn1 and Hn123/Hn123) was only slightly higher within the supergene than in flanking regions (electronic supplementary material, figure S6). Furthermore, phylogenetic topologies shift repeatedly along the inversions (electronic supplementary material, figure S7), consistent with ongoing recombination within each class of arrangement. By contrast, linkage disequilibrium within the supergene was much stronger than in collinear regions when we considered groups of individuals with alternative chromosomal arrangements, indicating recombination suppression between segments in different orientations (electronic supplementary material, figure S6).

Morphometric analyses were run on 109 specimens whose wings were in good condition, using colour pattern modelling (CPM), an algorithm for the quantification of colour pattern variations based on colour classification and colour pattern registration [32]. By providing naive descriptors of colour variation, CPM overcomes the limit and bias of phenotype description by human observers and, for instance, permits quantification of positional shifts of bi-dimensional colour patches. Briefly, starting from standardized pictures, wings were first extracted, their colours clustered into black, orange or yellow (the three colours present on the wings), and then aligned with each other on the basis of pattern similarity. Each pixel common to all aligned wings was considered a variable, resulting in the description of wing pattern variation by *ca* 10^5 variables. PCAs were used to reduce the high dimensionality of colour variation and showed that wing pattern polymorphism involves a mixture of qualitative and quantitative variation (figure 1*c*; electronic supplementary material, figures S8 and S9). Wing pattern polymorphism in *H. numata* indeed involves variations on different parts of the wing, with some features appearing more discrete (e.g. the presence of a broad yellow band) than others (e.g. the spread of hindwing black patterns; figure 1; electronic supplementary material, figures S8 and S9). The first phenotypic principal components were observed to describe obvious colour variations. For instance, the first principal component appears to quantify the overall amount of black patterns and the third principal component to describe the size of the broad yellow band on the forewing or its absence (figure 1*c*; electronic supplementary material, figures S8 and S9). As expected, specimens harbouring the same chromosomal arrangement were clustered in the phenotype space, highlighting that supergene inversions are major determinants of specimen phenotype (figure 1*c*; electronic supplementary material, figures S8 and S9). By contrast, no wing pattern feature was associated with a specific geographic locality and the origin of specimens appeared as a poor descriptor of individual phenotype (electronic supplementary material, figure S10).

In accordance with a previous study [43], PCAs and differentiation analyses (*F*st) showed a near absence of genetic structure in *H. numata* across the Neotropics (electronic supplementary material, figures S6 and S12–S13; see also [43]). For instance, the genomic differentiation (*F*st) between populations from French Guiana and from Peru was only of 0.02 despite being separated by approximately 2900 km (electronic supplementary material, figure S13). We found, however, a relatively strong differentiation (*F*st = 0.27) between the population from the Brazilian Atlantic forest and the populations from the rest of the range (electronic supplementary material, figure S12; see also [43]). Because this geographic structure could confound genotype–phenotype association studies, we removed specimens from the Brazilian Atlantic forest for subsequent analyses. In total, 100 phenotyped and genotyped specimens were retained for the genotype–phenotype association study.

In order to identify the loci associated with wing pattern variation, we used MV-PLINK [44] to perform multivariate genotype–phenotype associations using as phenotype the first six principal components describing the joint variation of fore and hind wings (variance explained: 58.08%) in the entire sample set and using all biallelic sites (36 928 374 sites with point mutations or indels). As expected following previous studies [21,29,36], the main region of association corresponded to the P supergene (figure 2*a*). In the analysis performed with all specimens regardless of their genotype at the supergene, owing to the absence of recombination between segments in opposite orientation generating strong linkage disequilibrium, a large number of mutations in the supergene were highly associated with wing pattern variations (figure 2*b*). In order to disentangle the effect of loci within the supergene and to remove the confounding effect of the inversions, we performed phenotype–genotype association studies separately on the specimens homozygous for chromosome arrangement Hn0 (no inversion, $n = 22$, hereafter referred to as Hn0 specimens) and on those homozygous or heterozygous for chromosome arrangement Hn123 (containing inversions P_1, P_2 and P_3, $n = 61$, hereafter referred to as Hn123 specimens). The dominance of inverted alleles to non-inverted alleles [32] allows the use of both homozygous and heterozygous specimens in the analyses. Using individual-based simulations mimicking the evolution of *H. numata*, we confirmed that this sampling and analysis scheme can support the detection and resolution of multiple wing pattern loci within a genomic region similar to the P supergene (see §4 and electronic supplementary material, figure S22). Specimens with an Hn1/Hn1 or Hn1/Hn0 genotype were not included in this analysis because they are all phenotypically very similar and do not display the combinatorial variation of wing pattern features that we aim to analyse (figure 1*c*; electronic supplementary material, figure S8).

Analysing genotype–phenotype associations separately on the Hn0 and Hn123 groups of specimens, we found that genetic variants located in different regions of the supergene were strongly associated with wing pattern variation (figure 3; electronic supplementary material, figures S14–S19). We first computed multivariate associations using two to six phenotypic principal components (figure 3*a*,*d*; electronic supplementary material, figures S14 and S15) and determined the significance of the association by performing 10^6 permutation tests [45]. Because selection around a causal mutation is expected to strengthen its linkage disequilibrium with nearby mutations, we should observe multiple associated variants in regions surrounding causal variants. To identify regions most likely to be involved in wing pattern variation, we therefore computed the density of associated variants along the genome in sliding windows, considering

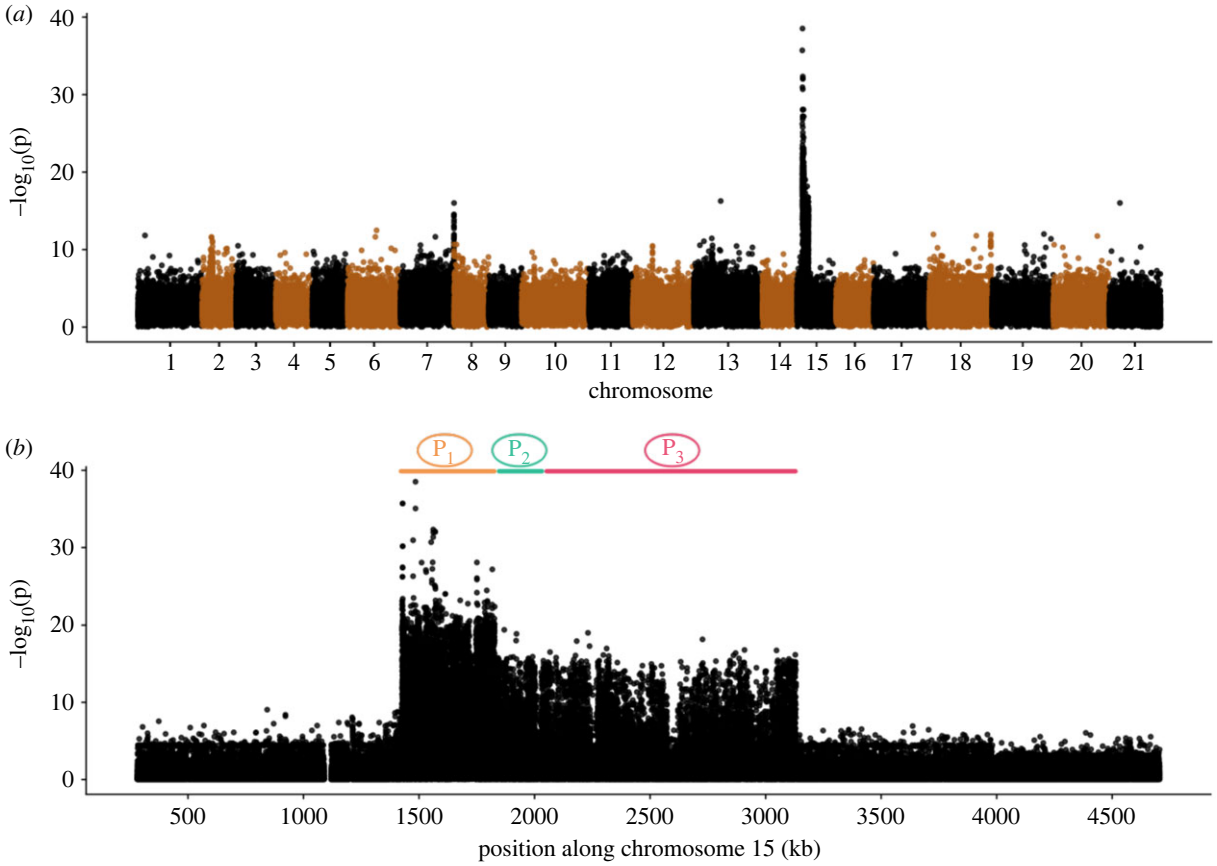

Figure 2. Genome-wide association of genetic and wing pattern variation. (*a*) Multivariate association study using as phenotype the first six principal components describing wing pattern variations (presented in figure 1*c* and electronic supplementary material, figure S8) and all specimens regardless of their genotype at the supergene. One major peak of association in noticeable on chromosome 15 corresponding to the supergene. One minor peak can be seen on chromosome 7. This is due to an assembly error in the reference genome Hmel2, and in reality, this region lies within the supergene region (see §4). (*b*) Focus on the peak of association on chromosome 15, corresponding to the position of the three polymorphic chromosomal inversions P_1, P_2 and P_3. (Online version in colour.)

