Title
Butterfly genome reveals promiscuous exchange of mimicry adaptations among species

Journal
Nature, 487(7405)

ISSN
0028-0836

Authors
Dasmahapatra, Kanchon K
Walters, James R
Briscoe, Adriana D
et al.

Publication Date
2012-07-01

DOI
10.1038/nature11041

Peer reviewed

Nature. Author manuscript; available in PMC 2013 January 05.

Published in final edited form as:
Nature. 2012 July 5; 487(7405): 94–98. doi:10.1038/nature11041.

Butterfly genome reveals promiscuous exchange of mimicry adaptations among species

The *Heliconius* Genome Consortium, **Kanchon K Dasmahapatra**[1,*], **James R. Walters**[2,*], **Adriana D. Briscoe**[3], **John W. Davey**[4], **Annabel Whibley**[5], **Nicola J. Nadeau**[2], **Aleksey V. Zimin**[6], **Daniel S. T. Hughes**[7], **Laura C. Ferguson**[8], **Simon H. Martin**[2], **Camilo Salazar**[2,9], **James J. Lewis**[3], **Sebastian Adler**[10], **Seung-Joon Ahn**[11], **Dean A. Baker**[12], **Simon W. Baxter**[2], **Nicola L. Chamberlain**[13], **Ritika Chauhan**[14], **Brian A. Counterman**[15], **Tamas Dalmay**[16], **Lawrence E. Gilbert**[17], **Karl Gordon**[18], **David G. Heckel**[11], **Heather M. Hines**[19], **Katharina J. Hoff**[10], **Peter W.H. Holland**[8], **Emmanuelle Jacquin-Joly**[20], **Francis M. Jiggins**[21], **Robert T. Jones**[5], **Durrell D. Kapan**[22,23], **Paul Kersey**[7], **Gerardo Lamas**[24], **Daniel Lawson**[7], **Daniel Mapleson**[25], **Luana S. Maroja**[26], **Arnaud Martin**[3], **Simon Moxon**[27], **William J. Palmer**[21], **Riccardo Papa**[28], **Alexie Papanicolaou**[18], **Yannick Pauchet**[11], **David A. Ray**[29,30], **Neil Rosser**[1], **Steven L. Salzberg**[31], **Megan A. Supple**[32], **Alison Surridge**[2], **Ayse Tenger-Trolander**[13], **Heiko Vogel**[11], **Paul A. Wilkinson**[33], **Derek Wilson**[7], **James A. Yorke**[6], **Furong Yuan**[3], **Alexi L. Balmuth**[34], **Cathlene Eland**[34], **Karim Gharbi**[34], **Marian Thomson**[34], **Richard A. Gibbs**[35], **Yi Han**[35], **Joy C. Jayaseelan**[35], **Christie Kovar**[35], **Tittu Mathew**[35], **Donna M. Muzny**[35], **Fiona Ongeri**[35], **Ling-Ling Pu**[35], **Jiaxin Qu**[35], **Rebecca L. Thornton**[35], **Kim C. Worley**[35], **Yuan-Qing Wu**[35], **Mauricio Linares**[36], **Mark L. Blaxter**[4,34], **Richard H. ffrench Constant**[14], **Mathieu Joron**[5], **Marcus R. Kronforst**[13], **Sean P. Mullen**[37], **Robert D. Reed**[3], **Steven E. Scherer**[35], **Stephen Richards**[35], **James Mallet**[1,38,§], **W. Owen McMillan**[9], and **Chris D. Jiggins**[2,9]

[1]Department of Genetics, Evolution and Environment, University College London, Gower Street, London, WC1E 6BT, UK.

[2]Department of Zoology, Downing Street, University of Cambridge, Cambridge, CB2 3EJ, UK.

[3]Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA.

[4]Institute of Evolutionary Biology, Ashworth Laboratories, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JT, UK.

[5]CNRS UMR 7205, Muséum National d'Histoire Naturelle, 45 rue Buffon, Paris 75005, France.

[6]Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA.

[7]European Bioinformatics Institute, Hinxton, CB10 1SD, UK.

[8]Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3PS, UK.

[9]Smithsonian Tropical Research Institute, Naos Island, Causeway Amador, Panamá, Republic de Panamá.

[10]Institut für Mathematik und Informatik, Universität Greifswald, 17487 Greifswald, Germany.

[11]Max Planck Institute for Chemical Ecology, D-07745 Jena, Germany.

[12]Ecology and Evolution, Imperial College London, London, UK.

[13]FAS Center for Systems Biology, Harvard University, Cambridge, MA 02138, USA.

[14]Centre for Ecology and Conservation, School of Biosciences, University of Exeter, Penryn, TR10 9EZ, UK.

Europe PMC Funders Author Manuscripts

[15]Department of Biology, Mississippi State University, Mississippi State, Mississippi 39762, USA.

[16]School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK.

[17]Section of Integrative Biology and Brackenridge Field Laboratory, University of Texas, Austin, TX 78712, USA.

[18]Black Mountain Laboratories, CSIRO Ecosystem Sciences, Clunies Ross Street, Canberra, ACT 2601, Australia.

[19]Department of Genetics, North Carolina State University, Raleigh, NC 27695, USA.

[20]Inra, UMR-A 1272 Inra-Université Pierre et Marie Curie, Physiologie de l'Insecte: Signalisation et Communication, Route de Saint-Cyr, Versailles, France.

[21]Department of Genetics, Downing Street, University of Cambridge, Cambridge, CB2 3EH, UK.

