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

Paul Jay, Manon Leroy, Yann Le Poul, Annabel Whibley, Mónica Arias, Mathieu Chouteau, Mathieu Joron

Association mapping of colour variations in a butterfly provides evidences that a supergene locks together a cluster of adaptive loci .

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

[1]CEFE, Université de Montpellier, CNRS, EPHE, IRD, Montpellier, France
[2]Faculty of Biology, LMU Munich, Planegg-Martinsried, Germany
[3]School of Biological Sciences, University of Auckland, Auckland, New Zealand
[4]LEEISA, USR 63456, Université de Guyane, CNRS, IFREMER, 275 route de Montabo, 797334 Cayenne, French Guiana

Keywords : Inversion | GWAS | Multivariate association | Wing colour pattern | Cluster of adaptive loci | Divergence hitchhiking

Correspondance: paul.yann.jay@gmail.com, mathieu.joron@cefe.cnrs.fr

Summary : 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 set 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 an 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 supergene 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.

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 homogenising 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, color 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 the complexity of untangling the genetic architecture of complex multidimensional 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 favor 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, *Heliconius 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 1b). Variation in wing patterning in *H. numata* involve 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 1b and (*32*)]. 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 (*33*, *34*), 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 on Figure 1a; (*29*, *35*)]. 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 3Mb 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 1a): Hn0, the standard arrangement without any inversion; Hn1, the first derived arrangement carrying the P1 inversion only; and Hn123, the second derived arrangement with the three adjacent inversions P1, P2 and P3. 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 than homozygote for the same chromosomal arrangements [Fig. 1A; (*32*, *36*)]. 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 (*35*, *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*, *35*), 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 multidimensional wing-pattern variation and performed genome wide association studies to associate phenotypic and genetic variations. We show that multiple genomic intervals are associated with colour variations in *H. numata* and that all these regions are situated within the P supergene, bringing evidence 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 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.

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, (*32*)]. Reads were mapped against the *H. melpomene* reference genome (Hmel 2).

Following previous studies (*29*, *39–41*), we used Principal component analyses (PCA) to identify individual genotypes at the supergene (Figure S1-4, 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, 11 Hn1/Hn0) and 72 homozygous or heterozygous for the second derived chromosomal arrangement (27 Hn123/Hn123, 37 Hn0/Hn123, 8 Hn1/Hn123) . Occurrence of individuals homozygous for each chromosomal rearrangement in substantial proportion in natural population (*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 three chromosomal arrangements (Hn0/Hn0, Hn1/Hn1 and Hn123/Hn123) was only slightly higher within the supergene than in flanking regions (Figure S6) . Furthermore, phylogenetic topologies shift repeatedly along the inversions (Figure S7), consistent with ongoing recombination within each class of arrangement. By contrast, the linkage disequilibrium within the supergene was much stronger than in colinear regions when we considered groups of individuals with alternative chromosomal arrangements, indicating recombination suppression between segments in different orientations (Figure S6).

Morphometric analyses were run on 109 samples whose wings were in good condition, using Color Pattern Modelling (CPM), an algorithm for the quantification of colour pattern variations based on colour classification and colour pattern registration (*36*). By providing naive descriptors of colour variation, CPM overcomes the limit and bias of phenotype description by a 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. Principal component analyses (PCAs) were used to reduce the high dimensionality of the colour variation and showed that wing pattern polymorphism involves a mixture of qualitative and quantitative variation (Figs 1C and S8-9). Wing pattern polymorphism in *H. numata* indeed involves variations on different parts of the wing, with some features appearing more discrete (e.g. presence of a broad yellow band) than others (e.g. spread of hindwing black patterns; Figure 1 and S8-9). The first phenotypic principal components were observed to describe obvious colour variations. For instance, the first principal component appear to quantify the overall amount of black patterns and the third principal component to describe the size of the broad yellow band on the fore wing or its absence (Figure 1c and S8-9). As expected, samples

harbouring the same chromosomal arrangement were clustered in the phenotype space, highlighting that supergene inversions are major determinants of specimen phenotype (Fig. 1C and S8-9). 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 (Fig. S10).

In accordance with a previous study (*43*), principal component analyses and differentiation analyses (Fst) showed a near absence of genetic structure in *H. numata* across the Neotropics [Figure S6 and S12-13 ; see also (*43*)]. For instance, the genomic differentiation (Fst) between populations from French Guiana and from Peru was only of 0.02 despite being separated by approximately 2900 km (Figure S13). We found however a relatively strong differentiation (Fst=0.27) between the population from the Brazilian Atlantic forest and the populations of the rest of the range [Fig. 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 samples 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*, *35*), the main region of association corresponded to the P supergene (Fig. 2A). In the analysis performed with all samples 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 (Fig. 2B). In order to disentangle the effect of the loci within the supergene and to remove the confounding effect of inversions, we performed phenotype-genotype associations separately on the specimens homozygous for the chromosome arrangement Hn0 (no inversion, n=22, referred hereafter as Hn0 samples) and on those homozygous or heterozygous for chromosome arrangement Hn123 (containing inversions P1, P2 and P3, n=61, hereafter referred to as Hn123 samples). The dominance of inverted alleles to non-inverted alleles (*36*) allows the retention 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 Methods and 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 1c and S8).

Performing 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 (Fig. 3A and S7). We first computed multivariate associations using two to six phenotypic principal components (Figure 3a and 3d and S14-15) 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 region 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 only variants with empirical p-value < $1e^{-6}$ (i.e. variants for which none permutation resulted in a better association). We found that several regions within P1, P2 and P3 were enriched in variants associated with wing-pattern variations, both in the Hn0 and Hn123 sample sets (Figure 3; see methods). Several regions were associated with phenotype changes in both sample sets (Fig. 7 and S7 and Table S2). For instance, the same intron in the gene *parn* was associated with colour variation in Hn0 and Hn123 samples. 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 (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 ten principal components describing the whole-genome structure [for population stratification effects; see Methods and Figure S5 ; (*46*)]. Because univariate analyses can only study 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 to 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 more specific part 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 (Figure S9, see Methods). 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 (Figure S16-19 and 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 (Figure S14-19). Computing the density of associated variants in these univariate analyses, we found the same regions enriched in wing-pattern associated variants than 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 associated genetic variants (SNPs or indels) on colour variation at each wing position (image pixel). This can be visualised in the form of heatmaps of variant effects on the wings, where 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 visualised for each colour independently but can also be summarized into an overall, non-directional effect per pixel (Fig. 3g-h). The limited phenotypic variation among Hn0 individuals did not enable proper dissection of the different wing pattern features in those forms (Figure 3h) and therefore all associated variants were found to affect similarly the wing pattern. By contrast, in specimens with the three inversions P1, P2 and P3, we found that associated variants in the different regions of the supergene affected different features of the wing phenotype (figure 3g). For instance, we found variants respectively associated with 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 hindwings (Fig. 3-4, 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 (Fig. 3-4, Table S2). In contrast, several variants within the supergene were associated with similar variations of the wing pattern. The co-association of distant genomic region with similar phenotypic variation in Hn123 samples could be caused by the correlation of multiple phenotype features in our dataset -i.e. the non-uniform phenotypic variation among Hn123 samples (Figure 1c)- leading the genetic variants controlling these independent features to apparently have the same effect (as observed in a much stronger proportion in Hn0 samples). 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 present 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-disks between samples 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. (Fig. 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, Table S2). Their differential expression during wing development strongly suggests that they participate in the control of wing pattern variation in *H. numata*.