only variants with empirical p-value $< 1 \times 10^{-6}$ (i.e. variants for which no permutation resulted in a better association). Both in the Hn0 and Hn123 sample sets, we found that several regions within P_1, P_2 and P_3 were enriched in variants associated with wing pattern variation (figure 3; see §4). Several regions were associated with phenotype changes in both sample sets (figure 3 and electronic supplementary material, figure S14–15 and table S2). For instance, the same intron in the gene *parn* was associated with colour variation in Hn0 and Hn123 specimens. Among the 448 variants with empirical p-value $< 10^{-6}$ in Hn123 multivariate associations, only 23 are non-synonymous mutations (1/178 in Hn0 associations). In others words, associated variants fell primarily in non-coding intergenic or intronic sequences (electronic supplementary material, table S2).

To confirm that associations were not caused by confounding cryptic geographic structure in our dataset and to take into account putative dominance effects of wing pattern loci, we also performed univariate associations using as covariates two genotype-dependent variables (for dominance effects) and the first 10 principal components describing whole-genome genetic structure (for population stratification effects, see §4 and electronic supplementary material, figure S5; [46]). Because univariate analyses can only consider a single variable at a time, they cannot identify variants that are associated with a phenotypic feature that would only be described by a combination of phenotypic principal components (in contrast with multivariate analyses). To analyse wing pattern variation in as much detail as possible with univariate analyses, in addition to the previous phenotypic PCAs performed on the entire wing phenotype, we therefore also computed phenotypic PCAs focusing on six partitions of wing pattern variation: variation found on hindwings only, variation found on forewings, variation in yellow patterning, variation in the forewing tip, variation in the forewing middle part and variation in the forewing base (electronic supplementary material, figure S9, see §4). We computed univariate analyses using each of the first three principal components of these PCAs as phenotypes. These analyses revealed sizeable peaks of association with different features of wing pattern located in different regions of the supergene (electronic supplementary material, figures S16–S19; figure 4). The vast majority of regions showing an association with wing pattern variation in these univariate analyses with dominance effect and population stratification were also associated in the previous multivariate analyses, confirming the low impact of population structure in our analyses (electronic supplementary material, figures S14 and S19). Computing the density of associated variants in these univariate analyses along the genome in sliding windows, we found the same regions enriched in wing pattern-associated variants as those found in multivariate analyses (figure 3).

To identify the wing pattern features affected by each associated genomic region, we computed the effect of the most strongly associated genetic variants (single nucleotide polymorphisms (SNPs) or indels) on colour variation at each wing position (image pixel). This can be visualized in the form of heatmaps of variant effects on the wings, where

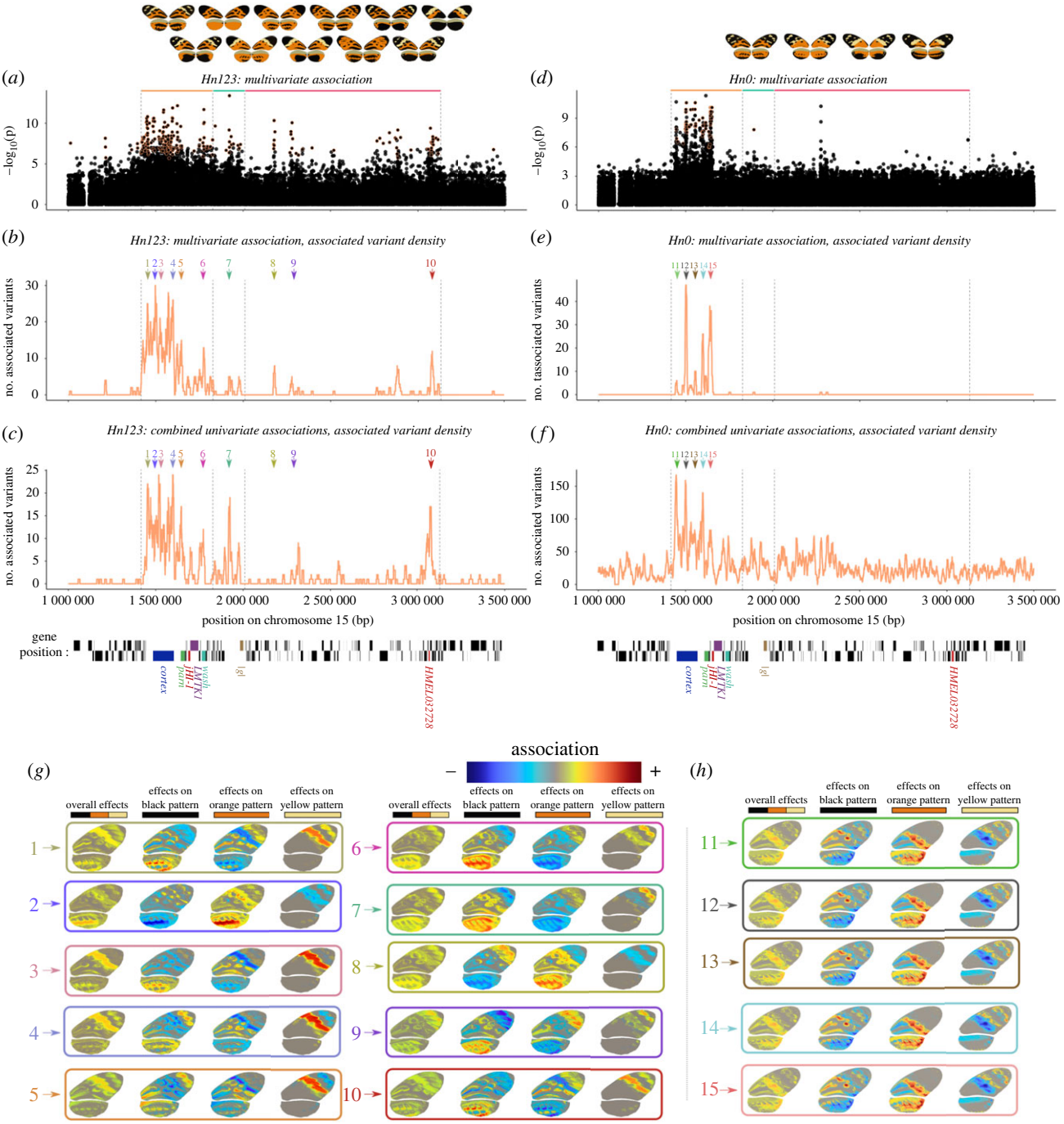

Figure 3. Distinct regions within the supergene are associated with variation in wing pattern features. (*a*,*d*) Multivariate association studies computed on Hn123 and Hn0 specimens, respectively, on the entire wing pattern variation (hindwing and forewing together, here using four principal component as multivariate phenotypes). The plotted *p*-value is the statistical *p*-value from the multivariate test of association. 1×10^{-6} permutations were performed for each variant. Variants highlighted in orange are variants with an empirical *p*-value $< 1 \times 10^{-6}$ (i.e. for which no permutation resulted in a lower statistical *p*-value). The positions of inversion breakpoints are represented by the dotted vertical lines. Associations computed with different numbers of phenotypic principal components are presented in electronic supplementary material, figures S14 and S15. (*b*,*e*) Density of significantly associated variants in multivariate analyses (with empirical *p*-value $< 1 \times 10^{-6}$) along the chromosome 15. Analyses were computed in 10 000 bp overlapping sliding window (with 100 bp overlap). All significantly associated variants in one or more of the multivariate association analyses (using 2, 3, 4, … , 6 phenotypic principal components; electronic supplementary material, figures S14 and S15) were used. (*c*,*f*) Density of significantly associated variants in univariate analyses (with empirical *p*-value $< 1 \times 10^{-6}$) along chromosome 15. Analyses computed in 10 000 bp overlapping sliding windows (with 100 bp overlap). All significantly associated variants in one or more of the univariate association analyses (focusing on different part of the wing and using the first, second or third using phenotypic principal component as phenotype; electronic supplementary material, figures S16–S19) were used. (*g*,*h*) Phenotypic effects of the top variant from each of the 15 regions that displayed a clear enrichment in significantly associated variants ((*b*–*f*) coloured arrows) to the wing pattern in Hn123 or Hn0 specimens, respectively. Heatmaps from blue to red represent, for every pixel, the strength and direction of association of the derived allele, that is how allelic change at a given genetic position affects this wing area. Overall effects are shown as well as colour-specific effects, the latter representing the extent to which allelic change is associated with the presence or absence of each colour at this wing area. Because blue and red represent the direction of the association, opposite directions (i.e. red and blue values) in the same wing area in two colour-specific heatmaps indicate that the focal locus is associated with a change from one colour to the other in this area. For instance, if the effect of a genetic variant on a given wing area is highlighted in blue when looking at the orange pattern but in red when looking at black pattern, that means that change at this variant is associated with a switch from orange to black at this wing area.