[22]Department of Entomology, Center for Comparative Genomics, California Academy of Sciences, 55 Music Concourse Drive, San Francisco, California 94118, USA.

[23]Center for Conservation and Research Training, Pacific Biosciences Research Center, University of Hawaii at Manoa, 3050 Maile Way, Gilmore 406, Honolulu, HI 96822, USA.

[24]Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Apartado 14-0434, Lima, Peru.

[25]School of Computing Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK.

[26]Department of Biology, Williams College, Williamstown, MA 01267, USA.

[27]Department of Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520, USA.

[28]Department of Biology and Center for Applied Tropical Ecology and Conservation, University of Puerto Rico, Rio Piedras, Puerto Rico.

[29]Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology, Mississippi State University, Mississippi State, Mississippi 39762, USA.

[30]Institute for Genomics, Biocomputing and Biotechnology, Mississippi State University, Mississippi State, Mississippi 39759, USA.

[31]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD 21205, USA.

[32]Biomathematics Program, North Carolina State University, Raleigh, NC 27695, USA.

[33]School of Biological Sciences, University of Bristol, Bristol, BS8 1UG, UK.

[34]The GenePool, Ashworth Laboratories, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JT, UK.

[35]Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA.

[36]Facultad de Ciencias Naturales y Matemáticas, Universidad del Rosario and Instituto de Genética, Universidad de los Andes, Bogotá, Colombia.

[37]Department of Biology, Boston University, 5 Cummington Street, Boston, MA 02215, USA.

[38]Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA.

Abstract

The evolutionary importance of hybridization and introgression has long been debated[1]. We used genomic tools to investigate introgression in *Heliconius*, a rapidly radiating genus of neotropical butterflies widely used in studies of ecology, behaviour, mimicry and speciation[2-5] . We sequenced the genome of *Heliconius melpomene* and compared it with other taxa to investigate chromosomal evolution in Lepidoptera and gene flow among multiple *Heliconius* species and races. Among 12,657 predicted genes for *Heliconius*, biologically important expansions of families of chemosensory and *Hox* genes are particularly noteworthy. Chromosomal organisation has remained broadly conserved since the Cretaceous, when butterflies split from the silkmoth lineage. Using genomic resequencing, we show hybrid exchange of genes between three co-mimics, *H. melpomene*, *H. timareta*, and *H. elevatus*, especially at two genomic regions that control mimicry pattern. Closely related *Heliconius* species clearly exchange protective colour pattern genes promiscuously, implying a major role for hybridization in adaptive radiation.

The butterfly genus *Heliconius* (Nymphalidae: Heliconiinae) is associated with a suite of derived life-history and ecological traits, including pollen-feeding, extended life-span, augmented ultraviolet colour vision, 'trap-lining' foraging behavior, gregarious roosting and complex mating behaviours, and provides outstanding opportunities for genomic studies of adaptive radiation and speciation[4, 6]. The genus is best known for the hundreds of different colour pattern races seen among its 43 species, with repeated examples of both convergent evolution among distantly related species and divergent evolution between closely related taxa[3]. Geographic mosaics of multiple colour pattern races, such as in *Heliconius melpomene* (Fig. 1), converge to similar mosaics in other species, and this led to the hypothesis of mimicry[2]. *Heliconius* are unpalatable and Müllerian mimicry of warning colour patterns enables species to share the cost of educating predators[3]. Divergence in wing pattern is also associated with speciation and adaptive radiation due to a dual role in mimicry and mate selection[3, 5]. A particularly recent radiation is the *melpomene*-silvaniform clade, where mimetic patterns often appear polyphyletic (Fig. 1a). Most species in this clade occasionally hybridise in the wild with other clade members[7]. Gene genealogies at a small number of loci indicate introgression between species[8], and one non-mimetic species, *H. heurippa*, has a hybrid origin[9]. Adaptive introgression of mimicry loci is therefore a plausible explanation for parallel evolution of multiple mimetic patterns in the *melpomene*-silvaniform clade.

A *Heliconius melpomene melpomene* stock from Darién, Panama (Fig. 1) was inbred via five generations of sib mating. A single male was sequenced to 38x coverage (after quality filtering) using combined 454 and Illumina technologies (Supplementary Information 1-8). The complete draft genome assembly of 269 Mb consists of 3,807 scaffolds with an N50 of 277 kb and contains 12,657 predicted protein-coding genes. RAD linkage mapping was used to assign and order 83% of the sequenced genome onto the 21 chromosomes (Supplementary Information 4). These data permit a considerably improved genome-wide chromosomal synteny comparison with the silkmoth *Bombyx mori*[10, 11]. Using 6,010 orthologues identified between *H. melpomene* and *B. mori* we found that 11 of 21 *H. melpomene* linkage groups show homology to single *B. mori* chromosomes and ten linkage groups have major contributions from two *B. mori* chromosomes (Fig. 2a and Supplementary Information 8), revealing several previously unidentified chromosomal fusions. These fusions on the *Heliconius* lineage most likely occurred after divergence from the sister genus *Eueides*[4], which has the lepidopteran modal karyotype of n=31[12]. Three chromosomal fusions are evident in *Bombyx* (Fig. 2a, *B. mori* chromosomes 11, 23 and 24), as required for evolution of the *Bombyx* n=28 karyotype from the ancestral n=31 karyotype. *Heliconius* and *Bombyx* lineages diverged in the Cretaceous >100 MYA[11], so the chromosomal structures of Lepidoptera genomes have remained highly conserved compared

to those of flies or vertebrates[13, 14]. In contrast, small-scale rearrangements were frequent. In the comparison with *Bombyx*, we estimate 0.05-0.13 breaks/Mb/MY, and with the Monarch butterfly, *Danaus plexippus*, 0.04-0.29 breaks/Mb/MY. Although lower than previously suggested for Lepidoptera[15], these rates are comparable to *Drosophila* (Supplementary Information 8).