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 multidimensional character such as wing patterns. However it requires the use of multivariate association methods to identify the genetic variants involved in these multidimensional 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 populations stratifications. 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 model » 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 colouration. 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 association. For

instance, if a locus controlling a given wing pattern feature is in linkage with a locus inducing mate rejection based on such 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 stresses 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 role of *cortex* and *wash* have been experimentally validated. Our result shows 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 an important role 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 variations in *H. numata* specimens with and without inversions, and in other species. This strongly 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 1B and 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 be indeed 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 the switch between combinations of coadapted alleles, which was long predicted but have received very few empirical demonstration (*13*).

Methods

Sampling and sequencing

Resequenced genome data from references (*29, 31*) 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), Brazil

(n=17) were used (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 2x100 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 (v4.3; ((*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 Heliconius melpomene genome published after the performance of the analyses presented here ((*61*); Figure S14). Aligment was perform with mummer V4.0 (*62*) with a 1000bp maximal gap between two adjacent matches and visualised with Circos (*63*)

Population genetic analyses

Principal component analyses 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 (Figure S5).

To quantify the geographical structure observed with genomic PCA, we performed Fst 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 (R2) was measured with plink v1.90b6.6 (*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 > 0.2 and with R2 > 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 Fst analyses and results from another study (*43*), we removed them from subsequent analyses.

Genomic PCA were also used to assess genotypes at the P supergene based on the genetic variation segregating within the regions of P1, P2 and P3 (Figure S1-4). Indeed, several study have shown that PCA can be used to discriminate between inversion genotypes at supergenes (*29*, *39–41*, *65*, *66*). This

was notably used in a 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, Hn123/Hn123. For simplicity, we refer hereafter to specimens without inversions (Hn0/Hn0) as the Hn0 samples, and to the samples with the three inversions (Hn1/Hn123, Hn0/Hn123 or Hn123/Hn123) as the Hn123 samples. Unless otherwise stated, every analysis presented hereafter was computed on these two sample sets separately.

Gene expression analyses

RNAseq data from (*47*) were reanalysed using the EdgeR R package [v3.16.5; (*48*)]. Gene expression in early pupal (24h) 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 wings 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 was normalized with the calcNormFactors function. The dispersion was estimated with the estimateDisp function. Data was fitted with a quasi-likelihood negative binomial generalized log-linear model with the glmQLFTest function and p-values were ajusted using the Benjamini & Hochberg correction for false discovery rate.

In order to study the evolutionary history of the supergene and determine whether recombination occur at the supergene, we computed sliding window phylogenies along the supergene, using 10 kb windows with RaxML (*67*). Only sample homozygous for the inversions or for their absence were used (Hn0/Hn0, Hn1/Hn1, 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, a Hn123/Hn123 individuals might have an a*urora*/t*arapotensis* 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 classes, phylogeny topology were highly variable along the supergene. To summarise these variations (topology weighting), we used Twisst (*68*) using the different morph as different taxa. We used the morphs s*ilvana*, *bicoloratus, aurora*, *timaeus, lyrcaeus*, t*arapotensis* and *messene* to perform such analyses *s*ince they were the only morphs with sufficient samples.

Phenotyping

Color pattern variation was described using Color Pattern Modeling (CPM) according to the developer's recommendations (*36*). 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. Wing 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 (*36*) for more details). This procedure is described as a pattern-based alignment, and results in the optimal compromise in the pixelwise superimposition of the different pattern elements constituting the wing pattern. This has the advantage of focusing the quantification on the change in pattern elements relatively 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 noticeable difference in the overall positioning of the pattern, due to the presence of androcony in males, (*36*)). Wing-pattern phenotypic variation could then be described as the color variation among all common pixel 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 principal component analysis (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 part/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) were 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 colour 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 samples harbouring different rearrangements, we performed PCAs on subset of samples, based on their genotype (presence/absence of the three inversions). Therefore, we computed PCAs only on samples homozygous or heterozygous for chromosome form

Hn123 and on samples homozygous for chromosome form Hn0, respectively (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 components) into pixel values using the eigenvalues of the phenotypic PCA. See MV-PLINK and CPM references for further details (*26*, *36*, *44*).

Multivariate association studies

To determine the genetic basis of colour pattern variation, multivariate genome wide association studies were performed using MV-PLINK (v1.6; (*44*)). MV-PLINK performs a Canonical Correlation Analysis (CCA) 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 patterns 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 isolating 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 minor allele frequency (MAF) > 0.02 and genotyping rate > 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: 500000-4000000, 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, $1e^6$ 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 2b and 2e), 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 analyses, 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 result sfrom the tight physical linkage of loci associated with wing pattern variation, especially around *cortex*. Hence, the list of associated regions (Figure 3g-h and 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.

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 v1.90b6.24 (*45*). We used the first 10 principal components of whole-genome PCA (see the *Population genetic analyses* sections and Figure S5) as covariates in these analyses (computed only with Hn123 or Hn0 samples). 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 summarising 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 aspect 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 Figure S9). For each of these specific aspect of wing variation, we computed association using the first three principal components in associations. Figure 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). Figure S17 and S19 show the results of these associations (only the result for the first principal components are shown). To distinguish regions that present an enrichment in significantly associated variants (Figure 2c and 2f), 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

analyses, 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).

Individual based simulations.

We used individual-based simulations to estimate our ability to identify the loci responsible for color variation in a supergene. In order to keep a tractable model, we used in simulations a simplified scenario mimicking *H. numata* evolution. We used SLiM V3.2 (*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 such as phenotypes beneficial in a population were counted selected 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 2Mb 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=1e-06. Each occurring mutations 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, 0.5 with uniform probabilities. In nature, the effective population size of *H. numata* have been estimated to be Ne=23,089,618 (*43*), with a mutation rate ~2.9 x 10^{-9} (*72*) and a recombination rate of ~0.6 x 10^{-5} (*73*). Considering the much smaller population size 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 the population 1 a single mutation mimicking a 1Mb inversion suppressing recombination over the region 500,000-1,500,000bp . 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.