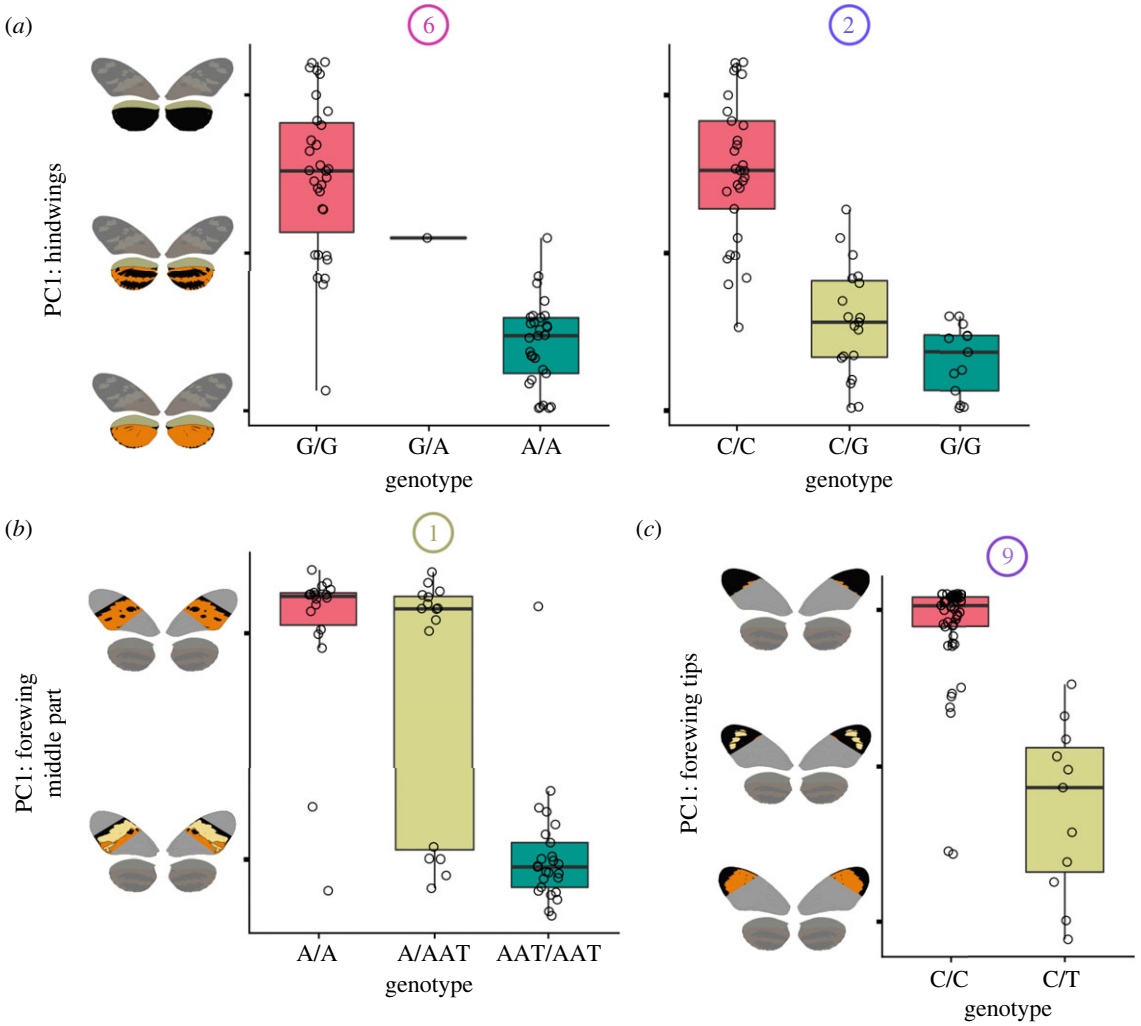

Figure 4. The effect of four selected variants on Hn123 wing pattern variation. Representation of the association of some genetic variants with specific wing pattern variation. See figure 3*g,h* and electronic supplementary material, figures S14–19 for additional representations of the association with specific aspect of wing patterns. The first principal components of analyses computed on different parts of the wing are used as proxy for the phenotype (*y*-axis): PCA computed on hindwings only (*a*), on the middle part of the forewings only (*b*) and on the tips of the forewings only (*c*). See electronic supplementary material, figure S9 for another representation of the principal components. Each dot is an individual specimen. Instead of annotating *y*-axis with eigenvalues (values of individual on principal components), schematic butterflies with average phenotypes along the principal components are displayed. Boxplot elements: central line, median; box limits, 25th and 75th percentiles; whiskers, 1.5 × interquartile range. (*a*) Effect of the most strongly associated variants in regions 6 (*cortex*, intron 2) and 2 (*wash*, intron 2) on the amount of black on hindwings. (*b*) Effect of the most strongly associated variants in region 1 (intergenic *HMEL032679-cortex*) on the forewing middle part. (*c*) Effect of the most strongly associated variants in region 9 (intergenic *HMEL022251-HMEL032698*) on the tip of the forewings. (Online version in colour.)

colour hue and brightness reflect the direction and the strength of the effect. Because each pixel can take three different colours (black, yellow and orange), the effect of variants on the presence or absence of these colours can be visualized for each colour independently but can also be summarized into an overall, non-directional effect per pixel (figure 3*g,h*). The limited phenotypic variation among Hn0 individuals did not enable proper dissection of the different wing pattern features in those forms (figure 3*h*), and therefore all associated variants were found to affect the wing pattern similarly. By contrast, in specimens with the three inversions P_1, P_2 and P_3, we found that associated variants in the different regions of the supergene affected different features of the wing phenotype (figure 3*g*). For instance, we found variants respectively associated with the presence of a broad yellow band on the middle of the forewing, with yellow patches on the tip of the forewing and with the variation in size of black patches on the hindwing (figures 3 and 4; electronic supplementary material, table S2).

Closely linked regions of the same gene were sometimes observed to correlate with different features of wing patterns. For instance, some regions of the gene *cortex* were associated with changes in yellow features on the forewing whereas closely linked regions of the same gene were associated with changes in black features on the hindwing (figures 3 and 4; electronic supplementary material, table S2). By contrast, several variants within the supergene were associated with similar variations of the wing pattern. The co-association of distant genomic regions with similar phenotypic variation in Hn123 specimens could be caused by the correlation of multiple phenotype features in our dataset—i.e. the non-uniform phenotypic variation among Hn123 specimens (figure 1*c*)—leading the genetic variants controlling these independent features to show the same effect (as observed in a much stronger proportion in Hn0 specimens). Co-association of closely linked variants could be simply caused by their linkage disequilibrium, owing to our relatively reduced sample size. This notably could be the case

around the gene *cortex*, which presents a high density of associated variants with similar phenotypic effect. Finally, co-associations of loci with similar phenotypic effect may also suggest that these loci are epistatic.

To estimate whether genetic variants associated with phenotypic variation may indeed control wing pattern development, we used previously published RNAseq data generated on *H. numata* [29,47]. Using edgeR [48], we compared the expression of genes in pupal wing-discs between specimens homozygous for Hn0, homozygous for Hn1 and homozygous for Hn123. We found that, among others, the genes *cortex*, *parn*, *wash*, *jhI-1* and *HMEL032728* display significant expression differences (electronic supplementary material, figure S10). These genes are within the supergene region and all include or are very close to sites associated with wing pattern variation in our analyses (figure 3; electronic supplementary material, table S2). Their differential expression during wing development strongly suggests that they participate in the control of wing pattern variation in *H. numata*.

