Nature Genetics
March 2021, Volume 53 Issue 3 Pages 288-293
https://doi.org/10.1038/s41588-020-00771-1
https://archimer.ifremer.fr/doc/00677/78932/

Mutation load at a mimicry supergene sheds new light on the evolution of inversion polymorphisms

Jay Paul [1, *] , Chouteau Mathieu [1, 2, *], Whibley Annabel [3], Bastide Héloïse [4], Parrinello Hugues [5], Llaurens Violaine [4], Joron Mathieu [1, *]

[1] CEFE, Université de Montpellier, CNRS, EPHE, IRD, Université Paul Valéry Montpellier 3, Montpellier, France
[2] LEEISA, Université de Guyane, CNRS, IFREMER, Cayenne, French Guiana
[3] School of Biological Sciences, University of Auckland, Auckland, New Zealand
[4] ISYEB, MNHN, CNRS, Sorbonne Universités, EPHE, Paris, France
[5] MGX, Biocampus Montpellier, Université de Montpellier, CNRS, INSERM, Montpellier, France

Corresponding authors, email addresses : paul.yann.jay@gmail.com ; mathieu.chouteau@cnrs.Fr ; mathieu.joron@cefe.cnrs.fr

Abstract :

Chromosomal inversions are ubiquitous in genomes and often coordinate complex phenotypes, such as the covariation of behavior and morphology in many birds, fishes, insects or mammals1,2,3,4,5,6,7,8,9,10,11. However, why and how inversions become associated with polymorphic traits remains obscure. Here we show that despite a strong selective advantage when they form, inversions accumulate recessive deleterious mutations that generate frequency-dependent selection and promote their maintenance at intermediate frequency. Combining genomics and in vivo fitness analyses in a model butterfly for wing-pattern polymorphism, Heliconius numata, we reveal that three ecologically advantageous inversions have built up a heavy mutational load from the sequential accumulation of deleterious mutations and transposable elements. Inversions associate with sharply reduced viability when homozygous, which prevents them from replacing ancestral chromosome arrangements. Our results suggest that other complex polymorphisms, rather than representing adaptations to competing ecological optima, could evolve because chromosomal rearrangements are intrinsically prone to carrying recessive harmful mutations.

Many organisms display concerted variation in their phenotypic traits. Consistent association of multiple phenotypic features, combining differences in behavior, morphology and physiology, may result in so-called syndromes, or complex traits with clear adaptive significance. This coordination is often controlled by chromosomal rearrangements. Examples include dimorphic social organization in several ant species[7], color displays and mating behaviors in many birds and butterflies[3-6,12], dimorphic flower morphology in plants[13], as well as the extreme cases provided by sexual dimorphism in numerous animals. Why and how these structurally-delimited complex polymorphisms arise is a long-standing puzzle in biology[11,14-17].

The so-called supergenes controlling the coordination of the multiple phenotypic features are characterized by suppression of recombination, often through polymorphic chromosomal rearrangements, which preserve alternative combinations of alleles at linked genes[4,7,11,13]. The encoded multi-feature phenotypes are often assumed to reflect the existence of multiple, distinct adaptive optima, and their maintenance in polymorphisms to result from antagonistic ecological factors such as differential survival or mating success[8,12,18,19]. Yet why and how polymorphic chromosomal rearrangements evolve and become associated with complex phenotypic variation is not understood. Moreover, many chromosomal rearrangements with unknown effects are maintained at intermediate frequencies in natural populations[20], which raises questions about the processes generating these protected genetic polymorphisms.

The Amazonian butterfly *Heliconius numata* displays one of the most spectacular wing pattern polymorphisms known, with up to seven morphs coexisting within a single locality, each one toxic and engaged in warning color mimicry with distinct groups of toxic species. The distinct morphs do not however all bring equal protection against predators because of differences in the abundance of co-mimics, which should oppose polymorphism and favor the fixation of the best

protected morph[21,22]. Morphs also display disassortative mate preferences, suggesting the forces fostering the maintenance of polymorphism also promoted a peculiar mating system[18]. Colour pattern polymorphisms in toxic prey such as *Heliconius* and monarch butterflies[23] or ladybirds[24] are a long-standing genetic paradox since those polymorphisms are iconic examples of the covariation of multiple phenotypic features[4-7,12], yet strong purifying selection on mimicry should inhibit their formation[14,21].

Polymorphism in *H. numata* is associated with chromosomal inversions at the mimicry locus (supergene *P*) on chromosome 15, forming three distinct haplotype classes[6]. The ancestral haplotype constitutes the class of recessive P alleles and is associated, for example, with the widespread morph *silvana*. Two classes of derived haplotypes are known, both associated with a chromosomal inversion called P_1 (~400kb, 21 genes), each conferring increased protection against predators via mimicry[22]. The first derived haplotype carries P_1 alone, and expresses the top dominant P allele determining the morph *bicoloratus*; the second class of derived haplotypes carries P_1 linked with additional yet still uncharacterized rearrangements and expresses alleles with intermediate levels of dominance, associated with a diversity of morphs, such as *tarapotensis* or *arcuella*[6,25]. Inversion polymorphism and supergene formation originated via the introgression of P_1 from the *H. pardalinus* lineage[26]. This incremental series of chromosomal rearrangements allows us to unravel in unprecedented detail the processes by which structural variation has become associated with directional and balancing selection.

From 10X linked reads, we generated de novo genome assemblies for 12 *H. numata* individuals of distinct morphs, revealing a history of supergene formation characterized by the accretion of three inversions with breakpoint reuse. Pairwise alignment of assemblies showed that the previously uncharacterized rearrangement associated with all intermediate dominant supergene alleles[6] is formed by two inversions: P_2 (200 kb, 15 genes), adjacent to P_1, and the longer P_3

(1150 kb, 71 genes), adjacent to P_2 (Figure 1A, Extended Data Figure 1). No enrichment of repetitive motifs or transposable elements were observed at the breakpoints.

