No genomic mosaicism in a putative hybrid butterfly species

Marcus R. Kronforst^{1,*}, Camilo Salazar², Mauricio Linares²
and Lawrence E. Gilbert¹

¹Section of Integrative Biology, University of Texas, Austin, TX 78712, USA

²Instituto de Genética, Universidad de los Andes, PO Box 4976, Bogotá, Colombia

Recent descriptions of hybrid animal species have spurred interest in this phenomenon, but little genomic data exist to support it. Here, we use frequency variation for 657 amplified fragment length polymorphism (AFLP) markers and DNA sequence variation from 16 genes to determine whether the genome of *Heliconius pachinus*, a suspected hybrid butterfly species, is a mixture of the putative parental species, *Heliconius cydno* and *Heliconius melpomene*. Despite substantial shared genetic variation among all three species, we show that the genome of *H. pachinus* is not a mosaic; both AFLP and DNA sequence data overwhelmingly associate *H. pachinus* with just one of the potential parents, *H. cydno*. This pattern also applies to the gene *wingless*, which is tightly linked to the locus that determines forewing colour—one specific *H. pachinus* trait that has been hypothesized to have originated from *H. melpomene*. As a whole, the data support a traditional, bifurcating model of speciation in which *H. pachinus* split from a common ancestor with *H. cydno* without a genetic contribution from *H. melpomene*. However, comparison of our data to DNA sequence data for another putative hybrid *Heliconius* species, *Heliconius heurippa*, suggests that the *H. heurippa* genome may be a mosaic.

Keywords: amplified fragment length polymorphism; *Heliconius*; introgressive hybridization; hybrid speciation; Lepidoptera

1. INTRODUCTION

Understanding how new species originate is a major focus of research in evolutionary biology. In general, the formation of new species follows a branching pattern, in which an ancestor gives rise to two (occasionally more) descendent lineages. An alternative to this standard bifurcation model of speciation is hybrid speciation, in which interbreeding between two species gives rise to a third distinct lineage. In plants, hybrid speciation as a result of allopolyploidy is well known (Rieseberg 1997; Levin 2002; Coyne & Orr 2004) and there are examples of allopolyploid hybrid animal species as well (Dowling & Secor 1997). Much less common is homoploid or recombinational hybrid speciation, in which hybridization generates a new species without a change in chromosome number. In theory, this process should be rare because a new homoploid hybrid lineage faces a variety of obstacles that are likely to prevent it from becoming established. Most critically, a hybrid lineage must become reproductively isolated from both parental species quickly otherwise it risks fusing with one of them (Buerkle *et al.* 2000). It also must be ecologically distinct otherwise it risks competitive exclusion (Coyne & Orr 2004). Owing to these obstacles, and a commonly held view that hybridization, in general, is quite rare in animals, hybrid

speciation in animals has received little attention historically (Dowling & Secor 1997). Recently, however, multiple descriptions of putative homoploid hybrid animal species have spurred interest in this phenomenon (Salzburger *et al.* 2002; Smith *et al.* 2003; Schwarz *et al.* 2005; Gompert *et al.* 2006; Mavarez *et al.* 2006; Meyer *et al.* 2006).

One animal group where hybridization is strongly suspected to have played an important role in driving phenotypic and species diversification is the neotropical butterfly genus *Heliconius* (Linares 1989; Gilbert 2003; Mavarez *et al.* 2006). Hybridization between two species in particular, *Heliconius cydno* and *Heliconius melpomene*, is hypothesized to have fuelled mimetic convergence within the genus (Gilbert 2003) and may be responsible for the origin of at least four species: *Heliconius pachinus*; *Heliconius heurippa*; *Heliconius timareta*; and *Heliconius tristero* (Gilbert 2003; Mavarez *et al.* 2006). Despite having diverged in a variety of ecological characteristics and mimetic wing patterns, *H. cydno* and *H. melpomene* continue to hybridize (Mallet *et al.* 1998, 2006) and exchange genes (Bull *et al.* 2006; Kronforst *et al.* 2006a) throughout Central America and Andean South America. The four putative hybrid species are all currently recognized as very close relatives of *H. cydno* based on morphology, behaviour, degree of interfertility and genetic data, but each displays colour pattern elements characteristic of a sympatric race of *H. melpomene*. For example, *H. pachinus* is ecologically and behaviourally identical to *H. cydno*, it is completely interfertile with *H. cydno* and DNA sequence data for mitochondrial and nuclear genes generally cluster *H. pachinus* haplotypes with those of

* Author and address for correspondence: Department of Ecology and Evolutionary Biology, Rice University, Houston, TX 77005, USA (mkronforst@mail.utexas.edu).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2006.0207> or via <http://www.journals.royalsoc.ac.uk>.