3. Discussion

In order to identify the genetic variants underlying wing pattern variation in *H. numata*, we quantified these patterns with CPM, a method that produces comprehensive and 'naive' descriptors of colour pattern variation. This method offers the advantage of handling characters that cannot be defined easily by human observers or using landmarks, such characters involving shape changes or colour patch translation. CPM is therefore powerful for describing multi-dimensional traits such as wing patterns. However, it requires the use of multivariate association methods to identify the genetic variants involved in these multi-dimensional character changes. Performing genome-wide multivariate association studies, we found multiple genomic regions associated with different features of wing pattern variation within the supergene interval. This observation was confirmed by univariate association analyses taking into account dominance effect and population stratification. This lends support to the long-standing hypothesis that the supergene P coordinates the variation of multiple, independent sites each associated with specific elements of wing pattern. Our results therefore support the 'supergene' or 'beads-on-a-string' model [12,13,49] and are generally inconsistent with the alternative hypothesis that the supergene could involve a single master gene with pleiotropic effects [50].

Since our findings are based on genotype–phenotype correlations, we cannot exclude that only some of the loci associated with wing pattern variation in our dataset would be functionally involved in wing coloration. An association could result, for instance, from a correlation between wing pattern traits and certain unmeasured traits. In this species, wing pattern forms share the same microhabitats and display similar behaviour [51]; however, they tend to mate disassortatively, i.e. females preferentially mate with males displaying a different wing pattern [42]. Depending on how this preference is genetically determined [52], it might cause spurious genotype–phenotype associations. For instance, if a locus controlling a given wing pattern feature is in linkage with a locus inducing mate rejection based on such a feature during courtship, genetic mapping may associate both loci with wing pattern variation. Nonetheless, this is unlikely to result in multiple non-causal associations. Moreover, such associations would reveal coordination of wing pattern variants and mate choice variants, which still stress the importance of chromosomal rearrangements in maintaining co-adapted loci in linkage.

Some of the wing pattern-associated regions include genes displaying RNA expression differences during wing development when comparing mimetic forms, suggesting that they are indeed involved in wing patterning. Among these genes, *cortex*, *parn* and *wash* are also associated with wing pattern variation in other *Heliconius* species [21,26,37,53], although only the roles of *cortex* and *wash* have been experimentally validated. Our results show that these genes associate with very different features of wing pattern in *H. numata* and also highlight that several other loci, such as the genes *parn*, *jhI-1* and *HMEL032728*, may also play important roles in wing patterning. These genes regulate general processes such transcription or cell division in insects. They here appear as candidates for wing patterning in *Heliconius*, but functional studies are required to better understand their role. Taken together, these results indicate that several loci associated with wing pattern variations in the supergene interval of *H. numata* are most likely functionally involved in wing patterning.

Some of the loci captured by *H. numata* inversions are associated with wing pattern variation in *H. numata* specimens with and without inversions, and in other species. This supports the hypothesis that P inversions have evolved because, via their effect of suppressing recombination, they maintain beneficial combinations of wing pattern alleles at these loci, forming good mimetic forms. The P_1 inversion is found in *H. pardalinus* [31] and recent studies have found inversions at a similar location, but with distinct breakpoints, in two other clades, respectively in *H. sara*, *H. demeter*, *H. hecalesia* and *H. telesiphe* [30], and in a very distant swallowtail butterfly, *Papilio clytia* [54]. The fact that multiple species have independently evolved inversions encompassing an orthologous region containing multiple wing pattern loci constitutes evidence that those chromosomal rearrangements have established because of their similar effect on the maintenance of coadaptation among loci in this region.

Genetic and phenotypic analyses, however, also suggest that some variations have evolved after the formation of inversions in *H. numata*. Indeed, certain phenotypes and genetic variants are only observed in a subset of the mimetic forms associated with the Hn123 rearrangement (figure 1*b*; electronic supplementary material, figure S7). For instance, forms *aurora*, *timaeus* and *tarapotensis* are all associated with the Hn123 chromosomal arrangements (inversions P_1, P_2 and P_3) but differ in several aspects of wing pattern (e.g. broad yellow band on forewings, or yellow spots in forewing tips) that we found to be genetically associated with different regions of the supergene. Because the formation of an inversion is a unique event leading to the capture of a single haplotype (i.e. inversions do not include any polymorphism upon formation; [29]), variants restricted to a subset of Hn123 forms have necessarily evolved after the formation of the inversions. These putative new wing pattern loci within the inversions may have been recruited by selection because of their tight linkage with other wing pattern loci. Following theoretical predictions [55,56], inversions could have evolved because they maintain in linkage loci initially

involved in wing patterning, and the resulting reduced recombination over a sizeable region may have subsequently favoured the recruitment of additional adaptive mutations in this region. This implies that the supergene interval contained a pre-existing set of loci with a role during wing development, some of which were previously undetected because of a lack of association with wing pattern variation in other taxa. Experimental assays are required to understand the implication of these loci in wing patterning across the *Heliconius* clade.

In summary, we found multiple loci associated with different wing pattern features in the *H. numata* supergene. Several of them are very close to genes differentially expressed in the wing discs of distinct forms and are also associated with wing pattern variation in related taxa. We found no unusual recombination pattern that could cause spurious wing pattern associations, and these associations are not expected to result from the correlation between wing pattern features and unmeasured traits. Overall, this indicates that several loci within the supergene may indeed be functionally involved in wing patterning. In agreement with theory [2,55,56], we found that the phenotypic diversity in *H. numata* is encoded by a tight cluster of loci, whose formation likely results both from the effect of chromosomal rearrangements in suppressing recombination between coadapted combinations of alleles at linked loci, and from the further recruitment of new adaptive alleles within these inversions. Our results provide empirical evidence that the preservation of coadaptation among linked loci is a key factor driving selection on chromosomal rearrangements when they form, and that supergenes allow switching between combinations of coadapted alleles, which was long predicted but has received scant empirical demonstration [13].

4. Methods

(a) Sampling and sequencing

Resequenced genome data from references [29,31,43] were used and were completed by 62 new specimens. In total, 131 specimens belonging to 16 mimetic variant of *H. numata* and from five geographical origins: Peru ($n = 85$), Ecuador ($n = 13$), Colombia ($n = 10$), French Guiana ($n = 6$) and Brazil ($n = 17$) were used (electronic supplementary material, table S1). Wings were preserved at room temperature and bodies were conserved in DMSO at −20°C.

DNA was extracted from thorax tissue using the Qiagen DNeasy blood and tissue extraction kit. Illumina Truseq paired-end libraries were prepared and sequenced in 2×100 base pair on an Illumina NovaSeq platform (Get Plage, INRA Toulouse). Reads were mapped on the *H. melpomene* (Hmel2) reference genome [57] with NextGenMap [58] with default parameters. Mapped reads were processed with GATK and SNP and indel detection was performed with the *unified genotyper*, following the procedure recommended by the authors [59]. Before filtering for the G*E association, the dataset included 46 999 947 SNPs. SnpEff (v. 4.3; [60]) was used to annotate genetic variants based on *H. melpomene* reference genome annotation.

Our reference genome Hmel2 turned out to contain an assembly error: a 45 853 bp contig considered to be on the chromosome 7 was in fact on chromosome 15 within the supergene. This explains the peak of association on chromosome 7 (figure 2). The misplacement of the scaffold on chromosome 7 was determined by mapping this scaffold with BLAST on the NCBI database and by aligning this scaffold on an improved *H. melpomene* genome published after the performance of the analyses presented here ([61]; electronic supplementary material, figure S14). Alignment was performed with mummer v. 4.0 [62] with a 1000 bp maximal gap between two adjacent matches and visualized with Circos [63].

(b) Population genetic analyses

PCAs were computed on genetic data using SNPRelate (v. 3.9; [64]) with standard option. This was used to detect whole-genome *H. numata* geographical structure (electronic supplementary material, figure S5).

To quantify the geographical structure observed with genomic PCA, we performed *F*st scans using scripts from https://github.com/simonhmartin/genomics_general, using 5000 bp windows with at least 500 SNPs per window, using all SNPs. The pairwise level of disequilibrum (R_2) was measured with PLINK v. 1.90b6.6 [45] between all biallelic SNP at the P region ($n = 762\,526$, including the three inversions and flanking regions), keeping only sites with minor allele frequency (MAF) greater than 0.2 and with $R_2 > 0.2$ to reduce computing time. Because populations from the Atlantic forest of Brazil were shown to be substantially differentiated from other *H. numata* populations based on the *F*st analyses and results from another study [43], we removed them from subsequent analyses.