During the whole duration of the simulations, the fitness of individual 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). Individual 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 population, individual homozygous for the Hn123 arrangement have a larval survival ~0.4 lower than individual heterozygous (Hn1/Hn123 or Hn0/Hn123; (*29*)). Individual without the inversion and without any colour derived mutation (ancestral state) had their fitness multiplied by a factor 0.5. These individual for instance represent the morph “silvana”, which are without inversion and are poorly protected against predator (*33*). Individuals with the inversion (homozygous or heterozygous) and without any colour derived mutation (ancestral state) have their fitness multiplied by a factor 0.8. Individuals with the inversion and the two derived colour mutation (being homozygous or heterozygous) had their fitness multiplied by a factor 0.3 (these individual are considered as “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 if they were in population 1, and by 0.7 if they were in population 2. Individuals with only the derived allele at the colour locus 2 had their fitness multiplied by 0.7 if they were in population 1, and by 1 if they are in population 2. Table S3 summarize these fitness parameters.

In sum, in our simulations, the inversion is a beneficial variant in both population, but is also associated with an homozygous disadvantage maintaining it at intermediate frequency (~10-50% in the parameter used). This mimics the situation encountered in *H. numata* (*29*). The two colour loci are epistatic (their fitness effect depend on the allele at the other locus) and are under disruptive selection. Each allele is favoured in a population and selected against in the other. Individuals with the two derived mutations are counter selected. Different combinations of alleles are maintained in population, 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 mimic 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 purpose two populations connected by high gene flow.

At simulation end (generation 120,000), we randomly selected 60 samples that harboured either only the derived colour allele 1 (homozygous or heterozygous) or only the derived colour allele 2 (homozygous or heterozygous) and this regardless of their population of origin. This mimics the sampling of individual with a mimetic forms know 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. Individual with the colour allele 1 were determined to have a phenotype of 1-γ, with γ being sampled between 0 and *PheVar* with uniform probabilities, *PheVar* being 0.2, 0.4 or 1.0. Individual with the colour allele 2 were determined to have a phenotype of -1+γ, with *gamma* being sampled between 0 and PheVar with uniform probabilities, PheVar being 0.2, 0.4 or 1.0. The PheVar variable therefore represent the fraction of phenotypic variation that is explained by factors other that the genotype at the colour locus. This was used to simulate variations in phenotype measures due to the environment for example.

Finally, we used PLINK v1.90b6.24 (REF) to perform association study using this simulated genomes and phenotypes using the same parameters and approach as used for the analyses of *H. numata* data. We display in figure S22 some examples of these associations. It shows that even using parameters that are prone to generate an 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 our approach. Only extreme parameter values, notably an high 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.

Acknowledgments

We are very grateful to Oscar Puebla, Floriane Coulmance, and to the Puebla lab for their careful review and useful comments on this manuscript. 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 research was supported by Agence Nationale de la Recherche (ANR) grants ANR-12-JSV7-0005 and ANR-18-CE02-0019-01 and European Research Council grant ERC-StG-243179 to MJ. 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. T. Lenormand, S. P. Otto, The evolution of recombination in a heterogeneous environment. *Genetics*. **156**, 423–438 (2000).

2. S. Yeaman, Genomic rearrangements and the evolution of clusters of locally adaptive loci. *PNAS*. **110**, E1743–E1751 (2013).

3. S. Lamichhaney, G. Fan, F. Widemo, U. Gunnarsson, D. S. Thalmann, M. P. Hoeppner, S. Kerje, U. Gustafson, C. Shi, H. Zhang, W. Chen, X. Liang, L. Huang, J. Wang, E. Liang, Q. Wu, S. M.-Y. Lee, X. Xu, J. Höglund, X. Liu, L. Andersson, Structural genomic changes underlie alternative reproductive strategies in the ruff (Philomachus pugnax). *Nat. Genet.* **48**, 84–88 (2016).

4. E. M. Tuttle, A. O. Bergland, M. L. Korody, M. S. Brewer, D. J. Newhouse, P. Minx, M. Stager, A. Betuel, Z. A. Cheviron, W. C. Warren, R. A. Gonser, C. N. Balakrishnan, Divergence and functional degradation of a sex chromosome-like supergene. *Curr. Biol.* **26**, 344–350 (2016).

5. K. Kunte, W. Zhang, A. Tenger-Trolander, D. H. Palmer, A. Martin, R. D. Reed, S. P. Mullen, M. R. Kronforst, *doublesex* is a mimicry supergene. *Nature*. **507**, 229–232 (2014).

6. J. Wang, Y. Wurm, M. Nipitwattanaphon, O. Riba-Grognuz, Y.-C. Huang, D. Shoemaker, L. Keller, A Y-like social chromosome causes alternative colony organization in fire ants. *Nature*. **493**, 664–668 (2013).

7. T. Kess, P. Bentzen, S. J. Lehnert, E. V. A. Sylvester, S. Lien, M. P. Kent, M. Sinclair-Waters, C. J. Morris, P. Regular, R. Fairweather, I. R. Bradbury, A migration-associated supergene reveals loss of biocomplexity in Atlantic cod. *Science Advances*. **5**, eaav2461 (2019).

8. H. Stefansson, A. Helgason, G. Thorleifsson, V. Steinthorsdottir, G. Masson, J. Barnard, A. Baker, A. Jonasdottir, A. Ingason, V. G. Gudnadottir, N. Desnica, A. Hicks, A. Gylfason, D. F. Gudbjartsson, G. M. Jonsdottir, J. Sainz, K. Agnarsson, B. Birgisdottir, S. Ghosh, A. Olafsdottir, J.-B. Cazier, K. Kristjansson, M. L. Frigge, T. E. Thorgeirsson, J. R. Gulcher, A. Kong, K. Stefansson, A common inversion under selection in Europeans. *Nat. Genet.* **37**, 129–137 (2005).

9. V. Castric, J. S. Bechsgaard, S. Grenier, R. Noureddine, M. H. Schierup, X. Vekemans, Molecular evolution within and between self-incompatibility specificities. *Molecular Biology and Evolution*. **27**, 11–20 (2010).

10. C. N. Huu, B. Keller, E. Conti, C. Kappel, M. Lenhard, Supergene evolution via stepwise duplications and neofunctionalization of a floral-organ identity gene. *PNAS*. **117**, 23148–23157 (2020).

11. C. Kappel, C. N. Huu, M. Lenhard, A short story gets longer: recent insights into the molecular basis of heterostyly. *J Exp Bot*. **68**, 5719–5730 (2017).

12. T. Schwander, R. Libbrecht, L. Keller, Supergenes and complex phenotypes. *Curr Biol*. **24**, R288-294 (2014).

13. R. Villoutreix, D. Ayala, M. Joron, Z. Gompert, J. L. Feder, P. Nosil, Inversion breakpoints and the evolution of supergenes. *Molecular Ecology*. **30**, 2738–2755 (2021).