The origin of butterflies was associated with a switch from nocturnal to diurnal behaviour, and a corresponding increase in visual communication[16]. *Heliconius* have increased visual complexity through expression of a duplicate UV opsin[6], in addition to the long wavelength, blue, and UV-sensitive opsins in *Bombyx*. We might therefore predict reduced complexity of olfactory genes, but in fact *Heliconius* and *Danaus*[17] genomes have more chemosensory proteins (CSPs) than any other insect genome: 33 and 34 CSPs respectively (Supplementary Information 9), versus 24 in *Bombyx* and 3- 4 in *Drosophila*[18]. Lineage-specific CSP expansions were evident in both *Danaus* and *Heliconius* (Fig. 2b). In contrast, all three lepidopteran genomes possess similar numbers of odorant binding proteins and olfactory receptors (Supplementary Information 9). *Hox* genes are involved in body plan development and show strong conservation across animals. We identified four additional *Hox* genes located between the canonical *Hox* genes *pb* and *zen*, orthologous to *shx* genes in *B. mori* (Supplementary Information 10)[19]. These *Hox* gene duplications in the butterflies and *Bombyx* share a common origin, and are independent of the two tandem duplications known in dipterans (*zen2*, *bcd*). Immunity-related gene families are similar across all three lepidopterans (Supplementary Information 11), contrasting with extensive duplications and losses within dipterans[20].

The *Heliconius* reference genome enabled rigorous tests for introgression among *melpomene*-silvaniform clade species. We used RAD resequencing to reconstruct a robust phylogenetic tree based on 84 individuals of *H. melpomene* and its relatives,sampling on average 12 Mb, or 4% of the genome (Fig 1a, Supplementary Information 12, 13, 18). We then tested for introgression between the sympatric co-mimetic postman races of *H. melpomene aglaope* and *H. timareta* ssp. nov. (Fig. 1) in Peru, employing 'ABBA-BABA' single nucleotide sites and Patterson's *D*-statistics (Fig. 3a), originally developed to test for admixture between Neanderthals and modern humans[21, 22] (Supplementary Information 12). Genome-wide we found an excess of ABBA sites, giving a significantly positive Patterson's $D = 0.037 \pm 0.003$ (two tailed *Z*-test for $D = 0$, $P = 1 \times 10^{-40}$), indicating greater genome-wide introgression between the sympatric mimetic taxa *H. m. amaryllis* and *H. timareta* ssp. nov., than between *H. m. aglaope* and *H. timareta* ssp. nov., which do not overlap spatially (Fig. 1b). These *D*-statistics yield an estimate of 2-5% of the genome exchanged[21] between the two taxa (Supplementary Information 12). Eleven of the 21 chromosomes have significantly positive *D*-statistics (Fig. 3b,); interestingly, the strongest signals of introgressions were found on two chromosomes containing the known mimicry loci *B*/*D* and *N*/*Yb* (Fig. 3b, Supplementary Information 15).

Perhaps the best known case of Müllerian mimicry is the geographic mosaic of ~30 bold postman and rayed colour pattern races of *H. melpomene* (Fig. 1b, Supplementary Information 22), which mimic a near-identical colour pattern mosaic in *H. erato* (Fig. 1a), among other *Heliconius*. Mimicry variation is generally controlled by a few loci with major effects. Mimetic pattern differences between the postman *H. melpomene amaryllis* and rayed *H. melpomene aglaope* races studied here (Fig 1a) are controlled by the *B*/*D* (red pattern) and *N*/*Yb* (yellow pattern) loci[23, 24]. These loci are located on the same two chromosomes showing the strongest *D*- statistics in our RAD analysis (Fig. 3b). To test whether mimicry loci might be introgressed between co-mimetic *H. timareta* and *H. melpomene* (Fig. 1a)[7], we resequenced the colour pattern regions *B*/*D* (0.7 Mb) and *N*/*Yb* (1.2 Mb), and 1.8 Mb of unlinked regions across the genome, from both postman and ray-

patterned *H. melpomene* and *H. timareta* from Peru and Colombia, and six silvaniform outgroup taxa (Fig. 1a, Supplementary Information 12). To test for introgression at the *B/D* mimicry locus we compared rayed *H. m. aglaope* and postman *H. m. amaryllis* as the ingroup with postman *H. timareta* ssp. nov. (as in Fig. 3a) and found large, significant peaks of shared fixed ABBA nucleotide sites combined with an almost complete lack of BABA sites (Fig. 4b). This provides evidence that blocks of shared sequence variation in the *B*/*D* region were exchanged between postman *H. timareta* and postman *H. melpomene*, in the genomic region known to determine red mimicry patterns between races of *H. melpomene*[23, 24] (Fig. 4a).