Genomic PCAs were also used to assess genotypes at the P supergene based on the genetic variation segregating within the regions of P_1, P_2 and P_3 (electronic supplementary material, figures S1–S4). Indeed, several studies have shown that PCA can be used to discriminate between inversion genotypes at supergenes [29,39–41,65,66]. This was notably used in the previous study on the mimicry supergene of *H. numata* [29]. This allowed us to detect seven genotypes at the supergene: Hn0/Hn0, Hn0/Hn1, Hn0/Hn123, Hn1/Hn1, Hn1/Hn123 and Hn123/Hn123. For simplicity, we refer hereafter to specimens without inversions (Hn0/Hn0) as the Hn0 specimens, and to the specimens with the three inversions (Hn1/Hn123, Hn0/Hn123 or Hn123/Hn123) as the Hn123 specimens. Unless otherwise stated, every analysis presented hereafter was computed on these two sample sets separately.

In order to study the evolutionary history of the supergene and determine whether recombination occurs at the supergene, we computed sliding window phylogenies along the supergene, using 10 kb windows with RAxML [67]. Only specimen samples homozygous for the inversions or for their absence were used (Hn0/Hn0, Hn1/Hn1 and Hn123/Hn123). Moreover, based on PCR genotyping [29] and breeding experiment results (data not shown), we removed samples that might be heterozygous for two supergene allele belonging to the same allelic class. For instance, considering that different Hn123 haplotypes encode the morphs *aurora* and *tarapotensis*, respectively, an Hn123/Hn123 individual might have an *aurora*/*tarapotensis* genotype, which could confound phylogenetic analysis studying the evolution of these two haplotypes. Because haplotypes of the same allelic class recombine and because there are likely many neutral polymorphisms segregating in each class, phylogeny topology was highly variable along the supergene. To summarize these variations (topology weighting), we used Twisst [68] using the different morph as different taxa. We used the morphs *silvana*, *bicoloratus*, *aurora*, *timaeus*, *lyrcaeus*, *tarapotensis* and *messene* to perform such analyses since they were the only morphs with sufficient samples.

(c) Gene expression analyses

RNAseq data from [47] were reanalysed using the edgeR R package (v. 3.16.5; [48]). Gene expression in early pupal (24 h) wings discs from *silvana* individuals (Hn0/Hn0, $n = 3$) was compared to gene expression in both *tarapotensis* and *aurora*

individuals (Hn123/Hn123, $n = 7$). Gene expression in prepupae wing discs from *bicoloratus* individuals (Hn1/Hn1, $n = 3$) was compared to gene expression in both *tarapotensis* and *aurora* individuals (Hn123/Hn123, $n = 8$). The data were normalized with the calcNormFactors function. The dispersion was estimated with the estimateDisp function. Data were fitted with a quasi-likelihood negative binomial generalized log-linear model with the glmQLFTest function and *p*-values were adjusted using the Benjamini & Hochberg correction for false discovery rate.

(d) Phenotyping

Colour pattern variation was described using CPM according to the developer's recommendations [32]. Briefly, from standardized images, CPM quantifies phenotypic variation among specimens by producing comprehensive descriptors of the colour patterns, which avoids the user making assumptions about the relevance of certain descriptors. Wing photographs were colour-segmented automatically and the resulting colour partitions were attributed to one of the three colours composing the tiger patterns displayed by *H. numata* morphs (and their co-mimics): black, orange and yellow. Pixelwise comparison of wing pattern requires the pattern images to be superimposed, i.e. in a common coordinate system. Wings were therefore aligned (by translation, scaling and rotation) to an average model (improved by recursion) by maximizing the mutual information between individual pattern and the model (see [32] for more details). This procedure is described as a pattern-based alignment and results in the optimal compromise in the pixel-wise superimposition of the different pattern elements constituting the wing pattern. This has the advantage of focusing the quantification on the change in pattern elements relative to each other. Compared to landmarks based alignments, or shape-based alignment, it is therefore relatively insensitive to wing shape variation, but also to variation in the overall positioning of the pattern elements on the wing (*H. numata* males and females have a noticeable difference in the overall positioning of the pattern, due to the presence of androcony in males [32]). Wing pattern phenotypic variation could then be described as the colour variation among all pixels common to all aligned wings. Colours were encoded using the one-hot-encoding technique, as a 3 binary numbers ([1,0,0] for black, [0,1,0] for orange, [0,0,1] for yellow). The high dimensionality phenotypic space (*ca* 10^5 pixels times 3) was summarized by PCAs.

In order to isolate variants associated with specific features of the wing pattern, we also performed a description of the phenotype limited to different parts/features of the wings. Besides the global analysis of fore- and hindwings together (i), we analysed separately the following partitions: (ii) forewings, (iii) hindwings, (iv) yellow patterns of the forewings, (v) base of the forewings, (vi) median area of the forewings (yellow band area) and (vii) apex of the forewings. Phenotyping of specific wing regions (i.e. partitions) was performed by feeding the PCA with pixel values belonging only to the region of interest. Colour-specific phenotyping was performed by hiding the variations other than the one of the colour of interest. This corresponds to setting the value of all other colours to 0 in the one-hot-encoding procedure. To take into account the genetic structure of the supergene P, which implies the absence of recombination between specimens harbouring different rearrangements, we performed PCAs on a subset of specimens, based on their genotype (presence/absence of the three inversions). Therefore, we computed PCAs only on specimens homozygous or heterozygous for chromosome form Hn123 and on specimens homozygous for chromosome form Hn0, respectively (electronic supplementary material, figure S8). To compute the effect of associated genetic variants on the colour variation at each wing position (image pixel), we translated the loads (contribution to the multivariate association result) attributed by MV-PLINKs to each phenotypic trait (i.e. to each phenotypic principal component) into pixel values using the eigenvalues of the phenotypic PCA. See MV-PLINK and CPM references for further details [26,32,44].

(e) Multivariate association studies

To determine the genetic basis of colour pattern variation, multivariate genome-wide association studies were performed using MV-PLINK (v. 1.6; [44]). MV-PLINK performs a canonical correlation analysis to test for an association between variation at multiple phenotypes at once and a single genetic variant. The description of variation in width, size and translation of wing pattern elements increases in accuracy with the number of principal components considered [69]. Nevertheless, including many non-informative variables (e.g. including many principal components explaining a low fraction of the phenotypic variance) in multivariate association causes non-informative association results, which hampers isolation of meaningful associations. Thus, we calculated G*E associations using two to six principal components as phenotypes. To take into account the supergene structure, GWAS were carried out independently within each genotypic group Hn0 (Hn0/Hn0, $n = 22$)) and Hn123 (Hn123/Hn0, Hn123/Hn1 or Hn123/Hn123, $n = 61$). Only variants with MAF greater than 0.02 and genotyping rate greater than 0.5 were conserved for analyses, resulting in 532 574 SNPs used for Hn0 association and 306 921 SNPs for Hn123 association at P region (Chromosome 15 : 500 000–4 000 000, including the three inversions and flanking regions). Bcftools was used to process the vcf files [70]. For each multivariate association (using 2, 3, 4, 5 or 6 phenotypic principal components), for each variant in the supergene region and flanking regions, 1×10^6 adaptive permutations were conducted (see PLINK manual, [45]). Only variants with no permutation resulting in a higher statistical *p*-value were considered to be significantly associated (i.e. with empirical *p*-value $< 10^{-6}$).

To distinguish regions that present an enrichment in significantly associated variants (figure 2*b*,*e*), we computed the density of significantly associated variants in overlapping sliding windows (10 000 bp, with 100 bp slide), considering for these analyses all variants that have been inferred as significantly associated in association analyses with 2, 3, … ,or 6 phenotypic principal components (variants that were found significantly associated in more than one association analysis, e.g. in analyses with 2, 3 and 4 phenotypic components, as frequently observed, were considered only once).

We defined a region of association as a 10 kb region displaying a clear enrichment in significantly associated variant compared to nearby regions. The choice of region size is arbitrary and a finer or larger region size can alter the results. We acknowledge that many regions present many significantly associated variants but not forming a clear peak. This likely results from the tight physical linkage of loci associated with wing pattern variation, especially around *cortex*. Hence, the list of associated regions (figure 3*g*,*h*; electronic supplementary material, table S2) is not intended to be exhaustive but reflects the regions that we think are of interest, because of clear peaks of association or because of their strong association with particular wing pattern features.

(f) Univariate association studies with covariates

To verify that the wing pattern-associated regions we found with multivariate analyses were not false positives due to a geographic structure in our dataset, we repeated our analyses but using a single phenotype and multiple geographical covariates with PLINK v. 1.90b6.24 [45]. We used the first 10 principal components of whole-genome PCA (see §4b and electronic

supplementary material, figure S5) as covariates in these analyses (computed only with Hn123 or Hn0 specimens). To analyse wing pattern variation in as much detail as possible with univariate analyses, in addition to the associations performed using as phenotype the principal components summarizing the variation of hind and forewing together (as done in multivariate analyses), we also performed associations using as phenotype the principal components of PCA focusing on more specific aspects of wing variation: variation found on hindwings only, variation found on forewings, variation in yellow patterning, variation in the forewing tip, variation in the forewing middle part and variation in the forewing base (see electronic supplementary material, figure S9). For each of these specific aspects of wing variation, we computed association using the first three principal components. Electronic supplementary material, figures S16 and S18 show the results of these associations (only the results for the first principal component are shown).