Genotyping of 66 re-sequenced specimens by sliding-window PCA along the supergene confirmed the prevalence of these inversions and the dominance of derived arrangements (denoted Hn1 and Hn123) to the ancestral arrangement (denoted Hn0) (Figure 1B-C, Extended Data Figure 2). Multiple genes in the inverted regions showed significant differential expression compared to ancestral segments, but this likely reflects divergence rather than direct breakpoint effects (Extended Data Figure 3). Indeed, none of the breakpoints of P_1, P_2 or P_3 fell within a gene, and no transcript found in Hn0 specimens was missing, disrupted, or differentially spliced in specimens with inversions.

Haplotype differentiation (Fst) and estimates of inversion ages, obtained by determining the most recent coalescence events between Hn0+Hn1 and Hn123, and within Hn123, suggest that the P supergene has evolved in three steps, involving the introgression of P_1 followed by the successive occurrence of P_2 and P_3 between ca. 1.8 and 3.0 Mya (Figure 1D, Extended Data Figure 4 and 5). While variation in selection acting on the inversions could lead to erroneous estimates, here the sequential accretion of P_2 and P_3 is also supported by their breakpoint reuse with P_1 and P_2, respectively. The three adjacent inversions of *H. numata* are therefore of distinct ages and originated in distinct lineages[26], which provides a rare opportunity to partition their mutational history and distinguish the consequences of their formation from those resulting from their maintenance in polymorphism.

Recombination suppression between structural alleles with opposite orientations[27] is expected to cause the accumulation of deleterious mutations and transposable elements (TEs) due to inefficient purging[28,29]. Consistent with this prediction, estimation of TE dynamics obtained by computing whole genome TE divergence supports a burst of insertion within the inversions of

recently active TEs, observed in particular for TEs belonging to the RC, DNA and LINE classes (Figure 2A-B, Extended Data Figure 6 and 7). Inverted haplotypes show a significant size increase (mean=+9.47%) compared to their corresponding non-inverted region in Hn0 (Figure 2C) and this expansion was caused primarily (71.8 %) by recent TE insertions from these classes (Figure 2A-B).

To investigate the impact of polymorphic inversions on the accumulation of deleterious mutations, we calculated, independently on inverted and non-inverted segments, the rate of non-synonymous to synonymous polymorphism (pN/pS), the rate of non-synonymous to synonymous substitution (dN/dS) and the direction of selection (DoS[30]). Consistent with a low efficiency of selection in eliminating deleterious variants, P_1, P_2, and P_3 were all found to be enriched in non-synonymous relative to synonymous polymorphisms compared to the whole genome and to non-inverted ancestral segments (pN/pS_{P1}=0.83, pN/pS_{P2}=0.54, pN/pS_{P3}=0.49, Figure 3A). Since inversions in *H. numata* are associated with wing pattern variations, a proportion of these non-synonymous mutations could be involved in wing patterning and be under positive selection. The inversions were nonetheless found to be overall under negative selection (DoS_{P1}=-0.136, DoS_{P2} = -0.087, DoS_{P3}=-0.079), with values reflecting their sequential origin (Figure 3A). Because P_1 was introgressed from *H. pardalinus*[26] , mutations that accumulated in P_1 before the introgression (i.e. shared with *H. pardalinus*) could be distinguished from those arising after supergene formation in *H. numata* (i.e. unique to Hn1 and Hn123). This revealed that non-synonymous mutations which existed in the P_1 segment before the introgression underwent a high rate of fixation in *H. pardalinus* (dN/dS = 0.78, Extended Data Figure 8), and in *H. numata* (dN/dS=1.33, Figure 3B), suggesting that both the formation of P_1 and its introgression led to the fixation of deleterious mutations. By contrast, 99.9 % of the mutations that accumulated in P_1 after its introgression -i.e. after supergene formation- remain polymorphic in Hn1-Hn123 and a high proportion of them are

non-synonymous (dN/dS=0.00, pN/pS=0.978, DoS=-0.49, Figure 3B). Taken together, these results suggest that the inversions have accumulated deleterious mutations during their evolution, presumably owing to bottlenecks generated by their formation and to recombination suppression with their ancestral, coexisting counterparts.

Inversions with an accumulated mutational load are expected to incur a fitness cost. Indeed, when comparing survival among P genotypes from 1016 genotyped F2 progeny, homozygotes for a derived haplotype showed a far lower survival than other genotypes, with only 6.2% of Hn1/Hn1 larvae and 31.3 % of the Hn123/Hn123 larvae surviving to the adult stage (Figure 4A). By contrast, ancestral homozygotes Hn0/Hn0 had a good survival rate (77.6%), as well as all heterozygous haplotype combinations (Hn0/Hn1; Hn1/Hn123; Hn0/Hn123). Inversions therefore harbor fully recessive variants with a strong impact on individual survival in homozygotes. The high survival of Hn1/Hn123 genotypes may indicate that Hn1 and Hn123 harbor different deleterious variants within P_1 or that variants in P_2 or P_3 compensate for the deleterious effects of P_1.