14. M. Wellenreuther, L. Bernatchez, Eco-evolutionary genomics of chromosomal inversions. *Trends Ecol. Evol.* **33**, 427–440 (2018).

15. K. Mather, The genetical architecture of heterostyly in primula sinensis. *Evolution*. **4**, 340–352 (1950).

16. Q. Yang, D. Zhang, Q. Li, Z. Cheng, Y. Xue, Heterochromatic and genetic features are consistent with recombination suppression of the self-incompatibility locus in Antirrhinum. *Plant J*. **51**, 140–151 (2007).

17. T. Kao, T. Tsukamoto, The molecular and genetic bases of S-RNase-based self-incompatibility. *The Plant Cell*. **16**, S72–S83 (2004).

18. T. Chookajorn, A. Kachroo, D. R. Ripoll, A. G. Clark, J. B. Nasrallah, Specificity determinants and diversification of the Brassica self-incompatibility pollen ligand. *PNAS*. **101**, 911–917 (2004).

19. S. Branco, H. Badouin, R. C. R. de la Vega, J. Gouzy, F. Carpentier, G. Aguileta, S. Siguenza, J.-T. Brandenburg, M. A. Coelho, M. E. Hood, T. Giraud, Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. *PNAS*. **114**, 7067–7072 (2017).

20. F. E. Hartmann, M. Duhamel, F. Carpentier, M. E. Hood, M. Foulongne Oriol, P. Silar, F. Malagnac, P. Grognet, T. Giraud, Recombination suppression and evolutionary strata around mating-type loci in fungi: documenting patterns and understanding evolutionary and mechanistic causes. *New Phytologist*. **229**, 2470–2491 (2021).

21. N. J. Nadeau, C. Pardo-Diaz, A. Whibley, M. A. Supple, S. V. Saenko, R. W. R. Wallbank, G. C. Wu, L. Maroja, L. Ferguson, J. J. Hanly, H. Hines, C. Salazar, R. M. Merrill, A. J. Dowling, R. H. ffrench-Constant, V. Llaurens, M. Joron, W. O. McMillan, C. D. Jiggins, The gene *cortex* controls mimicry and crypsis in butterflies and moths. *Nature*. **534**, 106–110 (2016).

22. R. D. Reed, R. Papa, A. Martin, H. M. Hines, B. A. Counterman, C. Pardo-Diaz, C. D. Jiggins, N. L. Chamberlain, M. R. Kronforst, R. Chen, G. Halder, H. F. Nijhout, W. O. McMillan, optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science*. **333**, 1137–1141 (2011).

23. E. L. Westerman, N. W. VanKuren, D. Massardo, A. Tenger-Trolander, W. Zhang, R. I. Hill, M. Perry, E. Bayala, K. Barr, N. Chamberlain, T. E. Douglas, N. Buerkle, S. E. Palmer, M. R. Kronforst, Aristaless controls butterfly wing color variation used in mimicry and mate choice. *Curr Biol*. **28**, 3469-3474.e4 (2018).

24. R. M. Merrill, K. K. Dasmahapatra, J. W. Davey, D. D. Dell'Aglio, J. J. Hanly, B. Huber, C. D. Jiggins, M. Joron, K. M. Kozak, V. Llaurens, S. H. Martin, S. H. Montgomery, J. Morris, N. J. Nadeau, A. L. Pinharanda, N. Rosser, M. J. Thompson, S. Vanjari, R. W. R. Wallbank, Q. Yu, The diversification of Heliconius butterflies: what have we learned in 150 years? *J Evol Biol*. **28**, 1417–1438 (2015).

25. A. Martin, R. Papa, N. J. Nadeau, R. I. Hill, B. A. Counterman, G. Halder, C. D. Jiggins, M. R. Kronforst, A. D. Long, W. O. McMillan, R. D. Reed, Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *PNAS*. **109**, 12632–12637 (2012).

26. B. Huber, A. Whibley, Y. L. Poul, N. Navarro, A. Martin, S. Baxter, A. Shah, B. Gilles, T. Wirth, W. O. McMillan, M. Joron, Conservatism and novelty in the genetic architecture of adaptation in Heliconius butterflies. *Heredity*. **114**, 515–524 (2015).

27. J. I. Meier, P. A. Salazar, M. Kučka, R. W. Davies, A. Dréau, I. Aldás, O. B. Power, N. J. Nadeau, J. R. Bridle, C. Rolian, N. H. Barton, W. O. McMillan, C. D. Jiggins, Y. F. Chan, Haplotype tagging reveals parallel formation of hybrid races in two butterfly species. *PNAS*. **118** (2021), doi:10.1073/pnas.2015005118.

28. M. Kirkpatrick, N. Barton, Chromosome inversions, local adaptation and speciation. *Genetics*. **173**, 419–434 (2006).

29. P. Jay, M. Chouteau, A. Whibley, H. Bastide, H. Parrinello, V. Llaurens, M. Joron, Mutation load at a mimicry supergene sheds new light on the evolution of inversion polymorphisms. *Nat. Genet.* **53**, 288–293 (2021).

30. N. B. Edelman, P. B. Frandsen, M. Miyagi, B. Clavijo, J. Davey, R. B. Dikow, G. García-Accinelli, S. M. Van Belleghem, N. Patterson, D. E. Neafsey, R. Challis, S. Kumar, G. R. P. Moreira, C. Salazar, M. Chouteau, B. A. Counterman, R. Papa, M. Blaxter, R. D. Reed, K. K. Dasmahapatra, M. Kronforst, M. Joron, C. D. Jiggins, W. O. McMillan, F. Di Palma, A. J. Blumberg, J. Wakeley, D. Jaffe, J. Mallet, Genomic architecture and introgression shape a butterfly radiation. *Science*. **366**, 594–599 (2019).

31. P. Jay, A. Whibley, L. Frézal, M. Á. Rodríguez de Cara, R. W. Nowell, J. Mallet, K. K. Dasmahapatra, M. Joron, Supergene evolution triggered by the introgression of a chromosomal inversion. *Current Biology*. **28**, 1839-1845.e3 (2018).

32. K. S. Brown, W. W. Benson, Adaptive polymorphism associated with multiple müllerian mimicry in Heliconius numata (Lepid. Nymph.). *Biotropica*. **6**, 205–228 (1974).

33. M. Chouteau, M. Arias, M. Joron, Warning signals are under positive frequency-dependent selection in nature. *PNAS*. **113**, 2164–2169 (2016).

34. M. Arias, J. W. Davey, S. Martin, C. Jiggins, N. Nadeau, M. Joron, V. Llaurens, How do predators generalize warning signals in simple and complex prey communities? Insights from a videogame. *Proceedings of the Royal Society B: Biological Sciences*. **287**, 20200014 (2020).