To account for any potential dominance effect of variant on wing pattern, we also performed analyses using two variables representing an additive effect and a dominance deviation (dominance-related covariates) in addition to the whole-genome covariate (using the '–genotypic' options in PLINK). Electronic supplementary material, figures S17 and S19 show the results of these associations (only the results for the first principal components are shown). To distinguish regions that present an enrichment in significantly associated variants (figure 2*c*,*f*), we computed the density of significantly associated variants in overlapping sliding window (10 000 bp, with 100 bp slide), considering for these analyses all variants that have been inferred as significantly associated in association analyses with one or more aspect of wing pattern variation (variants that were found significantly associated in more than one association analysis, e.g. in analyses with the second component describing yellow pattern variation and with the first components describing forewing variations, as frequently observed, were considered only once).

(g) Individual-based simulations

We used individual-based simulations to estimate our ability to identify the loci responsible for colour variation in a supergene. In order to maintain a tractable model, we used in simulations a simplified scenario mimicking *H. numata* evolution. We used SLiM v. 3.2 [71] to simulate during 120 000 generations the evolution of two panmictic populations, denoted 1 and 2, each of $N = 5000$ individuals in a Wright–Fisher model. The phenotype of the individuals (e.g. their wing pattern) was determined by three epistatic loci (one of these being an inversion). The two populations experienced disruptive selection: phenotypes beneficial in a population were selected against by natural selection in the other. The two populations were connected by a migration rate of $m = 0.1$. This relatively high migration rate between two populations was used to simulate the coexistence of alternative mimetic forms within the same population in *H. numata*. We simulated individuals with a single pair of 2 Mb chromosomes on which mutations occur at a rate u, with u ranging from 10^{-7} to 10^{-9} per bp and that were recombining at a rate $r = 1 \times 10^{-6}$. Each occurring mutation had its selection coefficient (s) drawn from a gamma distribution with a shape of 0.2 and a mean –0.03, and its dominance coefficient h randomly sampled among 0, 0.001, 0.01, 0.1, 0.25 and 0.5 with uniform probabilities. In nature, the effective population size of *H. numata* has been estimated to be Ne = 23 089 618 [43], with a mutation rate of *ca* 2.9×10^{-9} [72] and a recombination rate of *ca* 0.6×10^{-5} [73]. Considering the much smaller population size that we simulated here, we therefore expected to observe stronger linkage disequilibrium along the genome, and in particular within the supergene, in simulated populations than in real populations.

For each simulation, a burn-in period of 15 000 generations was run to allow the population to reach equilibrium for the number of segregating mutations. After this burn-in period, we introduced in population 1 a single mutation mimicking a 1 Mb inversion suppressing recombination over the region 500 000–1 500 000 bp. This inversion-mimicking, recombination modifier mutation was introduced on a single, randomly selected chromosome and, when heterozygous, it suppressed recombination across the region in which it resided (i.e. as a *cis*-recombination modifier). At generation 20 000, we introduced in the population 1 a single mutation at position 1 200 000 bp (denoted hereafter the colour locus 1) in a single chromosome harbouring the inversion (therefore within the inversion). At generation 25 000, we introduced in the population 2 a single mutation at position 1 400 000 (denoted hereafter colour locus 2) in a single chromosome harbouring the inversion (therefore also within the inversion). Therefore, all individuals with one or two derived colour mutations also harboured the inversion.

Throughout the simulations, the fitness of individuals depended on the deleterious mutations they carried and their genotype at the inversion and colour loci. We considered that the inversion and the derived colour alleles were dominant over the ancestral alleles (being the ancestral gene order in the case of the inversion). Individuals homozygous for the inversions had their fitness multiplied by a factor *Hdis*, with *Hdis* being 0.1, 0.3 or 0.5 (i.e. inversion homozygotes suffering from a disadvantage relative to inversion heterozygotes). In natural populations, individuals homozygous for the Hn123 arrangement have a larval survival approximately 0.4 lower than heterozygous individuals (Hn1/Hn123 or Hn0/Hn123; [29]). Individuals without the inversion and without any derived colour mutation (ancestral state) had their fitness multiplied by a factor 0.5. These individuals for instance represent the form *silvana*, without inversion and poorly protected against predators in Peru [34]. Individuals with the inversion (homozygous or heterozygous) and without any derived colour mutation (i.e with the ancestral allele) have their fitness multiplied by a factor 0.8. Individuals with the inversion and the two derived colour mutations (being homozygous or heterozygous) had their fitness multiplied by a factor 0.3 (these individuals are considered 'recombinant' and suffer from a maladapted wing pattern). Individuals with only the derived allele at the colour locus 1 had their fitness multiplied by 1 in population 1, and by 0.7 in population 2. Individuals with only the derived allele at the colour locus 2 had their fitness multiplied by 0.7 in population 1, and by 1 in population 2. Electronic supplementary material, table S3 summarizes these fitness parameters.

In sum, in our simulations, the inversion is a beneficial variant in both populations, but is also associated with a homozygous disadvantage maintaining it at intermediate frequency (approx. 10–50% in the parameter used). This mimics the situation encountered in *H. numata* in Peru [29]. The two colour loci are epistatic (their fitness effect depends on the allele at the other locus) and are under disruptive selection. Each allele is favoured in one population and selected against in the other. Individuals with the two derived mutations are selected against. Different combinations of alleles are maintained in populations, and the balance of these combinations of alleles is due to a mixture of migration-selection balance and inversion homozygous disadvantage. Several combinations are associated with the same gene order and can therefore recombine. However, such recombination is detrimental as it results in haplotypes with a reduced fitness (as observed in *H. numata*). In sum, this mimics the evolution of a supergene similar to the P supergene of *H. numata*, with the difference that in *H. numata*, the alternative mimetic forms segregate in the same population whereas here we consider for computing purposes two populations connected by high gene flow.

At simulation end (generation 120 000), we randomly selected 60 specimens that harboured either only the derived colour allele 1 (homozygous or heterozygous) or only the derived colour allele 2 (homozygous or heterozygous), regardless of their population of origin. This mimics the sampling of individuals with a mimetic form known to be associated with the Hn123 arrangements. For each of these individuals, we attributed a single quantitative phenotype value based on their genotype at the colour loci. Individuals with colour allele 1 were determined to have a phenotype of $1-\gamma$, with γ being sampled between 0 and *PheVar* with uniform probabilities, *PheVar* being 0.2, 0.4 or 1.0. Individuals with colour allele 2 were determined to have a phenotype of $-1+\gamma$, with γ being sampled between 0 and *PheVar* with uniform probabilities, *PheVar* being 0.2, 0.4 or 1.0. The *PheVar* variable therefore represents the fraction of phenotypic variation that is explained by factors other than the genotype at the colour locus. This was used to simulate variations in phenotype measurements due to the environment, for example.

Finally, we used PLINK v. 1.90b6.24 [45] to perform association studies using this simulated genomes and phenotypes using the same parameters and approach as used for the analyses of *H. numata* data. We display in the electronic supplementary material, figure S22 some examples of these associations. It shows that even using parameters that are prone to generate high LD at the inversion locus and poor associations (e.g. relatively low migration, low heritability of the wing pattern), we can still detect the independent wing pattern loci with this approach. Only extreme parameter values, notably a high level of homozygous disadvantage hampering the formation of homozygotes, prevented the detection of the wing pattern loci. In those cases, we observed an single large peak of association at the inversion locus, and not several independent peaks.

Ethics. The sampling of the butterflies used in this study and the associated researchers were allowed by the Peruvian government under the permits 236-2012-AG-DGFFS-DGEFFS, 201-2013-MINAGRIDGFFS/DGEFFS and 002-2015-SERFOR-DGGSPFFS.

Data accessibility. Code used to produce analyses have been deposited on GitHub: https://github.com/PaulYannJay/HeliconiusGWAS. The data used to generate the figures are available on figshare [74]. The raw sequence data were deposited in NCBI SRA and accession numbers are indicated in electronic supplementary material, table S1. The whole-genome VCF file is available upon request.

Supplementary figures and tables are provided in the electronic supplementary material [75].