H. cydno (Brower 1994; Brower & Egan 1997; Beltrán *et al.* 2002; Kronforst *et al.* 2006a). However, at six of seven Mendelian colour patterning loci, *H. pachinus* shares functionally equivalent alleles with *H. melpomene*. The *H. pachinus* colour pattern elements that may be derived from *H. melpomene* include the yellow colour of the forewing, the yellow colour of the hindwing, the position of the distal forewing shutter, the presence of the proximal forewing shutter, the absence of the hindwing shutter and the presence of red at the ventral base of the wings (Gilbert 2003). Similarly, a variety of data indicate that *H. heurippa*, *H. timareta* and *H. tristero* are closely related to *H. cydno* (Brower 1996a; Brower & Egan 1997; Salazar *et al.* 2005), yet each has colour pattern elements that appear to be derived from *H. melpomene*. For *H. heurippa*, these include the presence of red on the forewing and at the base of the ventral fore- and hindwing, the absence of brown on the ventral hindwing and the absence of blue iridescence. For *H. timareta*, these include the size and shape of the yellow forewing patch and the presence of red hindwing rays. *Heliconius tristero* displays a wing pattern virtually indistinguishable from the local race of *H. melpomene* but has mitochondrial haplotypes and genitalic morphology that place it in the *H. cydno* clade (Brower 1996a).

These data suggest a scenario in which introgressive hybridization may have moved alleles at certain colour patterning loci from *H. melpomene* to a largely *H. cydno* genetic background to produce the recombinant forms *H. pachinus*, *H. heurippa*, *H. timareta* and *H. tristero*. This hypothesis predicts that the genomes of these species should be a mosaic of the *H. cydno* and *H. melpomene* genomes. Furthermore, the hybrid-origin hypothesis makes specific predictions about what proportion of the genome should be derived from each parental species. Since female F₁ hybrids between *H. cydno* and *H. melpomene* are sterile, the most direct route from initial hybridization to a phenotype similar to any of these putative hybrid species involves one or two generations of backcrossing to pure *H. cydno*, followed by mating between these new fertile female hybrids and first-generation or backcross male hybrids. Indeed, a series of crosses like this have recently been shown to result in a wing colour pattern identical to that of the suspected hybrid species *H. heurippa* (Mavarez *et al.* 2006). Under this scenario, the hybrid-origin hypothesis predicts a mosaic genome with approximately 65–75% derived from *H. cydno* and 25–35% derived from *H. melpomene*.

Recently, we generated two comparative genomic datasets which allow us to test the hypothesis that the genome of one of these suspected hybrid species, *H. pachinus*, is a mosaic (Kronforst *et al.* 2006a). The first dataset consists of hundreds of amplified fragment length polymorphism (AFLP) markers from a total of 127 *H. cydno*, *H. pachinus* and *H. melpomene* individuals. The second dataset consists of comparative DNA sequence data for 15 genes from all three species and the closely related outgroup *Heliconius hecale*. Previously, we used these data to infer the extent of admixture and the direction of interspecific gene flow among *H. cydno*, *H. pachinus* and *H. melpomene* in Costa Rica (Kronforst *et al.* 2006a). The results of these analyses indicated that all three species exhibit detectable contemporary admixture and high levels of introgression. While these

findings suggest that the basic conditions necessary for the origin of new species via hybridization exist in *Heliconius*, they do not tell us whether this has occurred.

Here, we supplement these data with DNA sequences for an additional gene, *decapentaplegic*, and use them to test whether the genome of *H. pachinus* is a mixture of the genomes of *H. cydno* and *H. melpomene*. Furthermore, we explicitly test the hypothesis that the yellow forewing colour of *H. pachinus* is derived from *H. melpomene* by analysing DNA sequence data for the gene *wingless*, which is very tightly linked to the Mendelian locus that controls this trait (Kronforst *et al.* 2006b). Finally, we combine our DNA sequence data with published sequences for a second potential hybrid species, *H. heurippa*, to see whether our conclusions regarding the origin of the *H. pachinus* genome are likely to apply to this species as well. The results suggest that while there is a large amount of shared genetic variation among all of these species, the genome of *H. pachinus* is not a mosaic. In contrast, the limited data available for *H. heurippa* are consistent with the genomic mosaicism expected as a result of hybrid speciation.

2. MATERIAL AND METHODS

(a) Sampling

We collected 56 *H. cydno*, 44 *H. pachinus* and 27 *H. melpomene* specimens from locations throughout Costa Rica. All samples were collected in the field as adults between June and August 2000 or 2002. Tissue was preserved in 95% ethanol and total genomic DNA was extracted using a DNeasy Tissue Kit (Qiagen, Valencia, CA) or a standard phenol/chloroform extraction protocol.