For a reciprocal test, we used the same *H. melpomene* races as the ingroup to compare with rayed *H. timareta florencia* at the *B*/*D* region. In this case, correspondingly large and significant peaks of BABA nucleotide sites are accompanied by virtual absence of ABBA sites (Fig. 4c) indicating that variation at the same mimicry locus was also shared between rayed *H. timareta* and rayed *H. melpomene*. Equivalent results in the *N*/*Yb* colour pattern region, controlling yellow colour pattern differences, are in the expected directions for introgression and highly significant for the test using postman *H. timareta* ssp. nov. ($P = 6 \times 10^{-34}$), although not significant with rayed *H. timareta florencia* ($P = 0.13$, Supplementary Information 17). In contrast hardly any ABBA or BABA sites are present in either comparison across 1.8 Mb in 55 genomic scaffolds unlinked to the colour pattern regions (Supplementary Information 21). These concordant, but reciprocal patterns, where fixed ABBA and BABA substitutions occur almost exclusively within large genomic blocks at two different colour pattern loci (449 and 99 sites for *B*/*D* and *N*/*Yb* respectively, Figs. 4b,c and Supplementary Information 17) would be very hard to explain via convergent functional site evolution or under coalescent fluctuations. Instead, our results imply that derived colour pattern elements have introgressed recently between both rayed and postman forms of *H. timareta* and *H. melpomene*.

To test whether colour pattern loci might be shared more broadly across the clade, we used sliding-window phylogenetic analyses along the colour pattern regions. For regions flanking and unlinked to colour pattern loci, tree topologies are similar to the overriding signal recovered from the genome as a whole (Supplementary Information 18). Races of *H. melpomene* and *H. timareta* each form separate monophyletic sister groups and both are separated from the more distantly related silvaniform species (Fig. 4d). By contrast, within the region of peak ABBA/BABA differences, the topologies switch dramatically. Races of *H. melpomene* and *H. timareta* group according to wing pattern, while the species themselves become polyphyletic (Figs. 4e,f, Supplementary Information 19, 20). Remarkably, the rayed *H. elevatus*, a member of the silvaniform clade according to genome average relationships (Fig. 1a, Supplementary Information 18), groups with rayed races of unrelated *H. melpomene* and *H. timareta* in small sections within both *B*/*D* and *N*/*Yb* colour pattern loci (Fig. 4e, Supplementary Information 19, 20). These results are again most readily explained by introgression and fixation of mimicry genes.

We have developed a *de novo* reference genome sequence that will facilitate evolutionary and ecological studies in this key group of butterflies. We have demonstrated repeated exchange of small (~100 kb) adaptive genome regions among multiple species in an adaptive radiation. Our genome-scale analysis provides considerably greater power than previous tests of introgression [8, 25, 26]. As with *H. heurippa*[9], our evidence suggests that *H. elevatus* was formed during a hybrid speciation event. The main genomic signal from this rayed species places it closest to *H. pardalinus butleri* (Fig. 1a), but colour pattern genomic regions resemble those of rayed races of *H. melpomene* (Fig. 4e and Supplementary Information 18-20). Colour pattern is important in mating behaviour in *Heliconius*[5], and the transfer of mimetic pattern may have enabled the divergent sibling species *H. elevatus* to

coexist with *H. pardalinus* across the Amazon. Although it was long suspected that introgression might be important in adaptive radiation[1], our results from the most diverse terrestrial biome on the planet suggest that adaptive introgression is more pervasive than previously realized.

Methods summary

A full description of methods can be found in the Supplementary Information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Footnotes

[§]Correspondence and requests for materials should be addressed to J.M. (jmallet@oeb.harvard.edu).

[*]These authors contributed equally to this work.

Consortium leaders:
Chris D. Jiggins, W. Owen McMillan

Heliconius Genome Consortium PIs:
Richard ffrench-Constant, Marcus Kronforst, Mathieu Joron, James Mallet, Sean Mullen, Robert Reed, Mark Blaxter, Larry Gilbert, Mauricio Linares, Gerardo Lamas

Introgression study leader & corresponding author:
James Mallet

Lead investigators:
Kanchon Dasmahapatra, James Walters, Nicola Nadeau, Annabel Whibley, John Davey, Adriana Briscoe, Laura Ferguson, Daniel Hughes, Simon Martin, Camilo Salazar, James Lewis

Sequencing:
Stephen Richards, Steve Scherer, Alexi Balmuth, Marian Thomson, Karim Gharbi, Cathlene Eland, Mark Blaxter, Richard Gibbs, Yi Han, Joy Christina Jayaseelan, Christie Kovar, Tittu Mathew, Donna Marie Muzny, Fiona Ongeri, Ling-Ling Pu, Jiaxin Qu, Rebecca Lynn Thornton, Kim C. Worley, Yuan-Qing Wu

Assembly:
Aleksey Zimin, James Yorke, Steven Salzberg, Alexie Papanicolaou, Karl Gordon

RAD map and assembly verification:
John Davey, Simon Baxter, Mark Blaxter, Luana Maroja, Durrell D. Kapan, James Walters, Paul Wilkinson

Geographic distribution map:
Neil Rosser

Annotation:
James Walters, Daniel Hughes, Derek Wilson, Daniel Lawson, Katharina Hoff, Sebastian Adler, Paul Wilkinson

Genome browser and databases:
Daniel Hughes, James Lewis

Manual annotation and evolutionary analyses:
Olfactory proteins: Adriana Briscoe, Emmanuelle Jacquin-Joly, Furong Yuan
Hox genes: Laura Ferguson, Peter W. H. Holland, James Walters
Micro-RNAs: Alison Surridge, Tamas Dalmay, Daniel Mapleson, Simon Moxon
Immune genes: William Palmer, Francis Jiggins
P450 genes: Robert Jones and Ritika Chauhan
UGT genes: Heiko Vogel, Seung-Joon Ahn, David Heckel
Ribosomal Proteins: Yannick Pauchet