Authors' contributions. P.J.: conceptualization, data curation, methodology, supervision, visualization, writing—original draft and writing—review and editing; M.L.: conceptualization, data curation and investigation; Y.L.P.: methodology, software and writing—review and editing; A.W.: methodology and writing—review and editing; M.A.: resources and writing—review and editing; M.C.: resources and writing—review and editing; M.J.: conceptualization, funding acquisition, supervision and writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This research was supported by Agence Nationale de la Recherche (ANR) grant nos. ANR-12-JSV7-0005 and ANR-18-CE02-0019-01 and European Research Council grant no. ERC-StG-243179 to M.J.

Acknowledgements. We are very grateful to Oscar Puebla, Floriane Coulmance and the Puebla lab for their careful review and useful comments on this manuscript. We thank Camilo Salazar for providing specimens from Colombia, Chris Jiggins and his team for specimens from Ecuador, and André V. L. Freitas and Bárbara Huber for specimens from Brazil. We thank the Peruvian government for providing the necessary research permits (236-2012-AG-DGFFS-DGEFFS, 201-2013-MINAGRI-DGFFS/DGEFFS and 002-2015-SERFOR-DGGSPFFS). This project benefited from the Montpellier Bioinformatics Biodiversity platform supported by the LabEx CeMEB, ANR 'Investissements d'avenir' programme ANR-10-LABX-04-01.

References

1. Lenormand T, Otto SP. 2000 The evolution of recombination in a heterogeneous environment. *Genetics* **156**, 423–438. (doi:10.1093/genetics/156.1.423)
2. Yeaman S. 2013 Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proc. Natl Acad. Sci. USA* **110**, E1743–E1751. (doi:10.1073/pnas.1219381110)
3. Lamichhaney S *et al.* 2016 Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nat. Genet.* **48**, 84–88. (doi:10.1038/ng.3430)
4. Tuttle EM *et al.* 2016 Divergence and functional degradation of a sex chromosome-like supergene. *Curr. Biol.* **26**, 344–350. (doi:10.1016/j.cub.2015.11.069)
5. Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kronforst MR. 2014 *Doublesex* is a mimicry supergene. *Nature* **507**, 229–232. (doi:10.1038/nature13112)
6. Wang J, Wurm Y, Nipitwattanaphon M, Riba-Grognuz O, Huang YC, Shoemaker D, Keller L. 2013 A Y-like social chromosome causes alternative colony organization in fire ants. *Nature* **493**, 664–668. (doi:10.1038/nature11832)
7. Kess T *et al.* 2019 A migration-associated supergene reveals loss of biocomplexity in Atlantic cod. *Sci. Adv.* **5**, eaav2461. (doi:10.1126/sciadv.aav2461)
8. Stefansson H *et al.* 2005 A common inversion under selection in Europeans. *Nat. Genet.* **37**, 129–137. (doi:10.1038/ng1508)
9. Castric V, Bechsgaard JS, Grenier S, Noureddine R, Schierup MH, Vekemans X. 2010 Molecular evolution within and between self-incompatibility specificities. *Mol. Biol. Evol.* **27**, 11–20. (doi:10.1093/molbev/msp224)
10. Huu CN, Keller B, Conti E, Kappel C, Lenhard M. 2020 Supergene evolution via stepwise duplications and neofunctionalization of a floral-organ identity gene. *Proc. Natl Acad. Sci. USA* **117**, 23 148–23 157. (doi:10.1073/pnas.2006296117)
11. Kappel C, Huu CN, Lenhard M. 2017 A short story gets longer: recent insights into the molecular basis of heterostyly. *J. Exp. Bot.* **68**, 5719–5730. (doi:10.1093/jxb/erx387)
12. Schwander T, Libbrecht R, Keller L. 2014 Supergenes and complex phenotypes. *Curr. Biol.* **24**, R288–R294. (doi:10.1016/j.cub.2014.01.056)
13. Villoutreix R, Ayala D, Joron M, Gompert Z, Feder JL, Nosil P. 2021 Inversion breakpoints and the evolution of supergenes. *Mol. Ecol.* **30**, 2738–2755. (doi:10.1111/mec.15907)
14. Wellenreuther M, Bernatchez L. 2018 Eco-evolutionary genomics of chromosomal inversions. *Trends Ecol. Evol.* **33**, 427–440. (doi:10.1016/j.tree.2018.04.002)
15. Mather K. 1950 The genetical architecture of heterostyly in *Primula sinensis*. *Evolution* **4**, 340–352. (doi:10.1111/j.1558-5646.1950.tb01404.x)
16. Yang Q, Zhang D, Li Q, Cheng Z, Xue Y. 2007 Heterochromatic and genetic features are consistent with recombination suppression of the self-incompatibility locus in *Antirrhinum*. *Plant J.* **51**, 140–151. (doi:10.1111/j.1365-313X.2007.03127.x)
17. Kao T, Tsukamoto T. 2004 The molecular and genetic bases of S-RNase-based self-incompatibility. *Plant Cell* **16**, S72–S83. (doi:10.1105/tpc.016154)
18. Chookajorn T, Kachroo A, Ripoll DR, Clark AG, Nasrallah JB. 2004 Specificity determinants and diversification of the *Brassica* self-incompatibility pollen ligand. *Proc. Natl Acad. Sci. USA* **101**, 911–917. (doi:10.1073/pnas.2637116100)
19. Branco S *et al.* 2017 Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. *Proc. Natl Acad. Sci. USA* **114**, 7067–7072. (doi:10.1073/pnas.1701658114)
20. Hartmann FE, Duhamel M, Carpentier F, Hood ME, Foulongne-Oriol M, Silar P, Malagnac F, Grognet P,

Giraud T. 2021 Recombination suppression and evolutionary strata around mating-type loci in fungi: documenting patterns and understanding evolutionary and mechanistic causes. *New Phytologist* **229**, 2470–2491. (doi:10.1111/nph.17039)