(b) Molecular methods

We genotyped each of the 127 *H. cydno*, *H. pachinus* and *H. melpomene* specimens with AFLPs (Vos *et al.* 1995) using the ABI Plant Mapping Kit (PE Applied Biosystems, Foster City, CA) and 4 selective primer combinations, i.e. EcoRI-ACT/MseI-CAT, EcoRI-ACT/MseI-CTG, EcoRI-ACA/MseI-CAT and EcoRI-ACA/MseI-CTG. Fragments were separated with an ABI Prism 3100 genetic analyzer (PE Applied Biosystems). AFLP fragments between 50 and 500 bp were sized and scored using ABI GENEMAPPER software v. 3.7 (PE Applied Biosystems). Fragments with a peak height below 100 reflectance units were scored as an absence.

We sequenced multiple haplotypes for 1 mitochondrial and 15 nuclear loci from *H. cydno*, *H. pachinus*, *H. melpomene* and *H. hecale* (table 1 in electronic supplementary material). The loci are *apterous* (*ap*), *cubitus interruptus* (*ci*), *cinnabar* (*cn*), *Cytochrome oxidase* (mtDNA—CO), *decapentaplegic* (*dpp*), *Distal-less* (*Dll*), *Elongation factor 1 α* (*Ef1 α*), *engrailed* (*en*), *invected* (*inv*), *Mannose phosphate isomerase* (*Mpi*), *patched* (*ptc*), *scalloped* (*sd*), *scarlet* (*st*), *Triose phosphate isomerase* (*Tpi*), *white* (*w*) and *wingless* (*wg*). All 15 nuclear loci have been placed on a *H. cydno* genetic map (figure 1 in electronic supplementary material). Sequences have been deposited in GenBank under accession numbers AY744577–AY744672, AY745254–AY745278, AY745315–AY745335, AY745356–AY745490, DQ448305–DQ448516 and EF041105–EF041122. Additional details regarding sampling and molecular methods can be found in Kronforst (2005) and Kronforst *et al.* (2006a).

(c) Amplified fragment length polymorphism data analyses

To determine whether *H. pachinus* originated with a genetic contribution from both *H. cydno* and *H. melpomene*, we analysed the AFLP data in three ways. First, to assess whether *H. pachinus* is genetically intermediate to *H. cydno* and *H. melpomene*, we plotted the relative genetic distance among all *H. cydno*, *H. pachinus* and *H. melpomene* individuals using multidimensional scaling (MDS). For this analysis, we computed pairwise Euclidean square distances among individuals using ARLEQUIN v. 2.0 (Schneider *et al.* 2000) and these distances were then used in a metric MDS analysis which was performed with NCSS v. 2000 (Hintz 2001). To visualize relationships among species, we plotted individuals based on the two dimensions which encompassed the most inter-individual variation. This analysis was performed with three partitions of the AFLP data: the full dataset; the subset of loci that consisted of ancestry informative markers (AIMs)—loci that had significant allele frequency differences between *H. cydno* and *H. melpomene* ($p < 0.05$); and the subset of AIMs that retained a significant frequency difference between *H. cydno* and *H. melpomene* after Bonferroni correction for multiple comparisons. AIMs were identified by performing a locus-by-locus analysis of molecular variance (AMOVA) with the *H. cydno* and *H. melpomene* AFLP data using ARLEQUIN. The p -values for the locus-specific F_{ST} estimates that resulted from this analysis were based on a distribution of values obtained by permuting individuals between species 100 000 times.

For our second analysis, we used the model-based clustering approach implemented in the program STRUCTURE (Pritchard *et al.* 2000; Falush *et al.* 2003) to assess whether the *H. pachinus* genome is a mixture of the *H. cydno* and *H. melpomene* genomes. If *H. pachinus* originated due to hybridization between *H. cydno* and *H. melpomene*, we expect STRUCTURE-based clustering to partially assign each *H. pachinus* individual to each of the parental clusters when clustering is based on the assumption of two populations. Without using prior information regarding the identity of individuals, we inferred genetic clusters and individual genomic admixture for all 127 individuals, assuming both $K=2$ and 3 populations using the admixture model and 10^6 iterations of data collection following a 10 000 iteration burn-in. This analysis was performed separately with each of the three partitions of the AFLP dataset discussed previously.