Manual annotation group: Simon Baxter, Mark Blaxter, Adriana Briscoe, Nicola Chamberlain, Brian Counterman, Laura Ferguson, Heather Hines, Chris Jiggins, Frank Jiggins, Mathieu Joron, Durrell Kapan, Marcus Kronforst, Jim Mallet, Arnaud Martin, Sean Mullen, Nicola Nadeau, William Palmer, Riccardo Papa, Megan Supple, Ayse Trolander, Annabel Whibley, Furong Yuan
Transposable elements: Brian Counterman, David Ray
Orthologue predictions: Dean Baker
Synteny: Annabel Whibley, John Davey, David Heckel, Karl Gordon
Introgression analysis: Kanchon Dasmahapatra, Nicola Nadeau, John Davey, Simon Martin, Camilo Salazar, Chris Jiggins, Mathieu Joron, James Mallet

Author Information The genome sequence has been submitted to European Nucleotide Archive with accession numbers HE667773-HE672081. The annotated genome is available on our genome browser at http://butterflygenome.org/ and will also be made available in the next release of ENSEMBL Genomes. Additional short read sequences have been submitted to the European Nucleotide Archive with accession numbers ERP000993 and ERP000991.

The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature.

Acknowledgments

We thank the governments of Colombia, Peru and Panama for permission to collect and export butterflies. Sequencing was funded by contributions from consortium members. We thank Moisés Abanto for assistance in raising the inbred line. Individual laboratories were funded by the Leverhulme Trust (CDJ), John Fell Fund and Christ Church College, Oxford (LCF), The Royal Society (MJ, CDJ), NSF (WOM, MK, RR, SM, ADB), NIH (MK, SLS, JY), CNRS (MJ), ERC (MJ, PWHH), Banco de la República and COLCIENCAS (ML), and the BBSRC (JM, CDJ, MB, R ff-C).

References

1. Seehausen O. Hybridization and adaptive radiation. Trends Ecol. Evol. 2004; 19:198–207. [PubMed: 16701254]
2. Bates HW. Contributions to an insect fauna of the Amazon valley. Lepidoptera: Heliconidae. Trans. Linn. Soc. Lond. 1862; 23:495–566.
3. Turner JRG. Adaptation and evolution in *Heliconius*: a defense of neo-Darwinism. Ann. Rev. Ecol. Syst. 1981; 12:99–121.
4. Brown KS. The biology of *Heliconius* and related genera. Ann. Rev. Entomol. 1981; 26:427–456.
5. Jiggins CD, Naisbit RE, Coe RL, Mallet J. Reproductive isolation caused by colour pattern mimicry. Nature. 2001; 411:302–305. [PubMed: 11357131]
6. Briscoe AD, et al. Positive selection of a duplicated UV-sensitive visual pigment coincides with wing pigment evolution in *Heliconius* butterflies. Proc. Natl. Acad. Sci. USA. 2010; 107:3628–3633. [PubMed: 20133601]
7. Mallet, J. Speciation and Patterns of Diversity. Butlin, RK.; Schluter, D.; Bridle, JR., editors. Cambridge University Press; Cambridge: 2009. p. 177-194.
8. Kronforst MR. Gene flow persists millions of years after speciation in *Heliconius* butterflies. BMC Evol. Biol. 2008; 8:98. [PubMed: 18371203]
9. Salazar C, et al. Genetic evidence for hybrid trait speciation in *Heliconius* butterflies. PLoS Genet. 2010; 6:e1000930. [PubMed: 20442862]
10. International Silkworm Genome Consortium. The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. Insect Biochem. Molec. Biol. 2008; 38:1036–1045. [PubMed: 19121390]
11. Pringle EG, et al. Synteny and chromosome evolution in the Lepidoptera: evidence from mapping in *Heliconius melpomene*. Genetics. 2007; 177:417–426. [PubMed: 17603110]
12. Robinson, R. Lepidoptera Genetics. Pergamon Press; Oxford: 1971.

13. Deng Q, Zeng Q, Qian Y, Li C, Yang Y. Research on the karyotype and evolution of the *Drosophila melanogaster* species group. J. Genet. Genomics. 2007; 34:196–213. [PubMed: 17498617]
14. Kemkemer C, et al. Gene synteny comparisons between different vertebrates provide new insights into breakage and fusion events during mammalian karyotype evolution. BMC Evol. Biol. 2009; 9:84. [PubMed: 19393055]
15. d'Alençon E, et al. Extensive synteny conservation of holocentric chromosomes in Lepidoptera despite high rates of local genome rearrangements. Proc. Natl. Acad. Sci. USA. 2010; 107:7680–7685. [PubMed: 20388903]
16. Vane-Wright RI, Boppré M. Visual and chemical signalling in butterflies: functional and phylogenetic perspectives. Phil. Trans. Roy. Soc. Lond. B. 1993; 340:197–205.
17. Zhan S, Merlin C, Boore JL, Reppert SM. The monarch butterfly genome yields insights into long-distance migration. Cell. 2011; 147:1171–1185. [PubMed: 22118469]
18. Vieira FG, Rozas J. Comparative genomics of the odorant-binding and chemosensory protein gene families across the Arthropoda: origin and evolutionary history of the chemosensory system. Genome Biol. Evol. 2011; 3:476–490. [PubMed: 21527792]
19. Chai CL, et al. A genomewide survey of homeobox genes and identification of novel structure of the *Hox* cluster in the silkworm, *Bombyx mori*. Insect Biochem. Molec. Biol. 2008; 38:1111–1120. [PubMed: 19280701]
20. Sackton TB, et al. Dynamic evolution of the innate immune system in *Drosophila*. Nat. Genet. 2007; 39:1461–1468. [PubMed: 17987029]
21. Green RE, et al. A draft sequence of the Neandertal genome. Science. 2010; 328:710–722. [PubMed: 20448178]
22. Durand EY, Patterson N, Reich D, Slatkin M. Testing for ancient admixture between closely related populations. Molec. Biol. Evol. 2011; 28:2239–2252. [PubMed: 21325092]
23. Reed RD, et al. *optix* drives the repeated convergent evolution of butterfly wing pattern mimicry. Science. 2011; 333:1137–1141. [PubMed: 21778360]
24. Nadeau NJ, et al. Evidence for genomic islands of divergence among hybridizing species and subspecies of *Heliconius* butterflies obtained by large-scale targeted sequencing. Phil. Trans. Roy. Soc. B. 2012; 367:343–353. [PubMed: 22201164]
25. Kim M, et al. Regulatory genes control a key morphological and ecological trait transferred between species. Science. 2008; 322:1116–1119. [PubMed: 19008450]
26. Song Y, et al. Adaptive introgression of anticoagulant rodent poison resistance by hybridization between Old World mice. Curr. Biol. 2011; 21:1296–1301. [PubMed: 21782438]