These results bring key insights into why inversion polymorphisms could be maintained within populations. Inversions have largely been considered for their value in preserving combinations of co-adapted alleles through suppressed recombination, yet this also makes them prone to capturing deleterious mutations[28,29]. The inversions in *H. numata* allow the expression of dominant alleles with strongly positive effects on adult survival through wing-pattern mimicry[22], predicting their rapid fixation (Fig. 4B). Yet these inversions are also enriched in recessive deleterious variants, both captured from standing deleterious variation when inversion formed, and accumulating secondarily owing to recombination suppression in heterozygotes. Upon

formation, an inversion instantly locks a single sizeable haplotype which, if favored by selection, captures and brings to high frequency numerous recessive deleterious variants that were segregating via mutation-selection-drift balance in populations[31]. This initial mutational load is silent during the early establishment of inversions but becomes expressed as they reach higher frequencies and form homozygotes, thus preventing their fixation. Our results show that in *H. numata,* dominant mimetic haplotypes (Hn1 and Hn123) carry a recessive mutational load; therefore they only express net benefits when heterozygous, and enjoy highest fitness when rare in the population, explaining their maintenance in a polymorphism (negative frequency-dependent selection[32]; Fig. 4B). Lethality in inversion homozygotes and recombination suppression in heterozygotes limit the potential for inverted haplotypes to purge harmful variants via recombination, further promoting the accumulation of deleterious mutations within inversions. Inversion polymorphism therefore results from antagonistic ecological and genetic effects: positive selection on dominant alleles when heterozygous, but negative frequency-dependent selection due to recessive effects on intrinsic viability. This interplay between local adaptation and enrichment in deleterious elements is likely to be a very general mechanism determining the dynamics of inversions in populations. Among others, it could notably be involved in the maintenance of the supergenes known from the white throated sparrow[4] and fire ants[7], whose inversion polymorphisms are maintained despite a strong lethality of homozygotes.

Individuals carrying inversions at the P locus express disassortative mate preferences[18], which also balance inversion frequencies in the population (Fig. 4B) and participate to the strong deviation from Hardy-Weinberg equilibrium (lack of inversion homozygous) observed in natural populations[18]. Disassortative mating is likely to have evolved in response to the fitness costs associated with homozygous inversions, as selection may have favored mate preferences

lowering the likelihood of siring homozygous offspring[4,33]. The initial capture of genetic load in the inversions thus triggered cascading ecological effects and led to the long-term persistence of polymorphism. The low recombination regime associated with inversions also favored the insertion of transposons, increasing the size of inverted haplotypes. A similar pattern has also been observed in the Papaya neo sex-chromosomes[34] and in the fire ant supergene[35], indicating that this initial increase in size may be a general pattern in the early evolution of polymorphic chromosomes.

Our findings shed new light on the origin and evolution of complex polymorphisms controlled by supergenes and related architectures, such as sex-chromosomes. The benefits of structural variants in terms of recombination suppression between ecologically adaptive traits may explain why they are initially favored, whereas their maintenance as polymorphisms may be driven by another consequence of recombination suppression, namely their initial and gradually accumulating mutation load. These novel insights into the consequences of chromosomal inversions may explain why they are often polymorphic, overdominant, and linked with complex phenotypes in nature[2,4,12,36-38]. Besides ecological factors and mutation load, other selective forces such as meiotic drive[39], breakpoint effects[12] or position effects[40] may also balance the frequency of inversions in nature. A next challenge will be to identify the relative contribution of these processes in the maintenance of inversion polymorphisms. In a broader context, dissecting the opposing effects of suppressed recombination and how this determines the fate of chromosomal rearrangements may bring new light to our understanding of the variation in genome architecture across the tree of life[41].

Acknowledgments: We thank E. d'Alençon and M.-P. Dubois for their help in the lab, T. Aubier for DNA extraction, M. McClure, M. Tuatama, R. Mori-Pezo for their help during field

work, P. David, P. Nosil and R. Villoutreix for their careful and critical reading of the manuscript, K. Lhose, D. Laetsch, B. Nabholz, P.A. Gagnaire, M. Gautier, C. Lemaitre, F. Legeai, M. Elias and A.-S. Fiston-Lavier for insightful discussions. 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 and by fellowships from the Natural Sciences and Engineering Research Council of Canada, a Marie Sklodowska-Curie fellowship (FITINV, N 655857) and from an “Investissement d’Avenir” grant managed by Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01) to MC. VL was supported by the ANR grant DOMEVOL (ANR-JCJC-SVSE7-2013), HB by the Emergence program from Paris City council. This project benefited from the Montpellier Bioinformatics Biodiversity platform supported by the LabEx CeMEB, ANR "Investissements d’avenir" programme ANR-10-LABX-04-01. MGX (HP) acknowledges financial support from France Génomique National infrastructure, funded as part of ANR “Investissement d’avenir” programme ANR-10-INBS-09.

Author contributions: P.J., M.C., and M.J. designed the study. P.J., M.C., A.W., and M.J. wrote the paper. P.J., A.W., and M.J. generated the genomic data. M.C., H.B. and V.L. generated the RNAseq data. P.J. performed the genomic analyses with input from A.W.. M.C. managed butterfly rearing and performed fitness assays. P.J. performed selection coefficient analyses. H.P. performed whole genomes sequencing. All authors contributed to editing the manuscript.

Competing interests: Authors declare no competing interests.