Finally, we further investigated the relationship between *H. pachinus* and the putative parental species at each of the AIMs. The hybrid-origin hypothesis predicts that the allele frequencies should be similar between *H. pachinus* and *H. cydno* for some AIMs, while at others the allele frequencies between *H. pachinus* and *H. melpomene* should be similar. To test this hypothesis, we performed pairwise locus-by-locus AMOVAs between *H. pachinus* and each of the putative parental species for those loci that distinguished the parents. Based on these results, we assigned each locus to one of the four categories: those that associated *H. pachinus* with *H. cydno* (loci for which the allele frequency in *H. pachinus* was not significantly different from the frequency in *H. cydno* but was significantly different from the frequency in *H. melpomene*); those that associated *H. pachinus* with *H. melpomene*; those that were distinct in *H. pachinus* (the frequency in *H. pachinus* was different from both *H. cydno* and *H. melpomene*); and those for which *H. pachinus* was not significantly different from either putative parent. As controls for this analysis, we treated each of

the parental species as a potential hybrid between the other two species and repeated the steps: identifying AIMs that distinguished the 'parental' species; comparing the 'hybrid' allele frequency to each of the 'parent' allele frequencies at each AIM; and assigning each locus to one of the four possible categories. We performed these analyses with AIM datasets obtained before and after correcting for multiple comparisons.

(d) DNA sequence analyses

The datasets for the genes *Mpi*, *Tpi*, *CO* and *Ef1 α* were supplemented with sequences available on GenBank (accession numbers *Mpi*: AF413731, AF413734, AF413739–AF413744, AF516220, AY332417–AY332422, AY332461–AY332464; *Tpi*: AF413778, AF413782–AF413790, AY329804, AY329805, AY329839–AY329843; *CO*: U08482, U08483, U08500, U08518, U08520, U08523, U08524, U08544, AF413672–AF413674, AF413679, AF413683, AF413707; *Ef1 α* : AY090168). All of these sequences have previously been reported by Brower (1994, 1996b), Beltrán *et al.* (2002), Wahlberg *et al.* (2003) or Flanagan *et al.* (2004). To examine the evolutionary history of each locus, we estimated a gene genealogy using the neighbour-joining method implemented in MEGA v. 2.1 (Kumar *et al.* 2001) based on uncorrected pairwise proportional differences with gaps excluded in pairwise comparisons. The strength of support for each node was assessed by bootstrapping (1000 replicates). We quantified the overall genetic similarity among *H. cydno*, *H. pachinus*, *H. melpomene* and *H. hecale* at each locus by computing pairwise F_{ST} values using ARLEQUIN. The p -values for these F_{ST} estimates were based on a distribution of values obtained by permuting haplotypes between species 1000 times. For the gene *wingless*, we also tested the support for alternative evolutionary histories using a Bayesian phylogenetic approach. Using MRBAYES v. 3.0 (Huelsenbeck & Ronquist 2001), we constrained *H. pachinus* *wg* haplotypes to form a clade with those of either *H. cydno* or *H. melpomene*, and we compared the marginal likelihoods for these two hypotheses using Bayes factor (Kass & Raftery 1995; Nylander *et al.* 2004). For each analysis, four Metropolis-Coupled Markov chains were run for 750 000 generations following 250 000 burn-in generations, sampling every 100 generations, starting from a random tree. Parameters were estimated based on the GTR+I+ Γ model.

Finally, we supplemented our DNA sequence datasets for two genes, *Dll* and *inv*, with published sequences from another putative hybrid species, *H. heurippa*, to determine whether the genomic patterns uncovered for *H. pachinus* are likely to apply to other potential hybrid *Heliconius* species. These additional data, which are from Mavarez *et al.* (2006), include sequences from *H. heurippa* as well as the local races of the putative parental species, *Heliconius cydno cordula* and *Heliconius melpomene melpomene* (accession numbers *Dll*: DQ445384–DQ445415; *inv*: DQ445416–DQ445457). Using the combined data, we estimated neighbour-joining gene trees and computed pairwise F_{ST} values among *H. pachinus*, *H. heurippa*, *H. cydno*, *H. melpomene* and *H. hecale*.

3. RESULTS

(a) Amplified fragment length polymorphism data associate *H. pachinus* with only one potential parental species, *H. cydno*

We genotyped each *H. cydno*, *H. pachinus* and *H. melpomene* specimen for the presence or absence of a

