Butterfly genome reveals promiscuous exchange of mimicry adaptations among species

The Heliconius Genome Consortium*

The evolutionary importance of hybridization and introgression has long been debated. Hybrids are usually rare and unfit, but even infrequent hybridization can aid adaptation by transferring beneficial traits between species. Here we use genomic tools to investigate introgression in Heliconius, a rapidly radiating genus of neotropical butterflies widely used in studies of ecology, behaviour, mimicry and speciation. 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 12,669 predicted genes, biologically important expansions of families of chemosensory and Hox genes are particularly noteworthy. Chromosomal organization has remained broadly conserved since the Cretaceous period, when butterflies split from the Bombyx (silkmoth) lineage. Using genomic resequencing, we show hybrid exchange of genes between three co-mimics, Heliconius melpomene, Heliconius timareta and Heliconius elevatus, especially at two genomic regions that control mimicry pattern. We infer that closely related Heliconius species exchange protective colour-pattern genes promiscuously, implying that hybridization has an important role 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 lifespan, augmented ultraviolet colour vision, ‘trap-lining’ foraging behaviour, gregarious roosting and complex mating behaviours, and provides outstanding opportunities for genomic studies of adaptive radiation and speciation. The genus is best known for the hundreds of races with different colour patterns seen among its 43 species, with repeated examples of both convergent evolution among distantly related species and divergent evolution between closely related taxa. 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. Heliconius are unpalatable to vertebrate predators and Müllerian mimicry of warning colour patterns enables species to share the cost of educating predators. As a result of its dual role in mimicry and mate selection, divergence in wing pattern is also associated with speciation and adaptive radiation. A particularly recent radiation is the melpomene–silvaniform clade, in which mimetic patterns often seem to be polyphyletic (Fig. 1a). Most species in this clade occasionally hybridize in the wild with other clade members. Gene genealogies at

*Lists of participants and their affiliations appear at the end of the paper.
a small number of loci indicate introgression between species, and one non-mimetic species, Heliconius heurippa, has a hybrid origin. 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 through five generations of sib mating. We sequenced a single male to ×38 coverage (after quality filtering) using combined 454 and Illumina technologies (Supplementary Information, sections 1–8). The complete draft genome assembly, which is 269 megabases (Mb) in size, consists of 3,807 scaffolds with an N50 of 277 kb and contains 12,669 predicted protein-coding genes. Restriction-site-associated DNA (RAD) linkage mapping was used to assign and order 83% of the sequenced genome onto the 21 chromosomes (Supplementary Information, section 4). These data permit a considerably improved genome-wide chromosomal synteny comparison with the silkmoth Bombyx mori.

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 that ten linkage groups have major contributions from two B. mori chromosomes (Fig. 2a and Supplementary Information, section 8), revealing several previously unidentified chromosomal fusions. These fusions on the Heliconius lineage most probably occurred after divergence from the sister genus Eueides, which has the lepidopteran modal karyotype of n = 31 (ref. 12). Three chromosomal fusions are evident in Bombyx (B. mori chromosomes 11, 23 and 24; Fig. 2a), as required for evolution of the Bombyx n = 28 karyotype from the ancestral n = 31 karyotype. Heliconius and Bombyx lineages diverged in the Cretaceous, more than 100 million years ago, so the gross chromosomal structures of Lepidoptera genomes have remained highly conserved compared with those of flies or vertebrates. By contrast, small-scale rearrangements were previously suggested for Lepidoptera, these rates are comparable to those in Drosophila (Supplementary Information, section 8).

The origin of butterflies was associated with a switch from nocturnal to diurnal behaviour, and a corresponding increase in visual communication. Heliconius have increased visual complexity through expression of a duplicate ultraviolet opsin, in addition to the long-wavelength-, blue- and ultraviolet-sensitive opsins in Bombyx. We might therefore predict reduced complexity of olfactory genes, but in fact Heliconius and Danaus genomes have more chemosensory genes than any other insect genome: 33 and 34, respectively (Supplementary Information, section 9). For comparison, there are 24 in Bombyx and 3–4 in Drosophila. Lineage-specific expansions of chemosensory genes were evident in both Danaus and Heliconius (Fig. 2b). By contrast, all three lepidopteran genomes have similar numbers of odorant binding proteins and olfactory receptors (Supplementary Information, section 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, section 10). These Hox gene duplications in the butterflies and Bombyx have a common origin and are independent of the two tandem duplications known in dipterans (zen2 and bcd). Immunity-related gene families are similar across all three lepidopterans (Supplementary Information, section 11), whereas there are extensive duplications and losses within diptera.