References for main text:

1. Sturtevant, A. H. A Case of Rearrangement of Genes in Drosophila. *Proc Natl Acad Sci U S A* **7**, 235–237 (1921).
2. Dobzhansky, T. *Genetics of the Evolutionary Process: Columbia Classics edition*. 505 Pages (Columbia University Press, 1972).
3. Lamichhaney, S. *et al.* Structural genomic changes underlie alternative reproductive strategies in the ruff (Philomachus pugnax). *Nat. Genet.* **48**, 84–88 (2016).
4. Tuttle, E. M. *et al.* Divergence and Functional Degradation of a Sex Chromosome-like Supergene. *Current Biology* **26**, 344–350 (2016).
5. Kunte, K. *et al. doublesex* is a mimicry supergene. *Nature* **507**, 229–232 (2014).
6. Joron, M. *et al.* Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **477**, 203–206 (2011).
7. Wang, J. *et al.* A Y-like social chromosome causes alternative colony organization in fire ants. *Nature* **493**, 664–668 (2013).
8. Christie, M. R., McNickle, G. G., French, R. A. & Blouin, M. S. Life history variation is maintained by fitness trade-offs and negative frequency-dependent selection. *PNAS* **115**, 4441–4446 (2018).
9. Kess, T. *et al.* A migration-associated supergene reveals loss of biocomplexity in Atlantic cod. *Science Advances* **5**, eaav2461 (2019).
10. Stefansson, H. *et al.* A common inversion under selection in Europeans. *Nat. Genet.* **37**, 129–137 (2005).
11. Abbott, J. K., Nordén, A. K. & Hansson, B. Sex chromosome evolution: historical insights and future perspectives. *Proc. Biol. Sci.* **284**, (2017).
12. Küpper, C. *et al.* A supergene determines highly divergent male reproductive morphs in the ruff. *Nature Genetics* **48**, 79–83 (2016).
13. Li, J. *et al.* Genetic architecture and evolution of the *S* locus supergene in *Primula vulgaris*. *Nature Plants* **2**, 16188 (2016).
14. Fisher, R. A. *The genetical theory of natural selection*. (Clarendon Press, 1930). doi:10.5962/bhl.title.27468.
15. Charlesworth, D. & Charlesworth, B. Theoretical genetics of Batesian mimicry II. Evolution of supergenes. *J. Theor. Biol.* **55**, 305–324 (1975).
16. Ford, E. B. *Genetic Polymorphism*. (Faber & Faber: London, 1965).

17. Kopp, M. & Hermisson, J. The evolution of genetic architecture under frequency-dependent disruptive selection. *Evolution* **60**, 1537–1550 (2006).
18. Chouteau, M., Llaurens, V., Piron-Prunier, F. & Joron, M. Polymorphism at a mimicry supergene maintained by opposing frequency-dependent selection pressures. *PNAS* 201702482 (2017) doi:10.1073/pnas.1702482114.
19. Sinervo, B. & Lively, C. M. The rock–paper–scissors game and the evolution of alternative male strategies. *Nature* **380**, 240–243 (1996).
20. Giner-Delgado, C. *et al.* Evolutionary and functional impact of common polymorphic inversions in the human genome. *Nature Communications* **10**, 4222 (2019).
21. Mallet, J. & Barton, N. H. Strong natural selection in a warning-color hybrid zone. *Evolution* **43**, 421–431 (1989).
22. Chouteau, M., Arias, M. & Joron, M. Warning signals are under positive frequency-dependent selection in nature. *PNAS* **113**, 2164–2169 (2016).
23. Gordon, I. J. Polymorphism of the tropical butterfly, Danaus chrysippus L., in Africa. *Heredity* **53**, 583–593 (1984).
24. Majerus, M. E. N. *Ladybirds*. (Collins New Naturalist, 1994).
25. Nadeau, N. J. *et al.* The gene *cortex* controls mimicry and crypsis in butterflies and moths. *Nature* **534**, 106–110 (2016).
26. Jay, P. *et al.* Supergene Evolution Triggered by the Introgression of a Chromosomal Inversion. *Current Biology* **28**, 1839-1845.e3 (2018).
27. Kirkpatrick, M. How and Why Chromosome Inversions Evolve. *PLoS Biol* **8**, (2010).
28. Berdan, E. L., Blanckaert, A., Butlin, R. K. & Bank, C. Muller's Ratchet and the Long-Term Fate of Chromosomal Inversions. *bioRxiv* 606012 (2019) doi:10.1101/606012.
29. Faria, R., Johannesson, K., Butlin, R. K. & Westram, A. M. Evolving Inversions. *Trends in Ecology & Evolution* **34**, 239–248 (2019).
30. Stoletzki, N. & Eyre-Walker, A. Estimation of the neutrality index. *Mol. Biol. Evol.* **28**, 63–70 (2011).
31. Kirkpatrick, M. & Barton, N. Chromosome Inversions, Local Adaptation and Speciation. *Genetics* **173**, 419–434 (2006).

32. Llaurens, V., Whibley, A. & Joron, M. Genetic architecture and balancing selection: the life and death of differentiated variants. *Molecular Ecology* **26**, 2430–2448 (2017).
33. Maisonneuve, L., Chouteau, M., Joron, M. & Llaurens, V. Evolution and genetic architecture of disassortative mating at a locus under heterozygote advantage. *Evolution* **n/a** (doi: https://doi.org/10.1111/evo.14129), (2020).
34. Wang, J. *et al.* Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. *PNAS* **109**, 13710–13715 (2012).
35. Stolle, E. *et al.* Degenerative Expansion of a Young Supergene. *Mol Biol Evol* **36**, 553–561 (2019).
36. Lindtke, D. *et al.* Long-term balancing selection on chromosomal variants associated with crypsis in a stick insect. *Molecular Ecology* **26**, 6189–6205 (2017).
37. Knief, U. *et al.* A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nat Ecol Evol* **1**, 1177–1184 (2017).
38. Mérot, C., Llaurens, V., Normandeau, E., Bernatchez, L. & Wellenreuther, M. Balancing selection via life-history trade-offs maintains an inversion polymorphism in a seaweed fly. *Nat Commun* **11**, 1–11 (2020).
39. Helleu, Q. *et al.* Rapid evolution of a Y-chromosome heterochromatin protein underlies sex chromosome meiotic drive. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4110–4115 (2016).
40. Castermans, D. *et al.* Identification and characterization of the TRIP8 and REEP3 genes on chromosome 10q21.3 as novel candidate genes for autism. *European Journal of Human Genetics* **15**, 422–431 (2007).
41. Mérot, C., Oomen, R. A., Tigano, A. & Wellenreuther, M. A Roadmap for Understanding the Evolutionary Significance of Structural Genomic Variation. *Trends in Ecology & Evolution* **35**, 561–572 (2020).