Figure 1. Clustering of *H. cydno*, *H. pachinus* and *H. melpomene* individuals based on 173 AFLP loci that distinguished the putative parental species *H. cydno* and *H. melpomene*. (a) Metric multidimensional scaling (MDS) of all individuals based on pairwise Euclidean square distances. If the *H. pachinus* genome was a mixture of the *H. cydno* and *H. melpomene* genomes, MDS is expected to locate the *H. pachinus* cluster between the other two species, which it does not. Rather, the *H. pachinus* cluster is close to that of *H. cydno* and both are equidistant from *H. melpomene*. (b) STRUCTURE-based clustering of all individuals assuming both $K=3$ and $K=2$ populations. Each individual is represented by a narrow vertical column and the proportion of each of the colours signifies the posterior mean proportion of ancestry from each of the inferred populations. Genomic mosaicism of *H. pachinus* is expected to result in partial assignment of *H. pachinus* individuals to both *H. cydno* and *H. melpomene* when clustering assumes two populations. In contrast to this expectation, clustering based on the assumption of two populations assigns *H. pachinus* exclusively to a cluster with *H. cydno*. Figure 2 in electronic supplementary material shows results of the same analyses based on the full AFLP dataset and the subset of 62 loci that distinguish *H. cydno* and *H. melpomene* after correcting for multiple comparisons.

fragment at 664 AFLP loci. Of the markers, 7 were monomorphic (fixed for the presence of a fragment in all individuals) and 88 were singletons (fragment present in only one individual). All three analyses of the AFLP data supported a close genetic relationship between *H. pachinus* and *H. cydno* and none suggested a genetic contribution from *H. melpomene* in the origin of *H. pachinus*. MDS and STRUCTURE-based clustering both correctly identified the three genetic clusters corresponding to the three species. Neither analysis suggested that *H. pachinus* was genetically intermediate to *H. cydno* and *H. melpomene*; the *H. pachinus* cluster was not located between the *H. cydno* and *H. melpomene* clusters in the MDS plot and STRUCTURE clustering assuming $K=2$ populations did not partially assign *H. pachinus* individuals to each of the *H. cydno* and *H. melpomene* clusters (figure 1). These conclusions apply to analyses based on the full AFLP dataset (figure 2 in electronic supplementary material), the subset of 173 loci with a significant ($p<0.05$) frequency difference between *H. cydno* and *H. melpomene* (figure 1) and the subset of 62 loci with a significant ($p<0.00008$) frequency difference between *H. cydno* and *H. melpomene* after correcting for multiple comparisons (figure 2 in electronic supplementary material). Of the 173

loci that distinguished *H. cydno* and *H. melpomene* (only one of which was a fixed difference), 116 associated *H. pachinus* with one parent or the other; 83 had similar allele frequencies between *H. pachinus* and *H. cydno* while 33 had similar allele frequencies between *H. pachinus* and *H. melpomene* (table 1). The remaining 57 loci did not associate *H. pachinus* with a particular parent because the allele frequencies in *H. pachinus* were distinct from (35 loci) or not different from (22 loci) both parents. The control analyses, in which each of the parental species was treated as a potential hybrid, resulted in distributions very similar to the experimental analysis of *H. pachinus* (table 1). Correcting for multiple comparisons in all steps reduced the number of AIMs in each analysis but resulted in qualitatively similar distributions.

(b) DNA sequence data suggest no detectable contribution of *H. melpomene* in the origin of *H. pachinus*

We analysed a total of 557 *H. cydno*, *H. pachinus*, *H. melpomene* and *H. hecale* haplotypes from 16 loci (table 1 in electronic supplementary material). Neighbour-joining gene genealogies revealed substantially shared genetic variation among species. For instance, for none of

Table 1. Distribution of AFLP loci that associate the putative hybrid species *H. pachinus* with each of the suspected parental species, *H. cydno* and *H. melpomene*. (A locus-by-locus AMOVA identified AFLP loci with significant ($p < 0.05$) frequency differences between *H. cydno* and *H. melpomene*. Subsequent locus-by-locus AMOVAs between *H. pachinus* and each of the parental species allowed each locus to be assigned to one of the four categories; those that associated *H. pachinus* with *H. cydno*, those that associated *H. pachinus* with *H. melpomene*, those that were distinct in *H. pachinus* and those for which *H. pachinus* was not different from either putative parent. As controls, the same analysis was performed treating *H. cydno* and *H. melpomene* as potential hybrids between the other two species. Numbers in parentheses are after Bonferroni correction for multiple comparisons.)

treatment	hybrid	parent A	parent B	no. of loci distinguishing parents	loci linking hybrid to parent A	loci linking hybrid to parent B	loci distinct in hybrid	loci not distinct in hybrid
experimental	<i>pachinus</i>	<i>cydno</i>	<i>melpomene</i>	173 (62)	83 (43)	33 (7)	35 (2)	22 (10)
control 1	<i>cydno</i>	<i>pachinus</i>	<i>melpomene</i>	171 (73)	82 (45)	35 (10)	35 (3)	19 (15)
control 2	<i>melpomene</i>	<i>cydno</i>	<i>pachinus</i>	131 (23)	35 (5)	33 (7)	33 (4)	30 (7)

the 16 loci did haplotypes from all four species form monophyletic clades (figure 1 in electronic supplementary material). However, for most loci, there was a tendency for *H. pachinus* haplotypes to group with those from *H. cydno*. This association between *H. pachinus* and *H. cydno* was supported by locus-specific F_{ST} estimates; 15 out of 16 loci (all except *white*) exhibited significant differentiation among species and for all of these the level of differentiation between *H. cydno* and *H. pachinus* was less than that between *H. melpomene* and *H. pachinus* (table 2).