The Heliconius reference genome allowed us to perform 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 and Supplementary Information, sections 12–18). We then tested for introgression between the sympatric co-mimetic postman butterfly races of Heliconius melpomene amaryllis and H. timareta ssp. nov. (Fig. 1) in Peru, using ‘ABBA/BABA’ single nucleotide sites and Patterson’s D-statistics (Fig. 3a), originally developed to test for admixture between Neanderthals and modern humans (Supplementary Information, section 12). Genome-wide, we found an excess of ABBA sites, giving a significantly positive Patterson’s D of 0.037 ± 0.003 (two-tailed Z-test for D = 0, P = 1 × 10⁻⁴⁰), indicating greater genome-wide introgression between the sympatric mimetic taxa H. melpomene amaryllis and

**Figure 2 | Comparative analysis of synteny and expansion of the chemosensory genes.** a, Maps of the 21 Heliconius chromosomes (colour) and the 28 Bombyx chromosomes (grey) based on positions of 6,010 orthologue pairs demonstrate highly conserved synteny and a shared n = 31 ancestor (Supplementary Information, section 8). Dotted lines within chromosomes indicate major chromosomal fusions. b, Maximum-likelihood tree showing expansions of chemosensory protein (CSP) genes in the two butterfly genomes.
Heliconius m. amaryllis and H. melpomene aglaope (either ABBA or BABA). 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 are expected under a null hypothesis of no introgression, as depicted in the two gene genealogies. Distribution among chromosomes of Patterson's D-statistic (± s.e.), which measures excess of ABBA sites over BABA sites, 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. m. amaryllis.

D-statistic (± s.e.), which measures excess of ABBA sites over BABA sites, here for the comparison: 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. m. amaryllis.

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 ABBA or BABA). 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 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 sites over BABA sites, here for the comparison: 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. m. amaryllis.

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 unrelated to colour-pattern loci, tree topologies are similar to the predominant signal recovered from the genome as a whole (Supplementary Information, section 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, topologies within the region of peak ABBA/BABA differences group individuals by colour pattern, and the species themselves become polyphyletic (Fig. 4e, f and Supplementary Information, sections 19 and 20). Remarkably, the rayed H. elevatus, a member of the silvaniform clade according to genome average relationships (Fig. 1a and Supplementary Information, section 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 and Supplementary Information, section 18).
**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 northeast Peru. Divergence, \( F_{ST} \), is measured along the B/D region (Supplementary Information 14) and peaks in the region known to control red wing pattern elements between the genes *kinesin* and *optix*\(^2\). **b, c.** Distribution of fixed ABBA and BABA sites (see Fig. 3a) 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.

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 large (~100-kb) adaptive regions among multiple species in a recent radiation. Our genome-scale analysis provides considerably greater power than previous tests of introgression\(^8,25–27\). Our evidence suggests that *H. elevatus* like *H. hecippus*\(^a\), was formed during a hybrid speciation event. The main genomic signal from this rayed species places it closest to *Heliconius pardalus butleri* (Fig. 1a), but colour-pattern genomic regions resemble those of rayed races of *H. melpomene* (Fig. 4e and Supplementary Information, sections 18–21). Colour pattern is important in mating behaviour in *Heliconius*\(^c\), and the transfer of mimetic pattern may have enabled the divergent sibling species *H. elevatus* to coexist with *H. pardalus* across the Amazon basin. Although it was long suspected that introgression might be important in evolutionary radiations\(^1\), our results from the most diverse terrestrial biome on the planet suggest that adaptive introgression is more pervasive than previously realized.

The annotated genome version 1.1 is available on the *Heliconius* Genome Consortium’s genome browser at http://butterflygenome.org/ and this version will also be included in the next release of ENSEMBL Genomes. A full description of methods can be found in Supplementary Information.