Europe PMC Funders Author Manuscripts

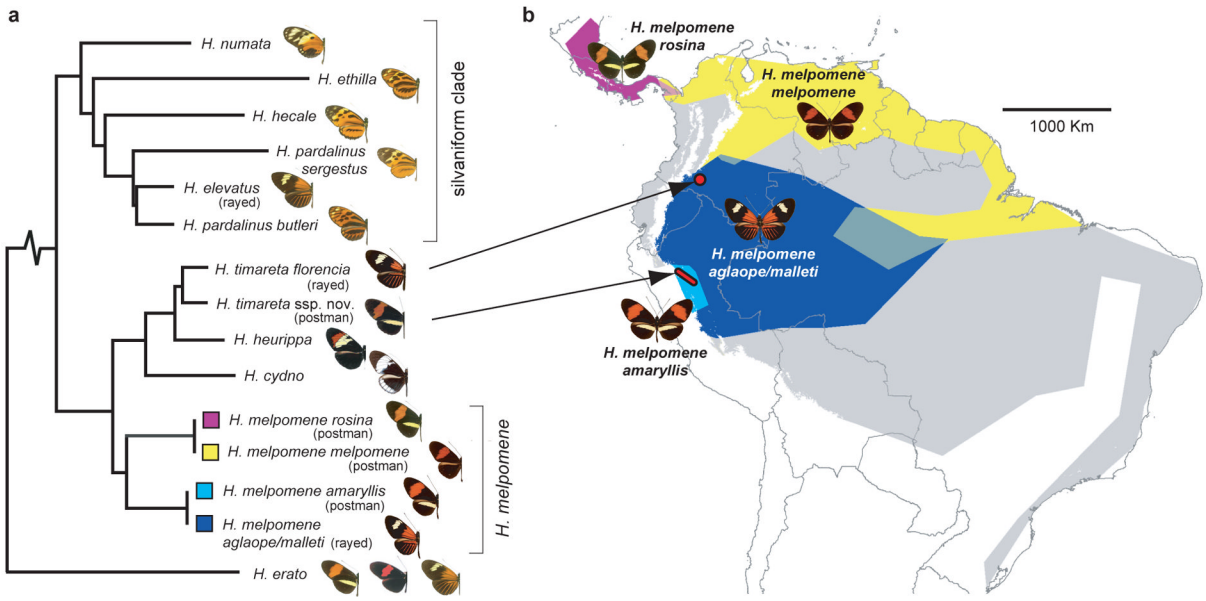

Figure 1. Distribution, mimicry and phylogenetic relationships of sequenced taxa

a Phylogenetic relationship of sequenced species and subspecies in the '*melpomene*-silvaniform clade' of *Heliconius*. *H. elevatus* falls in the 'silvaniform' clade, but its colour pattern mimics *melpomene-timareta* clade taxa. **b** Geographic distribution of 'postman' and 'rayed' *H. melpomene* races studied here (blue, yellow and purple), and the entire distribution of *H. melpomene* (grey). The *H. timareta* races investigated have limited distributions indicated by arrows (red) and mimic sympatric races of *H. melpomene*. *H. elevatus* and the other silvaniform species are distributed widely across the Amazon basin (Supplementary Information 22).