Figure legends for main text :

Fig. 1. Genomic architecture of the *H. numata* wing pattern polymorphism.

a, Alignment of the genome assemblies from 4 *H. numata* morphs across the supergene region on chromosome 15. **b,** Sliding window PCA computed along the supergene. For clarity, only a subset of morphs are shown here (full dataset presented in Extended Data Figure 2 and Supplementary Table 2). Each colored line represents the variation in the position of a specimen on the first PCA axis along

chromosome 15 . Within the inversions, individual genomes are characterized by one of three genotypes : homozygous for the inversion (down), heterozygous (middle), homozygous for the standard arrangement (top). The gene annotation track is shown under the plot, with the forward strand in the lower panel and the reverse strand in the upper panel. Each gene is represented by a different colour. **c,** Structure of the *H. numata* supergene P. Three chromosome types are found in *H. numata* populations, carrying the ancestral gene order (Hn0), inversion P_1 (Hn1), or inversions P_1, P_2 and P_3 (Hn123). **d,** Analysis of divergence times between inverted and standard segments of Hn123 and Hn0. The TMRCA between Hn123 and Hn0 and the most ancient common ancestor of Hn123 provides the upper and lower bounds, respectively, of the time of inversions formation. Boxplots display the distribution of estimated times computed on 5kb sliding windows across the supergene (n=600 windows, estimates plotted along the supergene presented in Extended Data Figure 5). All samples homozygous for standard or inverted segments were used for divergence estimation. Boxplot elements: central line: median, box limits: 25th and 75th percentiles, whiskers: 1.5x interquartile range. Time intervals are consistent with the stepwise accretion of P1, P2 and P3, but the simultaneous origin of P2 and P3 cannot be formally rejected.

Fig. 2. Variation in inversion size due to accumulation of transposable elements.

a, Proportion of transposable elements in the whole genome (n=12), in the 3 inversions (n_{P1}=7, n_{P2}=6, n_{P3}=6), and in the regions present uniquely in inversion P_1, P_2 or P_3 and not in ancestral non-inverted haplotypes -i.e. sequences that were inserted in P_1, P_2, or P_3. The different genomic regions (whole genome, inversions and insertions in inversions) display differences in TE proportion (ANOVA F-value=58.135 and p-value=2.2e-16). Insertions in inversions compared to non-inverted segments are mostly transposable elements. Since the variance in TE proportions was unequal among the different genomics regions (levene test p-value = 0.008767), Games-Howell post-hoc tests were used to perform all pairwise comparisons between the different genomics regions. Statistical tests on TE proportions differences were computed considering the inversions separately or together (P_1+P_2+P_3, considering therefore the whole supergene) and similarly for the insertions in inversions compared to non-inverted segments (insertions in the different inversions separately or together). Stars above boxes are used to

display significant differences (p-value<0.0005) in TE proportions between the three major regions (whole genomes, P_1+P_2+P_3, and all insertions in inversion compared to non-inverted segments). Letters above boxes (a/b/c) are used to display significant differences (p-value<0.05) in TE proportions between all regions. Shared letters (e.g. ab vs a) indicate a lack of significant differences in TE proportions (e.g. P_1 vs whole genome) whereas absence of shared letters (e.g. b vs a) indicate a significant differences in TE proportions (e.g. P_3 vs whole genome). Details of statistical testing in Supplementary Table 7. Boxplot elements: central line: median, box limits: 25th and 75th percentiles, whiskers: 1.5x interquartile range. **b**, Time of activity for the distinct transposable elements found in inversions or only in sequences that were inserted in P_1, P_2, or P_3. Timing of insertion is estimated by computing the divergence between TEs and their respective consensus sequence genome-wide. This distribution provides an estimate of the time of activity of transposable elements but does not enable the insertion times of specific elements to be estimated. Consequently, patterns of TE accumulation in P_1, P_2 and P_3 cannot be accurately distinguished (Extended Data Figure 6). Recently active TEs (RC, DNA and LINE) are those that have accumulated within inversions. **c**, Size comparisons of orthologous standard and inverted chromosomal segments. Inverted haplotypes are longer than haplotypes with the ancestral gene order.

Fig. 3. Accumulation of deleterious variants in inversions.

a, Direction of selection and ratio of non-synonymous to synonymous polymorphisms (pN/pS), computed on 500 kb windows genome-wide and in the inversions segments, for both inverted and non-inverted haplotypes (dataset presented in Supplementary Table 3). Only genes with coding sequences >5kb (n=6364) were retained in this analysis. Inversions tend to be under negative selection and to accumulate non-synonymous polymorphism. **b,** Ratios of non-synonymous to synonymous substitutions (dN/dS) and polymorphisms (pN/pS) on the different mutations partitions observed in the P_1 segment: all mutations observed in Hn0 (purple), all mutations observed in Hn1/Hn123 (red), all mutations shared by *H. pardalinus* (*H. pard.*) and Hn1/Hn123 and not observed in Hn0 (blue) and all mutations present uniquely

in Hn1/Hn123 (yellow). Mutations are polarized by comparisons with the *H. melpomene* reference genome (*H. mel.*). Shades of blue are used to display 0.95 and 0.975 quantiles.