(c) Analyses of wingless do not support the hybrid-origin hypothesis for *H. pachinus*

We explicitly tested alternative evolutionary hypotheses for the gene *wingless* because it is tightly linked to the locus that controls the white versus yellow forewing colour. The $\ln(L)$ of the phylogeny in which *H. pachinus* haplotypes were constrained to form a clade with those of *H. cydno* (which was the best tree) was -2526.73 , while that for the tree in which *H. pachinus* haplotypes were constrained to form a clade with those of *H. melpomene* was -2547.70 (figure 3 in electronic supplementary material). Bayes factor ($2 \ln B_{10}$) comparing these hypotheses was thus 41.94, which is indicative of very strong support for the tree grouping *H. pachinus* haplotypes with those from *H. cydno* versus the tree linking *H. pachinus* haplotypes with those from *H. melpomene*.

(d) The genome of *H. heurippa* may be a mosaic

By combining our *Dll* and *inv* sequence data for *H. cydno*, *H. pachinus*, *H. melpomene* and *H. hecale* from Costa Rica with 32 *Dll* and 42 *inv* sequences for *H. heurippa*, *H. cydno* and *H. melpomene* from Colombia, we found different distributions of genetic variation between the two putative hybrid species. Unlike *H. pachinus*, which shared only limited genetic variation with *H. melpomene* across 16 loci, *H. heurippa* exhibited a discordant pattern consistent with the hybrid-origin hypothesis. At *Dll*, the majority of *H. heurippa* haplotypes grouped with those from *H. melpomene*, while at *inv*, the majority grouped with *H. cydno* (figure 2). F_{ST} estimates supported this conclusion with less genetic differentiation between *H. heurippa* and *H. melpomene* at *Dll* and less differentiation between *H. heurippa* and *H. cydno* at *inv* (table 2 in electronic supplementary material).

4. DISCUSSION

Our analyses of two independent genomic datasets failed to reveal any robust evidence for genomic mosaicism in the putative hybrid butterfly species, *H. pachinus*. In fact, only the analysis in which we grouped AFLP loci according to whether they associated *H. pachinus* with *H. cydno* or *H. melpomene* provided results that were consistent with *H. pachinus* being a genetic mixture. That 72% of 116 AIMs associated *H. pachinus* with *H. cydno* while 28% associated *H. pachinus* with *H. melpomene* is remarkably consistent with the theoretical expectations under the hybrid-origin hypothesis (65–75% from *H. cydno* and 25–35% from *H. melpomene*). However, the control analyses showed that a similar distribution results regardless of which species is treated as a hybrid. This finding underscores the substantially shared genetic variation among all three of these species and it highlights the importance of proper controls in comparative genomic analyses of this type.

It is important to note that our finding of no genomic mosaicism in *H. pachinus* does not definitively answer the question of whether this species originated due to hybridization. This is primarily because hybridization and introgression following the origin of *H. pachinus* could have erased virtually all evidence of a hybrid origin. For instance, recent work has shown that *H. cydno*, *H. melpomene* and *H. pachinus* hybridize and exchange genes (Bull et al. 2006; Kronforst et al. 2006a), with interspecific gene flow between *H. cydno* and *H. pachinus* being particularly extensive. Coalescent simulations based on the isolation with migration model of speciation (Hey & Nielsen 2004) indicate that the population migration rate ($2Nm$) from *H. cydno* to *H. pachinus* is approximately 0.5 while that from *H. pachinus* to *H. cydno* is very large at approximately 4.3 (Kronforst et al. 2006a). Such extensive ongoing hybridization has the potential to largely homogenize the genomes of *H. cydno* and *H. pachinus* over time and thereby erase the genome-wide signature of hybrid speciation sought in this study.