35. M. Joron, L. Frezal, R. T. Jones, N. L. Chamberlain, S. F. Lee, C. R. Haag, A. Whibley, M. Becuwe, S. W. Baxter, L. Ferguson, P. A. Wilkinson, C. Salazar, C. Davidson, R. Clark, M. A. Quail, H. Beasley, R. Glithero, C. Lloyd, S. Sims, M. C. Jones, J. Rogers, C. D. Jiggins, R. H. ffrench-Constant, Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature*. **477**, 203–206 (2011).

36. Y. Le Poul, A. Whibley, M. Chouteau, F. Prunier, V. Llaurens, M. Joron, Evolution of dominance mechanisms at a butterfly mimicry supergene. *Nature Communications*. **5**, 5644 (2014).

37. S. M. Van Belleghem, P. Rastas, A. Papanicolaou, S. H. Martin, C. F. Arias, M. A. Supple, J. J. Hanly, J. Mallet, J. J. Lewis, H. M. Hines, M. Ruiz, C. Salazar, M. Linares, G. R. P. Moreira, C. D. Jiggins, B. A. Counterman, W. O. McMillan, R. Papa, Complex modular architecture around a simple toolkit of wing pattern genes. *Nature Ecology & Evolution*. **1**, 1–12 (2017).

38. M. Joron, R. Papa, M. Beltrán, N. Chamberlain, J. Mavárez, S. Baxter, M. Abanto, E. Bermingham, S. J. Humphray, J. Rogers, H. Beasley, K. Barlow, R. H. ffrench-Constant, J. Mallet, W. O. McMillan, C. D. Jiggins, A conserved supergene locus controls colour pattern diversity in Heliconius butterflies. *PLOS Biology*. **4**, e303 (2006).

39. J. Ma, C. I. Amos, Investigation of inversion polymorphisms in the human genome using principal components analysis. *PLOS ONE*. **7**, e40224 (2012).

40. D. Lindtke, K. Lucek, V. Soria Carrasco, R. Villoutreix, T. E. Farkas, R. Riesch, S. R. Dennis, Z. Gompert, P. Nosil, Long-term balancing selection on chromosomal variants associated with crypsis in a stick insect. *Molecular Ecology*. **26**, 6189–6205 (2017).

41. R. Faria, P. Chaube, H. E. Morales, T. Larsson, A. R. Lemmon, E. M. Lemmon, M. Rafajlović, M. Panova, M. Ravinet, K. Johannesson, A. M. Westram, R. K. Butlin, Multiple chromosomal rearrangements in a hybrid zone between Littorina saxatilis ecotypes. *Mol. Ecol.* **28**, 1375–1393 (2019).

42. M. Chouteau, V. Llaurens, F. Piron-Prunier, M. Joron, Polymorphism at a mimicry supergene maintained by opposing frequency-dependent selection pressures. *PNAS*, 201702482 (2017).

43. M. Á. R. de Cara, P. Jay, M. Chouteau, A. Whibley, B. Huber, F. Piron-Prunier, R. R. Ramos, A. V. L. Freitas, C. Salazar, K. L. Silva-Brandão, T. T. Torres, M. Joron, "Supergene formation is associated with a major shift in genome-wide patterns of diversity in a butterfly" (2021), p. 2021.09.29.462348, , doi:10.1101/2021.09.29.462348.

44. M. A. R. Ferreira, S. M. Purcell, A multivariate test of association. *Bioinformatics*. **25**, 132–133 (2009).

45. S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, P. C. Sham, PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. **81**, 559–575 (2007).

46. J. Hellwege, J. Keaton, A. Giri, X. Gao, D. R. Velez Edwards, T. L. Edwards, Population Stratification in Genetic Association Studies. *Current protocols in human genetics*. **95** (2017).

47. S. V. Saenko, M. Chouteau, F. Piron-Prunier, C. Blugeon, M. Joron, V. Llaurens, Unravelling the genes forming the wing pattern supergene in the polymorphic butterfly Heliconius numata. *Evodevo*. **10**, 16 (2019).

48. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. **26**, 139–140 (2010).

49. C. A. Clarke, P. M. Sheppard, Super-genes and mimicry. *Heredity*. **14**, 175–185 (1960).

50. M. J. T. N. Timmermans, S. W. Baxter, R. Clark, D. G. Heckel, H. Vogel, S. Collins, A. Papanicolaou, I. Fukova, M. Joron, M. J. Thompson, C. D. Jiggins, R. H. ffrench-Constant, A. P. Vogler, Comparative genomics of the mimicry switch in Papilio dardanus. *Proceedings of the Royal Society B: Biological Sciences*. **281**, 20140465 (2014).

51. M. Joron, Polymorphic mimicry, microhabitat use, and sex-specific behaviour. *Journal of Evolutionary Biology*. **18**, 547–556 (2005).

52. L. Maisonneuve, M. Chouteau, M. Joron, V. Llaurens, Evolution and genetic architecture of disassortative mating at a locus under heterozygote advantage. *Evolution*. **n/a**, doi:https://doi.org/10.1111/evo.14129.

53. L. Livraghi, J. J. Hanly, S. M. Van Bellghem, G. Montejo-Kovacevich, E. S. van der Heijden, L. S. Loh, A. Ren, I. A. Warren, J. J. Lewis, C. Concha, L. Hebberecht, C. J. Wright, J. M. Walker, J. Foley, Z. H. Goldberg, H. Arenas-Castro, C. Salazar, M. W. Perry, R. Papa, A. Martin, W. O. McMillan, C. D. Jiggins, Cortex cis-regulatory switches establish scale colour identity and pattern diversity in Heliconius. *eLife*. **10**, e68549 (2021).

54. N. W. VanKuren, D. Massardo, S. Nallu, M. R. Kronforst, Butterfly mimicry polymorphisms highlight phylogenetic limits of gene reuse in the evolution of diverse adaptations. *Molecular Biology and Evolution*. **36**, 2842–2853 (2019).

55. S. Yeaman, S. Aeschbacher, R. Bürger, The evolution of genomic islands by increased establishment probability of linked alleles. *Mol Ecol*. **25**, 2542–2558 (2016).

56. R. Bürger, A. Akerman, The effects of linkage and gene flow on local adaptation: a two-locus continent-island model. *Theor Popul Biol*. **80**, 272–288 (2011).

57. J. W. Davey, M. Chouteau, S. L. Barker, L. Maroja, S. W. Baxter, F. Simpson, R. M. Merrill, M. Joron, J. Mallet, K. K. Dasmahapatra, C. D. Jiggins, Major improvements to the Heliconius melpomene genome assembly used to confirm 10 chromosome fusion events in 6 million years of butterfly evolution. *G3: Genes, Genomes, Genetics*. **6**, 695–708 (2016).