Fig. 4. Fitness variation associated with chromosomal inversions at the supergene in H. numata. **a,** Larval survival rates for the different supergene genotypes. Box plot display the distribution of larval survival in the different broods. Grey dots represent the survival rates in each brood. The number of broods and larvae analyzed for each genotype is indicated above the plot (dataset presented in Supplementary Table 4). Difference in survival between genotype was assessed with a generalized linear mixed models analysis followed by a Tukey's HSD post-hoc test. GLMM analysis confirmed that genotype was a significant predictor of survival ((χ^2 = 459.776; df = 5; p = 2.0e-16) while experimental cross design was unimportant (χ^2 = 0.8117; df = 2; p = 0.666), validating the joint analysis of all families and crosses. Individuals homozygote for the Hn1 haplotype displayed the lowest survival with survival frequency of 0.06 ± 0.04 ($p \leq 0.001$ when comparing to all other genotypes). On the other extreme of the survival spectrum Hn0 homozygotes and all heterozygotes (Hn123/Hn0, Hn1/Hn123, Hn1/Hn0) had a significantly higher survival rate ranging from 0.69 ± 0.06 to 0.83 ± 0.05 ($p \leq 0.001$ when comparing to Hn1/Hn1 and Hn123/Hn123). In between these two survival extremes, Hn123 homozygotes displayed a survival rate of 0.32 ± 0.06. Details of statistical testing in Supplementary Table 8. Boxplot elements: central line: median, box limits: 25th and 75th percentiles, whiskers: 1.5x interquartile range. **b,** Frequency dependent selection coefficients associated with supergene alleles. Colored dots indicate frequencies and selection coefficients measured in H. numata populations of Tarapoto, Peru. Curves indicate selection coefficients estimates according to allele frequencies. Three supergene alleles were considered, so selection coefficients for a given allele depend on the relative frequencies of the other two alleles. Colored areas around curves represent the distribution (quantile 0.1 to 0.9) of selection coefficients depending on these relative frequencies, and the curves themselves the distribution mean. Since selection coefficients are calculated relative to the population mean, they all tend toward 0 when allele frequencies tend toward 1.When selection coefficients are below 0 (grey area), haplotypes are

295 selected against, resulting in a decrease in haplotype frequency (and conversely for selection coefficients above 0).

Only methods:

Sampling and sequencing

To investigate the structure of the P supergene allele, we intercrossed wild-caught individuals in cages in order to obtain F2 (or later generation) autozygous individuals (i.e. with two identical copies of the supergene allele). Samples were either conserved in NaCl saturated DMSO solution at 20°C or snap frozen alive in liquid nitrogen and conserved at -80°C (Supplementary Table 1). DNA was extracted from the whole butterfly bodies except the head with a protocol adapted from Ref. [42], with the following modification. Butterflies were ground in a frozen mortar with liquid nitrogen, 150 mg of tissue powder was mixed with 900µl of preheated buffer and 6µl of RNaseA. Tube were incubated during 120 minutes at 50°C for lysis, and then at -10°C for 10 minutes, with the addition of 300µl of Potassium acetate for the precipitation. One volume of binding buffer was added with 100µl of Serapure beads solution. Three washing cycles were used and DNA was resuspended in 100µl of EB buffer. Samples 35 and 36 were prepared using the NEBNExt FFPE DNA Repair MIX (NEB) . DNA fragment shorter than 20Kb were removed for sample 35 and 36, and shorter than 40kb for samples 26 and 28. 10x Chromium linked-read libraries of 10 autozygous individuals corresponding to 8 different morphs, as well as 2 wild-caught homozygous individuals, were prepared and 2x150bp paired-end reads were sequenced using Illumina HiSeq 2500. Draft genomes (Supplementary Table 1) were assembled using Supernova v2.1.1[43]. In order to untangle the association between supergene genotype and wing pattern phenotype, PCR markers diagnostic for the distinct inversions[18] were amplified for 490 wild individuals from various locations (Supplementary Table 5). Briefly, amplicon size differences for a marker designed within the *H. numata* orthologue of HM00025 (cortex,

Genbank accension FP236845.2, Supplementary Table 6) enables discriminating between the distinct supergene arrangements (Hn1 (~1200bp), Hn123 (~800bp) and Hn0(~ 600 bp)) [18].

Whole genome assembly analysis

The assembled genomes were compared to the *H. melpomene* reference genome (v2.5, http://lepbase.org/) and to each other using BLAST v2.7.1+[44] and LAST v956[45]. Sequence variation among individuals allowed us to determine inversion breakpoint positions with 50-200 bp resolution. Because for some specimens, the supergene was dispersed across multiple scaffolds, we used Ragout2[46] to re-scaffold their supergene assembly, using as reference the four individual assemblies with the highest quality assembly statistics (n°38, 29, 40, and 26). Because sample n° 41 was heterozygous for inversion P1, P2 and P3, its genome was not considered for analyses focused on these regions (TE proportions and regions size). Genome quality analysis was assessed with BUSCO v3.0.2[47] using the insecta odb9 database. MAKER[48] (v2.31.10) was used to annotate the genomes, using protein sequences obtained from the *H. melpomene* genome (v2.5, http://lepbase.org/) in combination with an *H. numata* transcriptome dataset[49]. RepeatModeler v1.0.11[50] was used to identify unannotated TEs in the 12 *H. numata* genomes. Unknown repeat elements detected by RepeatModeler were compared by BLAST[44] (-evalue cut-off $1e^{-10}$) to a transposase database (Tpases080212) from Ref. [51]. Identified TEs were merged with the *Heliconius* repeat database[52] and redundancy was filtered using CDHIT v4.8.1[53] with a 80 % identity threshold. RepeatMasker v1.332[50] was then used to annotate transposable elements and repeats using this combined database and results were parsed with scripts from https://github.com/4ureliek/Parsing-RepeatMasker-Outputs.git. TE sequence divergence was converted into divergence time considering a substitutions rate of 1.9×10^{-9} substitutions/ site/generation and 4 generations/year[54].