Similarly, our analysis of the evolutionary history of the gene *wingless* does not absolutely preclude the possibility that the yellow forewing colour of *H. pachinus* originated from *H. melpomene*. The gene *wingless* is tightly linked to the locus that controls the white versus yellow forewing colour in *H. cydno* and *H. pachinus* (Kronforst et al. 2006b) and it is an excellent candidate gene for this trait. On the developing forewings of another nymphalid butterfly, *Junonia coenia*, *wingless* is expressed in bands that eventually contain the

Table 2. Pairwise genetic differentiation among *H. pachinus*, *H. cydno*, *H. melpomene* and *H. hecale* from Costa Rica at 16 sequenced loci. (Comparisons between the putative hybrid species, *H. pachinus*, and each of the suspected parental species, *H. cydno* and *H. melpomene*, are listed first for each locus. Of these two, the comparison exhibiting the least differentiation is given in bold. There was no differentiation among species at the locus *white* (*w*), so no comparison is given in bold.)

locus	comparison	F_{ST}	probability
<i>ap</i>	<i>cydno</i> versus <i>pachinus</i>	0.246	0.021
	<i>melpomene</i> versus <i>pachinus</i>	0.479	0.000
	<i>cydno</i> versus <i>melpomene</i>	0.546	0.000
	<i>cydno</i> versus <i>hecale</i>	0.950	0.031
	<i>pachinus</i> versus <i>hecale</i>	0.886	0.031
	<i>melpomene</i> versus <i>hecale</i>	1.000	0.024
	overall	0.785	0.000
<i>ci</i>	<i>cydno</i> versus <i>pachinus</i>	− 0.074	0.854
	<i>melpomene</i> versus <i>pachinus</i>	0.107	0.096
	<i>cydno</i> versus <i>melpomene</i>	0.129	0.041
	<i>cydno</i> versus <i>hecale</i>	0.100	0.104
	<i>pachinus</i> versus <i>hecale</i>	0.122	0.102
	<i>melpomene</i> versus <i>hecale</i>	0.117	0.128
	overall	0.090	0.027
<i>cn</i>	<i>cydno</i> versus <i>pachinus</i>	− 0.008	0.625
	<i>melpomene</i> versus <i>pachinus</i>	0.208	0.078
	<i>cydno</i> versus <i>melpomene</i>	0.357	0.029
	<i>cydno</i> versus <i>hecale</i>	0.760	0.999
	<i>pachinus</i> versus <i>hecale</i>	0.729	0.999
	<i>melpomene</i> versus <i>hecale</i>	0.864	0.999
	overall	0.443	0.002
<i>CO</i>	<i>cydno</i> versus <i>pachinus</i>	0.686	0.000
	<i>melpomene</i> versus <i>pachinus</i>	0.911	0.000
	<i>cydno</i> versus <i>melpomene</i>	0.852	0.000
	<i>cydno</i> versus <i>hecale</i>	0.910	0.000
	<i>pachinus</i> versus <i>hecale</i>	0.954	0.000
	<i>melpomene</i> versus <i>hecale</i>	0.973	0.003
	overall	0.818	0.000
<i>Dll</i>	<i>cydno</i> versus <i>pachinus</i>	0.109	0.000
	<i>melpomene</i> versus <i>pachinus</i>	0.500	0.000
	<i>cydno</i> versus <i>melpomene</i>	0.485	0.000
	<i>cydno</i> versus <i>hecale</i>	0.757	0.000
	<i>pachinus</i> versus <i>hecale</i>	0.771	0.000
	<i>melpomene</i> versus <i>hecale</i>	0.429	0.000
	overall	0.542	0.000
<i>dpp</i>	<i>cydno</i> versus <i>pachinus</i>	− 0.031	0.711
	<i>melpomene</i> versus <i>pachinus</i>	0.083	0.445
	<i>cydno</i> versus <i>melpomene</i>	0.006	0.726
	<i>cydno</i> versus <i>hecale</i>	0.941	0.024
	<i>pachinus</i> versus <i>hecale</i>	0.760	0.006
	<i>melpomene</i> versus <i>hecale</i>	0.946	0.005
	overall	0.703	0.000
<i>Ef1α</i>	<i>cydno</i> versus <i>pachinus</i>	0.345	0.037
	<i>melpomene</i> versus <i>pachinus</i>	0.764	0.009
	<i>cydno</i> versus <i>melpomene</i>	0.678	0.015
	<i>cydno</i> versus <i>hecale</i>	0.671	0.063
	<i>pachinus</i> versus <i>hecale</i>	0.852	0.050
	<i>melpomene</i> versus <i>hecale</i>	0.706	0.055
	overall	0.679	0.000
<i>en</i>	<i>cydno</i> versus <i>pachinus</i>	− 0.179	0.848
	<i>melpomene</i> versus <i>pachinus</i>	0.462	0.008
	<i>cydno</i> versus <i>melpomene</i>	0.512	0.009
	<i>cydno</i> versus <i>hecale</i>	0.751	0.004
	<i>pachinus</i> versus <i>hecale</i>	0.724	0.001
	<i>melpomene</i> versus <i>hecale</i>	0.926	0.000
	overall	0.671	0.000
<i>inv</i>	<i>cydno</i> versus <i>pachinus</i>	0.448	0.002
	<i>melpomene</i> versus <i>pachinus</i>	0.631	0.000
	<i>cydno</i> versus <i>melpomene</i>	0.595	0.000
	<i>cydno</i> versus <i>hecale</i>	0.478	0.000
	<i>pachinus</i> versus <i>hecale</i>	0.509	0.000

(Continued.)