21. Nadeau NJ *et al.* 2016 The gene *cortex* controls mimicry and crypsis in butterflies and moths. *Nature* **534**, 106–110. (doi:10.1038/nature17961)
22. Reed RD *et al.* 2011 *optix* drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* **333**, 1137–1141. (doi:10.1126/science.1208227)
23. Westerman EL *et al.* 2018 *Aristaless* controls butterfly wing color variation used in mimicry and mate choice. *Curr. Biol.* **28**, 3469–3474.e4. (doi:10.1016/j.cub.2018.08.051)
24. Merrill RM *et al.* 2015 The diversification of *Heliconius* butterflies: what have we learned in 150 years? *J. Evol. Biol.* **28**, 1417–1438. (doi:10.1111/jeb.12672)
25. Martin A *et al.* 2012 Diversification of complex butterfly wing patterns by repeated regulatory evolution of a *Wnt* ligand. *Proc. Natl Acad. Sci. USA* **109**, 12 632–12 637. (doi:10.1073/pnas.1204800109)
26. Huber B *et al.* 2015 Conservatism and novelty in the genetic architecture of adaptation in *Heliconius* butterflies. *Heredity* **114**, 515–524. (doi:10.1038/hdy.2015.22)
27. Meier JI *et al.* 2021 Haplotype tagging reveals parallel formation of hybrid races in two butterfly species. *Proc. Natl Acad. Sci. USA* **118**, e2015005118. (doi:10.1073/pnas.2015005118)
28. Kirkpatrick M, Barton N. 2006 Chromosome inversions, local adaptation and speciation. *Genetics* **173**, 419–434. (doi:10.1534/genetics.105.047985)
29. Jay P, Chouteau M, Whibley A, Bastide H, Parrinello H, Llaurens V, Joron M. 2021 Mutation load at a mimicry supergene sheds new light on the evolution of inversion polymorphisms. *Nat. Genet.* **53**, 288–293. (doi:10.1038/s41588-020-00771-1)
30. Edelman NB *et al.* 2019 Genomic architecture and introgression shape a butterfly radiation. *Science* **366**, 594–599. (doi:10.1126/science.aaw2090)
31. Jay P, Whibley A, Frézal L, de Cara MÁ, Nowell RW, Mallet J, Dasmahapatra KK, Joron M. 2018 Supergene evolution triggered by the introgression of a chromosomal inversion. *Curr. Biol.* **28**, 1839–1845.e3. (doi:10.1016/j.cub.2018.04.072)
32. Le Poul Y, Whibley A, Chouteau M, Prunier F, Llaurens V, Joron M. 2014 Evolution of dominance mechanisms at a butterfly mimicry supergene. *Nat. Commun.* **5**, 5644. (doi:10.1038/ncomms6644)
33. Brown KS, Benson WW. 1974 Adaptive polymorphism associated with multiple Müllerian mimicry in *Heliconius numata* (Lepid. Nymph.). *Biotropica* **6**, 205–228. (doi:10.2307/2989666)
34. Chouteau M, Arias M, Joron M. 2016 Warning signals are under positive frequency-dependent selection in nature. *Proc. Natl Acad. Sci. USA* **113**, 2164–2169. (doi:10.1073/pnas.1519216113)
35. Arias M, Davey JW, Martin S, Jiggins C, Nadeau N, Joron M, Llaurens V. 2020 How do predators generalize warning signals in simple and complex prey communities? Insights from a videogame. *Proc. R. Soc. B* **287**, 20200014. (doi:10.1098/rspb.2020.0014)
36. Joron M *et al.* 2011 Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **477**, 203–206. (doi:10.1038/nature10341)
37. Van Belleghem SM *et al.* 2017 Complex modular architecture around a simple toolkit of wing pattern genes. *Nat. Ecol. Evol.* **1**, 1–12. (doi:10.1038/s41559-016-0052)
38. Joron M *et al.* 2006 A conserved supergene locus controls colour pattern diversity in *Heliconius* butterflies. *PLoS Biol.* **4**, e303. (doi:10.1371/journal.pbio.0040303)
39. Ma J, Amos CI. 2012 Investigation of inversion polymorphisms in the human genome using principal components analysis. *PLoS ONE* **7**, e40224. (doi:10.1371/journal.pone.0040224)
40. Lindtke D, Lucek K, Soria-Carrasco V, Villoutreix R, Farkas TE, Riesch R, Dennis SR, Gompert Z, Nosil P. 2017 Long-term balancing selection on chromosomal variants associated with crypsis in a stick insect. *Mol. Ecol.* **26**, 6189–6205. (doi:10.1111/mec.14280)
41. Faria R *et al.* 2019 Multiple chromosomal rearrangements in a hybrid zone between *Littorina saxatilis* ecotypes. *Mol. Ecol.* **28**, 1375–1393. (doi:10.1111/mec.14972)
42. Chouteau M, Llaurens V, Piron-Prunier F, Joron M. 2017 Polymorphism at a mimicry supergene maintained by opposing frequency-dependent selection pressures. *Proc. Natl Acad. Sci. USA* **114**, 8325–8329. (doi:10.1073/pnas.1702482114)
43. de Cara MÁR *et al.* 2021 Supergene formation is associated with a major shift in genome-wide patterns of diversity in a butterfly. *bioRxiv* 2021.09.29.462348. (doi:10.1101/2021.09.29.462348)
44. Ferreira MAR, Purcell SM. 2009 A multivariate test of association. *Bioinformatics* **25**, 132–133. (doi:10.1093/bioinformatics/btn563)
45. Purcell S *et al.* 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575. (doi:10.1086/519795)
46. Hellwege J, Keaton J, Giri A, Gao X, Velez Edwards DR, Edwards TL. 2017 Population stratification in genetic association studies. *Curr. Protocols Hum. Genet.* **95**, 1.22.1–1.22.23. (doi:10.1002/cphg.48)
47. Saenko SV, Chouteau M, Piron-Prunier F, Blugeon C, Joron M, Llaurens V. 2019 Unravelling the genes forming the wing pattern supergene in the polymorphic butterfly *Heliconius numata*. *Evodevo* **10**, 16. (doi:10.1186/s13227-019-0129-2)
48. Robinson MD, McCarthy DJ, Smyth GK. 2010 edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140. (doi:10.1093/bioinformatics/btp616)
49. Clarke CA, Sheppard PM. 1960 Super-genes and mimicry. *Heredity* **14**, 175–185. (doi:10.1038/hdy.1960.15)
50. Timmermans MJTN *et al.* 2014 Comparative genomics of the mimicry switch in *Papilio dardanus*. *Proc. R. Soc. B* **281**, 20140465. (doi:10.1098/rspb.2014.0465)
51. Joron M. 2005 Polymorphic mimicry, microhabitat use, and sex-specific behaviour. *J. Evol. Biol.* **18**, 547–556. (doi:10.1111/j.1420-9101.2005.00880.x)
52. Maisonneuve L, Chouteau M, Joron M, Llaurens V. 2021 Evolution and genetic architecture of disassortative mating at a locus under heterozygote advantage. *Evolution* **75**, 149–165. (doi:10.1111/evo.14129)
53. Livraghi L *et al.* 2021 Cortex *cis*-regulatory switches establish scale colour identity and pattern diversity in *Heliconius*. *eLife* **10**, e68549. (doi:10.7554/eLife.68549)
54. VanKuren NW, Massardo D, Nallu S, Kronforst MR. 2019 Butterfly mimicry polymorphisms highlight phylogenetic limits of gene reuse in the evolution of diverse adaptations. *Mol. Biol. Evol.* **36**, 2842–2853. (doi:10.1093/molbev/msz194)
55. Yeaman S, Aeschbacher S, Bürger R. 2016 The evolution of genomic islands by increased establishment probability of linked alleles. *Mol. Ecol.* **25**, 2542–2558. (doi:10.1111/mec.13611)
56. Bürger R, Akerman A. 2011 The effects of linkage and gene flow on local adaptation: a two-locus continent–island model. *Theor. Popul. Biol.* **80**, 272–288. (doi:10.1016/j.tpb.2011.07.002)
57. Davey JW *et al.* 2016 Major improvements to the *Heliconius melpomene* genome assembly used to confirm 10 chromosome fusion events in 6 million years of butterfly evolution. *G3* **6**, 695–708. (doi:10.1534/g3.115.023655)
58. Sedlazeck FJ, Rescheneder P, von Haeseler A. 2013 NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics* **29**, 2790–2791. (doi:10.1093/bioinformatics/btt468)
59. DePristo MA *et al.* 2011 A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498. (doi:10.1038/ng.806)
60. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster strain w1118; iso-2; iso-3*. *Fly (Austin)* **6**, 80–92. (doi:10.4161/fly.19695)
61. Davey JW, Barker SL, Rastas PM, Pinharanda A, Martin SH, Durbin R, McMillan WO, Merrill RM, Jiggins CD. 2017 No evidence for maintenance of a sympatric *Heliconius* species barrier by chromosomal inversions. *Evolution Letters.* **1**, 138–154. (doi:10.1002/evl3.12)
62. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004 Versatile and open software for comparing large genomes. *Genome Biol.* **5**, R12. (doi:10.1186/gb-2004-5-2-r12)
63. Krzywinski M, Schein J, Birol İ, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009 Circos: an information aesthetic for comparative genomics.

Genome Res. **19**, 1639–1645. (doi:10.1101/gr.092759.109)

64. Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. 2012 A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328. (doi:10.1093/bioinformatics/bts606)
65. Todesco M *et al.* 2020 Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature* **584**, 602–607. (doi:10.1038/s41586-020-2467-6)
66. Mérot C, Oomen RA, Tigano A, Wellenreuther M. 2020 A roadmap for understanding the evolutionary significance of structural genomic variation. *Trends Ecol. Evol.* **35**, 561–572. (doi:10.1016/j.tree.2020.03.002)
67. Stamatakis A. 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313. (doi:10.1093/bioinformatics/btu033)
68. Martin SH, Belleghem SMV. 2017 Exploring evolutionary relationships across the genome using topology weighting. *Genetics* **206**, 429–438. (doi:10.1534/genetics.116.194720)
69. Le Poul Y. 2014 Selection for mimicry in butterflies: quantitative approaches on colour pattern resemblance and diversity. Doctoral thesis, Muséum National d'Histoire Naturelle, Paris, France.
70. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079. (doi:10.1093/bioinformatics/btp352)
71. Haller BC, Messer PW. 2019 SLiM 3: forward genetic simulations beyond the Wright–Fisher model. *Mol. Biol. Evol.* **36**, 632–637. (doi:10.1093/molbev/msy228)
72. Keightley PD, Pinharanda A, Ness RW, Simpson F, Dasmahapatra KK, Mallet J, Davey JW, Jiggins CD. 2015 Estimation of the spontaneous mutation rate in *Heliconius melpomene*. *Mol. Biol. Evol.* **32**, 239–243. (doi:10.1093/molbev/msu302)
73. Wilfert L, Gadau J, Schmid-Hempel P. 2007 Variation in genomic recombination rates among animal taxa and the case of social insects. *Heredity* **98**, 189–197. (doi:10.1038/sj.hdy.6800950)
74. Jay P. 2022 Datasets for Association mapping of colour variation in a butterfly provides evidence that a supergene locks together a cluster of adaptive loci. FigShare. (doi:10.6084/m9.figshare.19706320.v1)
75. Jay P, Leroy M, Le Poul Y, Whibley A, Arias M, Chouteau M, Joron M. 2022 Association mapping of colour variation in a butterfly provides evidence that a supergene locks together a cluster of adaptive loci. FigShare. (doi:10.6084/m9.figshare.c.5983511)