Europe PMC Funders Author Manuscripts

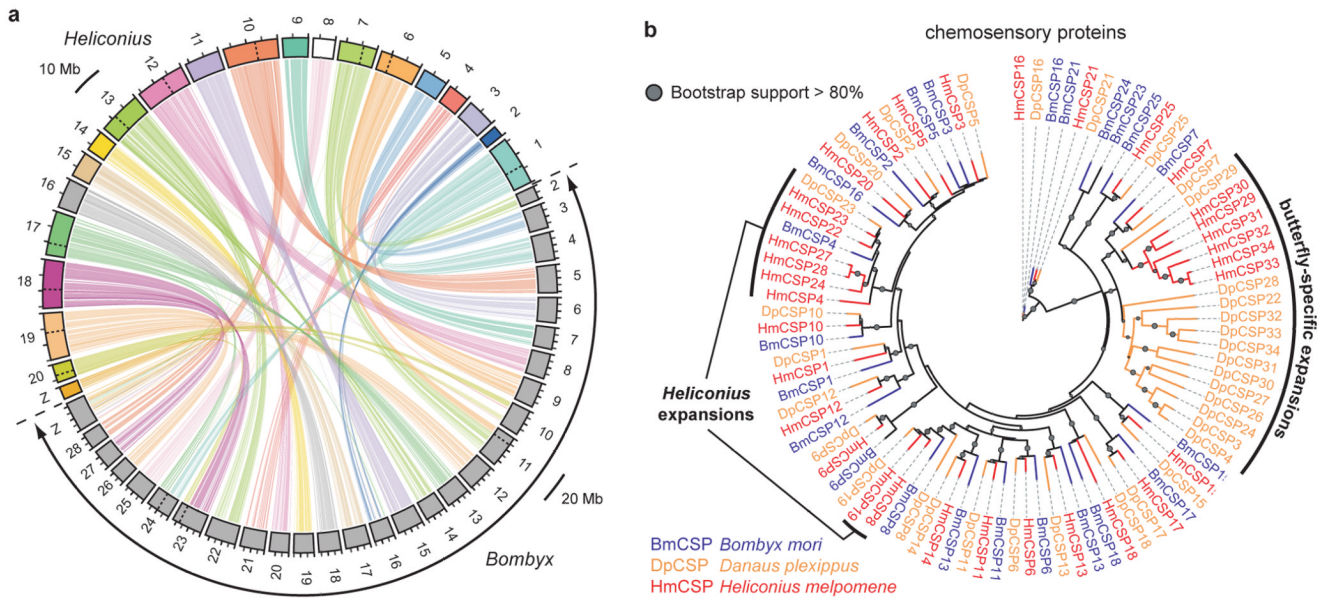

Figure 2. Comparative analysis of synteny and expansion of the chemosensory genes

a Maps of the 21 *Heliconius* chromosomes (above) and of the 28 *Bombyx* chromosomes (below in grey) based on positions of 6010 orthologue pairs demonstrate highly conserved synteny and a shared n=31 ancestor (Supplementary Information 8). Dotted lines within chromosomes indicate major chromosomal fusions. **b** Maximum likelihood tree showing expansion of the chemosensory protein genes (CSP) in two butterfly genomes.

Europe PMC Funders Author Manuscripts

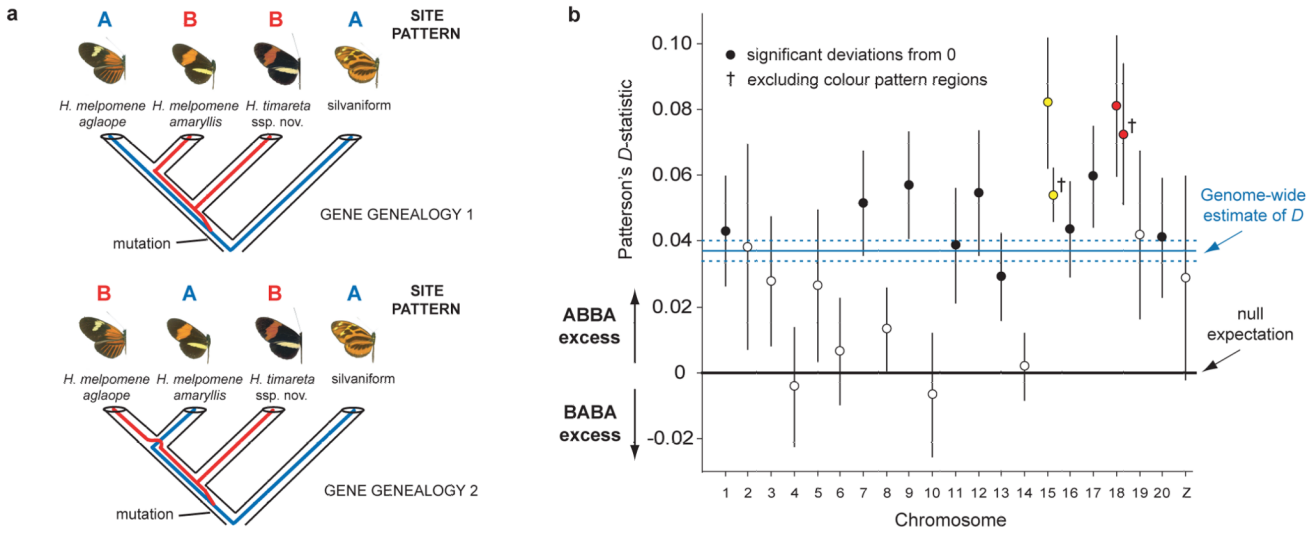

Figure 3. Four-taxon ABBA-BABA test of introgression

a ABBA and BABA nucleotide sites employed in the test are derived (– – **B** –) in *H. timareta*, compared with the silvaniform outgroup (– – – **A**), but differ among *H. melpomene amaryllis* and *H. melpomene aglaope* (either **AB**BA or **BA**BA). As this almost exclusively restricts attention to sites polymorphic in the ancestor of *H. timareta* and *H. melpomene*, equal numbers of ABBA and BABA sites[22] are expected under a null hypothesis of no introgression, as depicted in the two gene genealogies. **b** Distribution among chromosomes of Patterson's *D*-statistic ± s.e., which measures excess of ABBA vs. BABA sites[22], here for the comparison *H. m. aglaope*; *H. m. amaryllis*; *H. timareta* ssp. nov.; silvaniform. Chromosomes containing the two colour pattern regions (*B*/*D* red; *N*/*Yb* yellow) have the two highest *D*-statistics; the combinatorial probability of this occurring by chance is 0.005. The excess of ABBA sites ($0 < D < 1$) indicates introgression between sympatric *H. timareta* and *H. melpomene amaryllis*.