345

Population Genomic Analysis.

Whole genome re-sequence data from *H. numata* and other *Heliconius* species from Ref. [26] were used, as well as *37* new wild-caught *H. numata* specimens (Supplementary Table 2). For the latter samples, butterfly bodies were conserved in NaCl saturated DMSO solution at -20°C and
350 DNA was extracted using QIAGEN DNeasy blood and tissue kits according to the manufacturer's instructions with RNase treatment. Illumina Truseq paired-end whole genome libraries were prepared and 2x100bp reads were sequenced on the Illumina HiSeq 2000 platform. Reads were mapped to the *H. melpomene* Hmel2 reference genome[55] using Stampy v1.0.28[56] with default settings except for the substitution rate which was set to 0.05 to allow for expected
355 divergence from the reference. Alignment file manipulations were performed using SAMtools v0.1.3[57]. After mapping, duplicate reads were excluded using the MarkDuplicates tool in Picard v1.1125 (http://broadinstitute.github.io/picard) and local indel realignment using IndelRealigner was performed with GATK v3.5[58]. Invariant and polymorphic sites were called with GATK HaplotypeCaller, with options --min base quality score 25 --min mapping quality score 25 -stand
360 emit conf 20 --heterozygosity 0.015.

VCF data were processed using bcftools v1.9-20-g34a5b0d[59]. Genotype at the supergene was assessed using PCA computed with the SNPRelate R package[60] (v1.24.0), using 5kb windows for sliding window analyses and only biallelic sites. This simple method based on the structure of nucleotide variation proved powerful and efficient to detect polymorphic rearrangements without
365 prior knowledge. Using Phylobayes v2.3[61], on 5kb sliding windows, we estimated 1) the most recent coalescence event between *Hn0+Hn1* and *Hn123*, which corresponds to age of the last recombination between Hn0+Hn1 and Hn123, and 2) the time to the most recent common

ancestor (TMRCA) of all Hn123 haplotypes. This provides respectively the upper (1) and the lower (2) bounds of the date of the inversion event (Figure 1D, Extended Data Figure 5).

In order to compute the Fst and standard population genetic analyses, we manually curated the phasing of heterozygous individuals since computational phasing packages such as SHAPEIT or BEAGLE were found to introduce frequent phase switch errors. For each heterozygous SNP in inversion regions, if one and only one of the two alleles is observed in more than 80 % of individuals without inversions (Hn0), this allele is considered being on the haplotype 1, the other being on haplotype 2. For SNPs which did not fit this criterion, each allele was placed randomly on one of the two haplotypes.

RNAseq analysis

RNAseq data from Ref. [49] were reanalysed using the EdgeR R package[63] (v3.16.5). Gene expression in wing discs from *silvana* individuals (Hn0/Hn0) was compared to gene expression in *tarapotensis* and *aurora* individuals (Hn123/Hn123). Gene ontology (GO) annotation was performed with GO FEAT[64] using default parameters. Enrichment in GO for differentially expressed genes was assessed with Goseq v1.26.0[65] using the Wallenius approximation and a false discovery rate cutoff set at 0.05.

Deleterious mutation accumulation

SnpEff v4.3t[66] with default parameters was used to annotate the *H. numata* SNPs using the *H. melpomene* Hmel2 reference genome annotation. We computed the ratio of synonymous to non-synonymous variants (pN/pS), the rate of synonymous to non-synonymous substitution (dN/dS) compared to *H. melpomene*, and the direction of selection with DoS = Dn/(Dn + Ds) − Pn/(Pn +

*P*s)[30], using all individuals, or only those homozygous for a given inversion type, for every gene larger than 5kb (to ensure each gene comprises several SNPs). Whole genome distribution was computed on 500kb non-overlapping sliding windows.

395

Fitness Assay

H. numata specimens used for the fitness analyses originated from the Tarapoto valley, San Martin, Peru. Broods were designed to control for genome-wide inbreeding depression and are illustrated in Supplementary Figure 1. Raw results are presented in Supplementary Table 4. First,
400 P heterozygous F_1 butterflies were generated by crossing wild F_0 males to captive bred virgin females. Unrelated F_1 male-female pairs were then selected for their *P* genotype and hand paired to generate an F_2 progeny. We specifically designed these crosses to generate an F_2 progeny containing both homozygotes and heterozygotes, within a single family. Larvae were monitored twice a day to assess survival or mortality. Upon death or butterfly emergence, individuals were
405 stored in 96% ethanol until genotyping. We generated a total of 486 F_2 progeny from 6 independent replicate broods for the Hn0/Hn1 x Hn0/Hn1 cross, 504 F2 progeny from 6 broods for the Hn1/Hn123 x Hn1/Hn123 cross and 454 F2 progeny from 7 broods from the Hn1/Hn123 x Hn0/Hn1 cross. 1,016 F2 progeny could be genotyped using a methodology from [18] (see Sampling and Sequencing Method section).