58. F. J. Sedlazeck, P. Rescheneder, A. von Haeseler, NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics*. **29**, 2790–2791 (2013).

59. M. A. DePristo, E. Banks, R. E. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, M. J. Daly, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. **43**, 491–498 (2011).

60. P. Cingolani, A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, D. M. Ruden, 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 (2012).

61. J. W. Davey, S. L. Barker, P. M. Rastas, A. Pinharanda, S. H. Martin, R. Durbin, W. O. McMillan, R. M. Merrill, C. D. Jiggins, No evidence for maintenance of a sympatric Heliconius species barrier by chromosomal inversions. *Evolution Letters*. **1**, 138–154 (2017).

62. S. Kurtz, A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway, C. Antonescu, S. L. Salzberg, Versatile and open software for comparing large genomes. *Genome Biology*. **5**, R12 (2004).

63. M. Krzywinski, J. Schein, İ. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, M. A. Marra, Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).

64. X. Zheng, D. Levine, J. Shen, S. M. Gogarten, C. Laurie, B. S. Weir, A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*. **28**, 3326–3328 (2012).

65. M. Todesco, G. L. Owens, N. Bercovich, J.-S. Légaré, S. Soudi, D. O. Burge, K. Huang, K. L. Ostevik, E. B. M. Drummond, I. Imerovski, K. Lande, M. A. Pascual-Robles, M. Nanavati, M. Jahani, W. Cheung, S. E. Staton, S. Muños, R. Nielsen, L. A. Donovan, J. M. Burke, S. Yeaman, L. H. Rieseberg, Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature*. **584**, 602–607 (2020).

66. C. Mérot, R. A. Oomen, A. Tigano, M. Wellenreuther, A roadmap for understanding the evolutionary significance of structural genomic variation. *Trends Ecol. Evol.* **35**, 561–572 (2020).

67. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. **30**, 1312–1313 (2014).

68. S. H. Martin, S. M. V. Belleghem, Exploring evolutionary relationships across the genome using topology weighting. *Genetics*. **206**, 429–438 (2017).

69. Y. Le Poul, thesis, Paris, Muséum national d'histoire naturelle (2010).

70. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools. *Bioinformatics*. **25**, 2078–2079 (2009).

71. B. C. Haller, P. W. Messer, SLiM 3: Forward genetic simulations beyond the wright–fisher model. *Mol. Biol. Evol.* **36**, 632–637 (2019).

72. P. D. Keightley, A. Pinharanda, R. W. Ness, F. Simpson, K. K. Dasmahapatra, J. Mallet, J. W. Davey, C. D. Jiggins, Estimation of the spontaneous mutation rate in Heliconius melpomene. *Molecular Biology and Evolution*. **32**, 239–243 (2015).

73. L. Wilfert, J. Gadau, P. Schmid-Hempel, Variation in genomic recombination rates among animal taxa and the case of social insects. *Heredity*. **98**, 189–197 (2007).

Author contributions: P.J. and M.J. designed the study; P.J. and M.J. wrote the paper; P.J. and A.W. generated the genomic data; Butterflies were collected by P.J., M.J., M.C. and M.A.; P.J. and M.L. performed the association studies; P.J., M.L. and Y.L.P. performed the phenotyping; All authors contributed to editing the manuscript; M.J. supervised the study.

Data availability: Codes used to produce analyses have been deposited on Github: https://github.com/PaulYannJay/HeliconiusGWAS. The raw sequence data were deposited in NCBI SRA and accession

numbers are indicated in Supplementary table 1 (the new genomes will be deposited upon manuscript acceptance). The whole genome VCF file is available upon request.

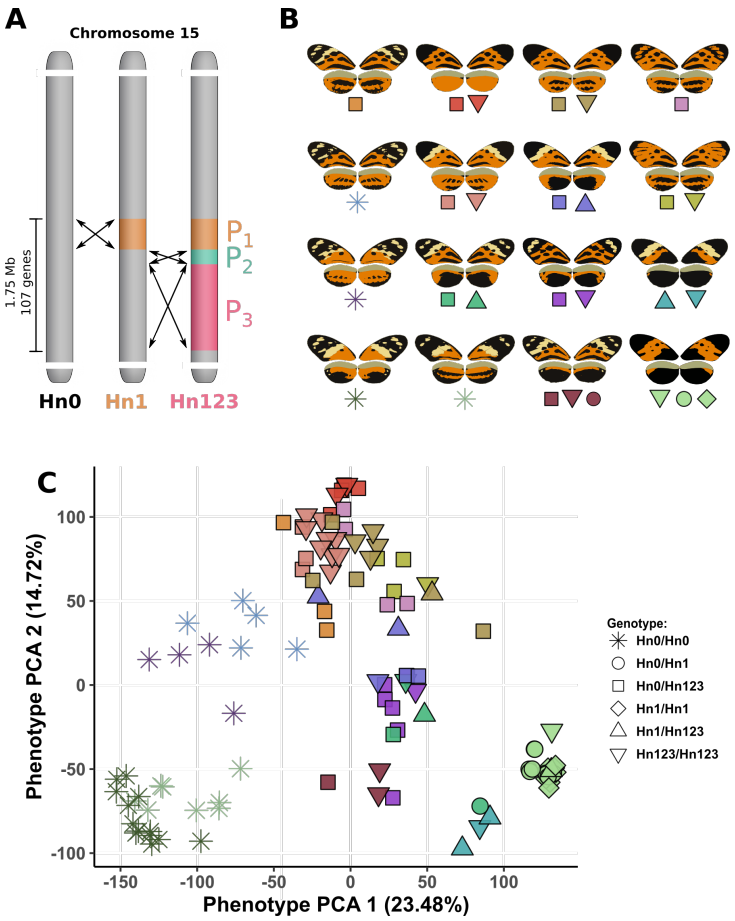


Figure 1 | Genetic architecture and wing pattern diversity in H. numata

a, Genetic architecture of the H. numata mimicry supergene P characterized by three polymorphic inversions of respective size 400kb, 200kb and 1150 kb (REF) **b,** Schematic diversity of wing patterns of *H. numata* in our dataset. **c**, Two-dimension approximation of the morphological space representing the phenotype diversity observed in H. numata. The dotplot displays results from a Principal Component Analysis (the first two components are displayed here) computed on wing pattern variations as obtained using CPM (*36*). For display purposes, butterflies were manually classified into mimetic forms based on the literature (*32*); 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 Figure S8. PCAs computed on samples with the same supergene arrangement and on more specific parts of wing pattern are presented in in Figure S9.