Received 26 October 2011; accepted 12 March 2012.
Published online 16 May 2012.

1. Seehausen, O. Hybridization and adaptive radiation. *Trends Ecol. Evol.* **19**, 198–207 (2004).
2. Bates, H. W. Contributions to an insect fauna of the Amazon valley. *Lepidoptera: Heliconiidae*. *Trans. Linn. Soc. Lond.* **23**, 495–566 (1862).
3. Turner, J. R. G. Adaptation and evolution in *Heliconius*: a defense of neo-Darwinism. *Annu. Rev. Ecol. Syst.* **12**, 99–121 (1981).
4. Brown, K. S. The biology of *Heliconius* and related genera. *Annu. Rev. Entomol.* **26**, 427–457 (1981).
5. Jiggins, C. D., Naish, R. E., Coe, R. L. & Mallet, J. Reproductive isolation caused by colour pattern mimicry. *Nature* **411**, 302–305 (2001).
6. Briscoe, A. D. et al. Positive selection of a duplicated UV-sensitive visual pigment coincides with wing pigment evolution in Heliconius butterflies. Proc. Natl Acad. Sci. USA 107, 3628–3633 (2010).

7. Mall compression in Drosophila melanogaster: evidence from mapping in Heliconius melpomene. Genetics 177, 417–426 (2007).

8. Robinson, R. Lepidoptera Genetics 575–598 (Pergamon, 1971).

9. Deng, Y., Zeng, Q., Ouyang, Y., Li, R. & Yang, Y. Research on the karyotype and evolution of the Drosophila melanogaster species group. J. Genet. Genomics 34, 196–213 (2007).

10. Kanchon K. Dasmahapatra, James R. Walters, Adriana Briscoe, John W. Davey, Annabel Whibley, Nicola J. Nadeau, Alexey V. Vrbovsky, D. Graham Burt, T. Hughes, Larry J. Genet. J. 177, 169–174 (Cambridge Univ. Press, 2009).

11. Kronforst, M. Gene flow persists millions of years after speciation in Heliconius butterflies. BMC Evol. Biol. 8, 58 (2008).

12. Saiz, F., et al. Genetic evidence for hybrid trac trait speciation in Heliconius butterflies. PLoS Genet. 6, e1000930 (2010).

13. National. Silkworm Genome Consortium. The genome of a lepidopteran model insect, the silkworm Bombyx mori. Insect Biochem. Mol. Biol. 38, 1036–1045 (2008).

14. Pringle, E. Q. et al. Synteny and chromosome evolution in the Lepidoptera: evidence from mapping in Heliconius melpomene. Genetics 177, 417–426 (2007).

15. Robinson, R. Lepidoptera Genetics 575–598 (Pergamon, 1971).

16. Sackton, T. B. Nature Rev. Genet. 39, 1461–1466 (2008).

17. Green, R. E. et al. A draft sequence of the Neandertal genome. Science 328, 710–712 (2010).

18. Durand, E. V., Patterson, N., Reich, D. & Slatkin, M. Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28, 2239–2251 (2011).

19. Reed, R. D. et al. optix drives the repeated convergent evolution of butterfly wing pattern mimicry. Science 333, 1137–1141 (2011).

20. Nadeau, N. J. et al. Evidence for genomic islands of divergence among hybridizing species and subspecies of Heliconius butterflies obtained by large-scale targeted sequencing. Phil. Trans. R. Soc. B 367, 343–353 (2012).

21. Bull, V. et al. Polyphyletic and gene flow between non-sibling Heliconius species. BMC Evol. Biol. 8, 206 (2008).

22. Kim, M. et al. Regulatory genes control a key morphological and ecological trait transferred between species. Science 322, 1116–1119 (2008).

23. Song, Y. et al. Adaptive introgression of anticoagulant rodent poison resistance by hybridization between Old World mice. Curr. Biol. 21, 1296–1301 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the governments of Colombia, Peru and Panama for permission to collect the butterflies. Sequencing was funded by contributions from consortium members. We thank M. Abanto for assistance in raising the inbred line. permission to collect the butterflies. Sequencing was funded by contributions from www.nature.com/nature.