410

Frequency dependent selection coefficient

Frequency dependent selection coefficient associated with supergene alleles were calculated based on adult survival data (predation) from [22], on mating success data from ref. [18] and on larval survival data from this study, all three estimated from the populations around Tarapoto, Peru.
415 Adult survival estimates were based on a probability of predation attempt during a 30 days

period and one mating event was considered. For adult survival and mate choice, we assumed perfect dominance among haplotypes, such as Hn1 dominant to Hn123 and Hn0, and Hn123 dominant to Hn0, as suggested by Ref. [67] and PCR analyses (Supplementary Table 5). Only morphs *silvana*, *bicoloratus* and *tarapotensis* were used, encoded by haplotypes Hn0, Hn1, and Hn123, respectively. To reflect the incomplete association between haplotype and phenotype (Fig. 1, Extended Data Figure 2 and Supplementary Table 5), simulations were also run with variation in the level of genotype-phenotype association (Supplementary Figure 9). We assumed mate choice was based only on phenotype and not on genotype (e.g. Hn1/Hn0 and Hn1/Hn1 individuals perform equally). The three haplotypes Hn0, Hn1 and Hn123 were analysed and their genotype frequencies were considered to be at Hardy-Weinberg equilibrium. This hypothesis allowed us to visualise how the different selective forces generate departure from HW equilibrium. Three supergene haplotypes were considered, so selection coefficients for a given haplotype depended on the relative frequencies of the other two haplotypes. To estimate selection on a given haplotype, the relative frequencies of the other two haplotypes were allowed to vary between 99:1 and 1:99. For instance, when calculating selection coefficients for Hn0 when its frequency was 0.10 (f_{Hn0}=0.10), Hn1 and Hn123 frequency (respectively f_{Hn1} and f_{Hn123}) ranged from 0.891 (=0.99*(1-f_{Hn0})) and 0.009 (=0.01*(1-f_{Hn0})), with $f_{Hn0} + f_{Hn1} + f_{Hn123}$=1. Colored areas around curves in Figure 4B represent the distribution (quantile 0.1 to 0.9) of these different selection coefficients depending on the relative frequencies of the other two alleles. Selection coefficients were calculated relative to the population mean.

Statistics and reproducibility

To compare genotype survival between families and crosses (Figure 4a) we performed generalized linear mixed models analysis followed by a *Tukey's* HSD post-hoc test (package

"lme4" [68]; in R version 3.1.3[69]), with the survival of an individual with a given genotype as the response variable (binomial response with logit link). The significance of the predictors was tested using likelihood ratio tests. The genotype was a covariate predictor, crosses was a fixed effect and family identity a random effect to control for non-independence of measures. ANOVA was used to test for differences in TE proportions in different genomic regions with 6 degrees of freedom. Equality of variance was assesses with a Levene's test. Pairwise comparisons among genomic regions was performed with a Games-Howell post-hoc test. Block Jackknife resampling was used to test for inversion enrichment in differentially expressed genes, using 100 blocks of 400 transcripts. Plots were created with ggplot2[70].

Data availability:

The raw sequence data were deposited in NCBI SRA and accession numbers are indicated in Supplementary table 2. The whole genome VCF file is available upon request. Whole genome assemblies were deposited in NCBI under accession number PRJNA676017. All data underlying the fitness assays are available in Supplementary Table 4.

Code Availability:

Repeat masker results were parsed with scripts from https://github.com/4ureliek/Parsing-RepeatMasker-Outputs.git. Scripts used to compute the main analyses of this study are available at https://github.com/PaulYannJay/Mutation-load-analysis

460 # Methods-only references:

42. Mayjonade, B. *et al.* Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. *BioTechniques* **61**, 203–205 (2016).
43. Weisenfeld, N. I., Kumar, V., Shah, P., Church, D. M. & Jaffe, D. B. Direct determination of diploid genome sequences. *Genome Res.* **27**, 757–767 (2017).
44. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
45. Kiełbasa, S. M., Wan, R., Sato, K., Horton, P. & Frith, M. C. Adaptive seeds tame genomic sequence comparison. *Genome Res.* **21**, 487–493 (2011).
46. Kolmogorov, M. *et al.* Chromosome assembly of large and complex genomes using multiple references. *Genome Res.* (2018) doi:10.1101/gr.236273.118.
47. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
48. Cantarel, B. L. *et al.* MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res* **18**, 188–196 (2008).
49. Saenko, S. V. *et al.* Unravelling the genes forming the wing pattern supergene in the polymorphic butterfly Heliconius numata. *Evodevo* **10**, 16 (2019).
50. Smit, AFA, Hubley, R & Green, P. RepeatMasker Open-4.0. 2013-2015.
51. Campbell, M. S. *et al.* MAKER-P: a tool kit for the rapid creation, management, and quality control of plant genome annotations. *Plant Physiol.* **164**, 513–524 (2014).
52. Lavoie, C. A., Platt, R. N., Novick, P. A., Counterman, B. A. & Ray, D. A. Transposable element evolution in Heliconius suggests genome diversity within Lepidoptera. *Mob DNA* **4**, 21 (2013).
53. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
54. Martin, S. H. *et al.* Natural Selection and Genetic Diversity in the Butterfly Heliconius melpomene. *Genetics* **203**, 525–541 (2016).

55. Davey, J. W. *et al.* 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).
56. Lunter, G. & Goodson, M. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res.* **21**, 936–939 (2011).
57. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
58. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491–498 (2011).
59. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
60. Zheng, X. *et al.* A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328 (2012).
61. Lartillot, N. & Philippe, H. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* **21**, 1095–1109 (2004).
62. Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E. & Lercher, M. J. PopGenome: An Efficient Swiss Army Knife for Population Genomic Analyses in R. *Mol Biol Evol* **31**, 1929–1936 (2014).
63. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
64. Araujo, F. A., Barh, D., Silva, A., Guimarães, L. & Ramos, R. T. J. GO FEAT: a rapid web-based functional annotation tool for genomic and transcriptomic data. *Sci Rep* **8**, (2018).
65. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol* **11**, R14 (2010).
66. Cingolani, P. *et al.* 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).
67. Le Poul, Y. *et al.* Evolution of dominance mechanisms at a butterfly mimicry supergene. *Nature Communications* **5**, 5644 (2014).
68. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**, 1–48 (2015).
69. R Core Team. R: A language and environment for statistical computing. (2013).

70. Wickham., H. *ggplot2: Elegant Graphics for Data Analysis*. (2016).