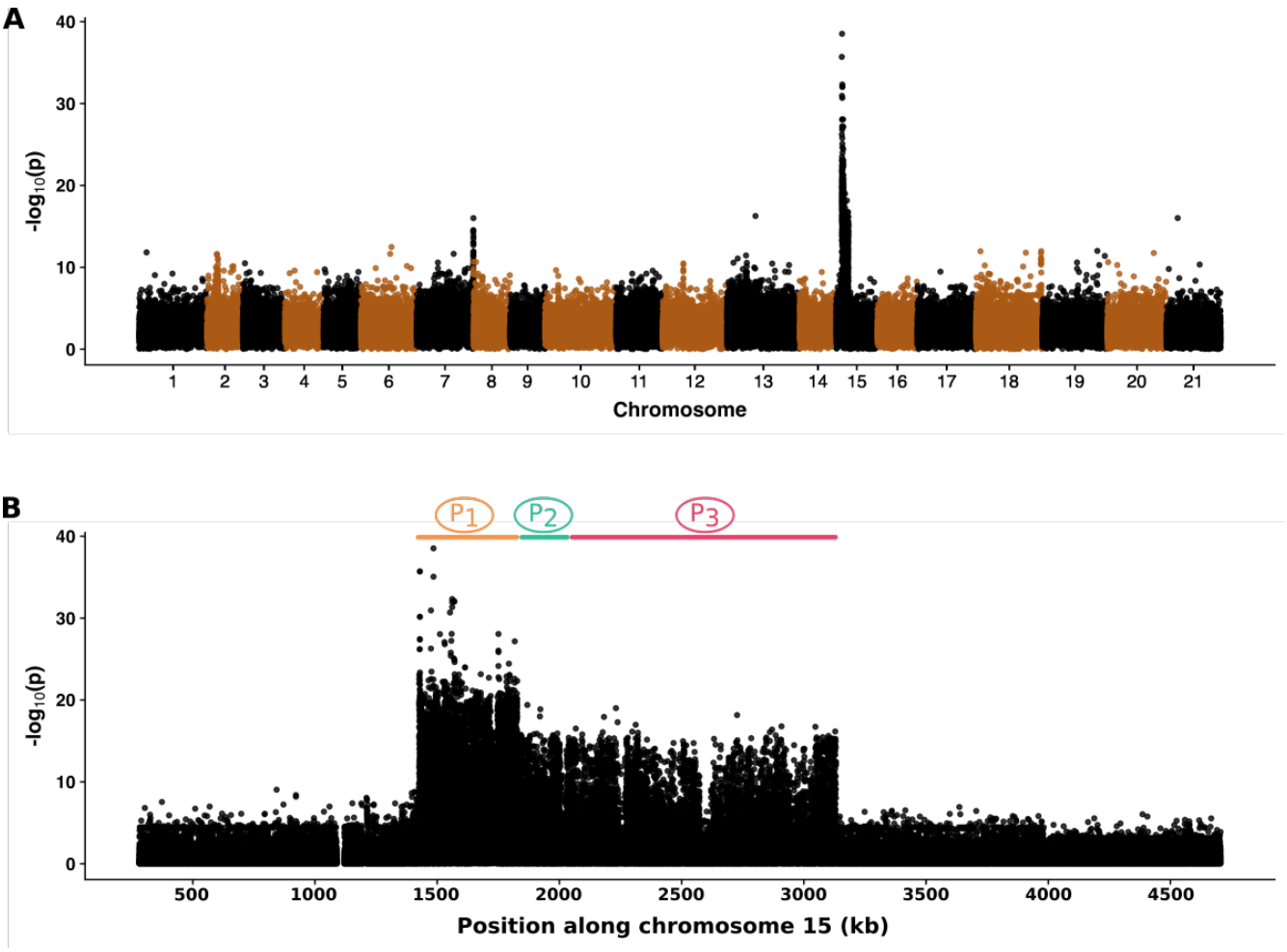


Figure 2 | Genome wide association of genetic and wing pattern variation.
A, Multivariate association study using as phenotype the first six principal components describing wing pattern variations (presented in Figure 1c and S8) and all samples regardless of their genotype at the supergene. One major 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, and in reality this region is laying within the supergene region (see methods). **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.