Europe PMC Funders Author Manuscripts

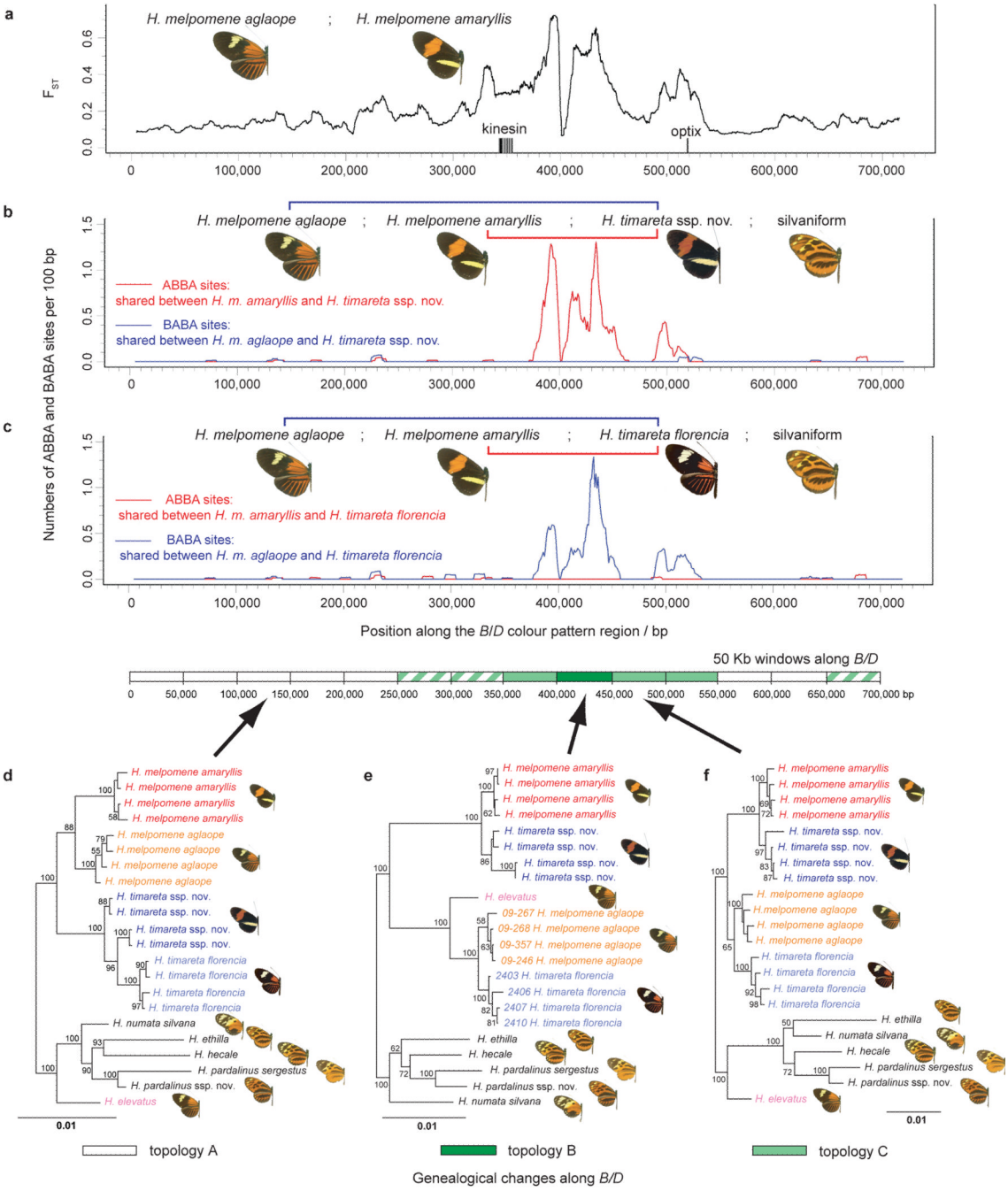

Figure 4. Evidence for adaptive introgression at the *B/D* mimicry locus

a Genetic divergence between *H. melpomene* races *aglaope* (rayed) and *amaryllis* (postman) across a hybrid zone in N.E. Peru. Divergence, F_{ST}, is measured along the *B/D* region (Supplementary Information 14). F_{ST} peaks in the region known to control red wing pattern elements between the genes *kinesin* and *optix*[23]. **b, c** Distribution of fixed ABBA and BABA sites (see Fig. 4a) along *B/D* for two comparisons. Excesses of ABBA in **b** and BABA in **c** are highly significant (two-tailed *Z*-tests for $D = 0$; $D = 0.90 \pm 0.13$, $P = 5 \times 10^{-14}$ and $D = -0.91 \pm 0.10$, $P = 9 \times 10^{-24}$ respectively), indicating introgression. **d, e, f,** Genealogical change along *B/D* investigated with maximum likelihood based on 50 kb

Europe PMC Funders Author Manuscripts

windows. Three representative tree topologies are shown. Topology A, the species tree, is found within the white windows. In topologies B (dark green window) and C (light green windows) taxa group by colour pattern rather than species. Within striped windows *H. melpomene* and/or *H. timareta* are paraphyletic but the taxa do not group by colour pattern. Support is shown for nodes with > 50% bootstrap support (Supplementary Information 19).

Europe PMC Funders Author Manuscripts