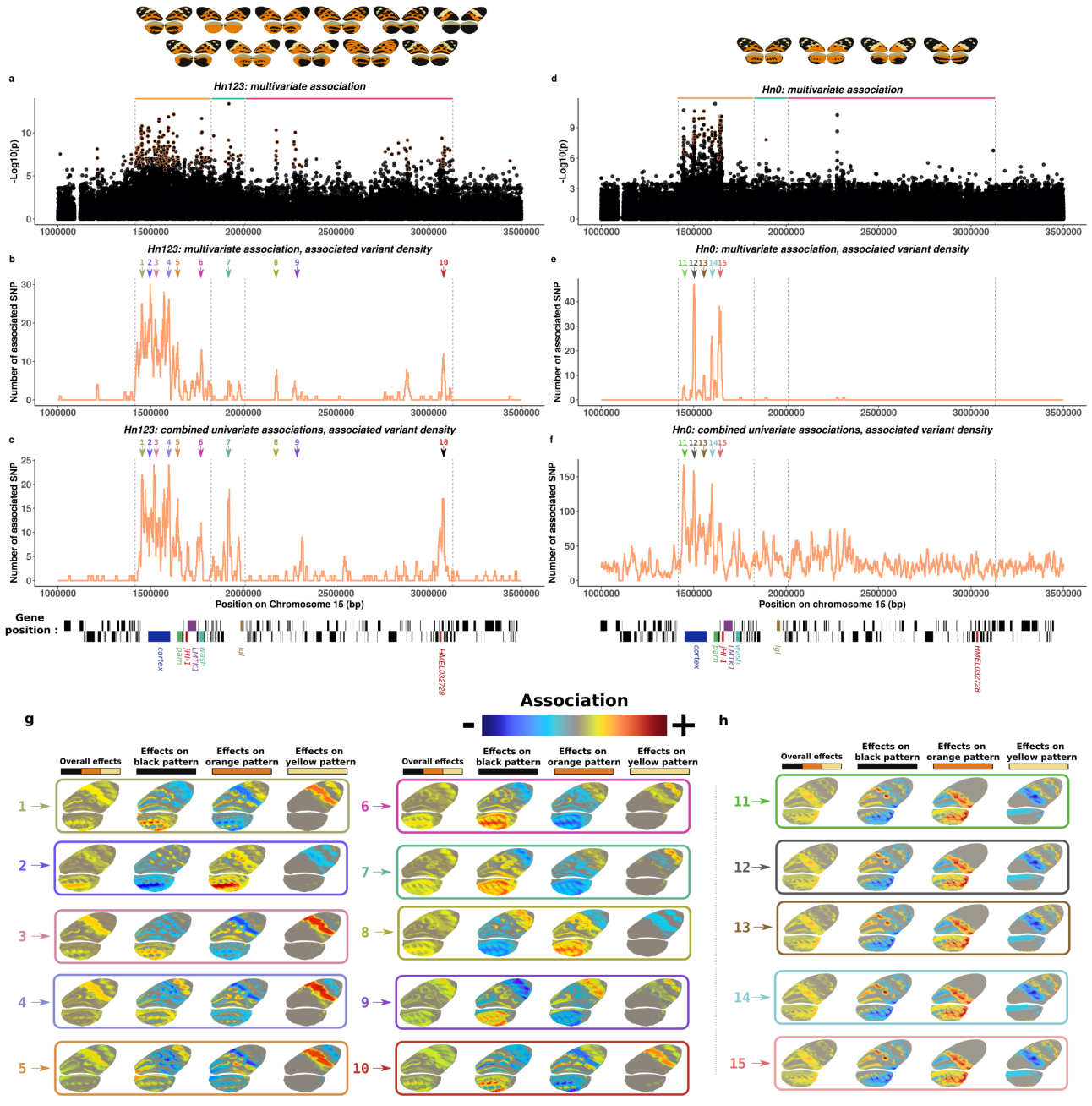
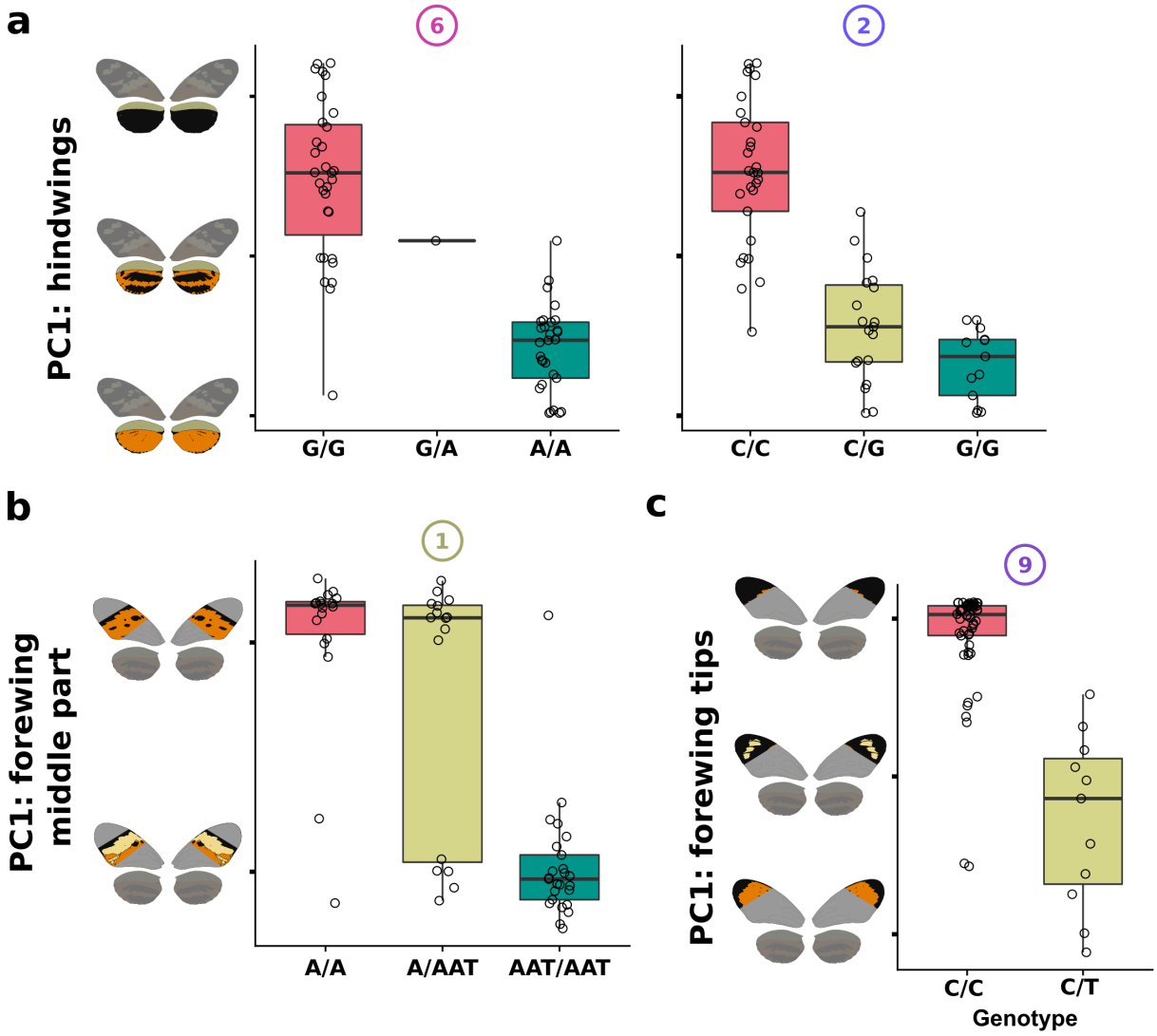


Figure 3 | Distinct regions within the supergene are associated with variation in wing pattern features.

a,d, Multivariate association studies computed on Hn123 and Hn0 samples, 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. $1e^{-6}$ permutations were performed for each variant. Variants highlighted in orange are variants with an empirical p-value < $1e^{-6}$ (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

figures S14-15**.** **b,e** Density of significantly associated variants in multivariate analyses (with empirical p-value < $1e^{-6}$) along the chromosome 15. Analyses computed in 10,000bp overlapping sliding window (with 100bp overlap). All significantly associated variants in one or more of the multivariate association analyses (using 2,3, 4, …, 6 phenotypic principal components; figure S14-15) were used. **c,f,** Density of significantly associated variants in univariate analyses (with empirical p-value < $1e^{-6}$) along chromosome 15. Analyses computed in 10,000bp overlapping sliding windows (with 100bp 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; figure S16-19) 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 (panels **b-f,** coloured arrows) to the wing pattern in Hn123 or Hn0 samples, respectively. Heatmaps from blue to red represent, for every pixel, the strength and direction of association of the derived allele, that is how the allelic change at a given genetic position affect 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 direction (i.e. red and blue values) in the same wing area in two colour-specific heatmaps indicates 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.



690 **Figure 4 | Effect of four selected variants on Hn123 wing pattern variations.** Representation of the association of some genetic variants with specific wing-pattern variation. See Figure 3g-h and 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 of the phenotype (y-axis): PCA computed on hindwings only (A), on the middle part of the 695 forewings only (B), on the tips of the forewings only (C). See Figure S9 for another representation of the principal components. Each dot is an individual. Instead of annotating y-axis with eigenvalues (values of individual on principal components), schematic butterflies with average phenotypes along the principal components were displayed. Boxplot elements: central line, median; box limits, 25th and 75th percentiles; whiskers, 1.5 × interquartile range. **a**, Effect of the most strongly associated 700 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.