Table 2. (Continued.)

locus	comparison	F_{ST}	probability
<i>Mpi</i>	<i>melpomene</i> versus <i>hecale</i>	0.567	0.000
	overall	0.537	0.000
	<i>cydno</i> versus <i>pachinus</i>	0.188	0.040
	<i>melpomene</i> versus <i>pachinus</i>	0.365	0.004
	<i>cydno</i> versus <i>melpomene</i>	-0.045	0.575
	<i>cydno</i> versus <i>hecale</i>	0.193	0.081
	<i>pachinus</i> versus <i>hecale</i>	0.502	0.000
	<i>melpomene</i> versus <i>hecale</i>	0.258	0.015
<i>ptc</i>	overall	0.259	0.002
	<i>cydno</i> versus <i>pachinus</i>	0.218	0.011
	<i>melpomene</i> versus <i>pachinus</i>	0.832	0.002
	<i>cydno</i> versus <i>melpomene</i>	0.620	0.004
	<i>cydno</i> versus <i>hecale</i>	0.877	0.000
	<i>pachinus</i> versus <i>hecale</i>	0.964	0.000
	<i>melpomene</i> versus <i>hecale</i>	0.926	0.002
	overall	0.850	0.000
<i>sd</i>	<i>cydno</i> versus <i>pachinus</i>	0.097	0.132
	<i>melpomene</i> versus <i>pachinus</i>	0.485	0.000
	<i>cydno</i> versus <i>melpomene</i>	0.216	0.000
	<i>cydno</i> versus <i>hecale</i>	0.659	0.000
	<i>pachinus</i> versus <i>hecale</i>	0.874	0.000
	<i>melpomene</i> versus <i>hecale</i>	0.874	0.000
	overall	0.690	0.000
<i>st</i>	<i>cydno</i> versus <i>pachinus</i>	0.019	0.280
	<i>melpomene</i> versus <i>pachinus</i>	0.260	0.009
	<i>cydno</i> versus <i>melpomene</i>	0.126	0.095
	<i>cydno</i> versus <i>hecale</i>	0.722	0.000
	<i>pachinus</i> versus <i>hecale</i>	0.787	0.000
	<i>melpomene</i> versus <i>hecale</i>	0.655	0.000
	overall	0.591	0.000
<i>Tpi</i>	<i>cydno</i> versus <i>pachinus</i>	- 0.010	0.306
	<i>melpomene</i> versus <i>pachinus</i>	0.420	0.000
	<i>cydno</i> versus <i>melpomene</i>	0.354	0.000
	<i>cydno</i> versus <i>hecale</i>	0.867	0.004
	<i>pachinus</i> versus <i>hecale</i>	0.880	0.005
	<i>melpomene</i> versus <i>hecale</i>	0.962	0.004
	overall	0.682	0.000
<i>w</i>	<i>cydno</i> versus <i>pachinus</i>	-0.004	0.637
	<i>melpomene</i> versus <i>pachinus</i>	-0.074	0.503
	<i>cydno</i> versus <i>melpomene</i>	0.125	0.088
	<i>cydno</i> versus <i>hecale</i>	0.174	0.093
	<i>pachinus</i> versus <i>hecale</i>	-0.034	0.418
	<i>melpomene</i> versus <i>hecale</i>	0.066	0.224
	overall	0.044	0.253
<i>wg</i>	<i>cydno</i> versus <i>pachinus</i>	0.277	0.002
	<i>melpomene</i> versus <i>pachinus</i>	0.680	0.003
	<i>cydno</i> versus <i>melpomene</i>	0.578	0.000
	<i>cydno</i> versus <i>hecale</i>	0.503	0.000
	<i>pachinus</i> versus <i>hecale</i>	0.616	0.003
	<i>melpomene</i> versus <i>hecale</i>	0.679	0.005
	overall	0.534	0.000

ommochrome pigment ommatin D (Carroll *et al.* 1994; Nijhout 1997). The yellow pigment on *Heliconius* wings is also an ommochrome, while the white colour is structural (Gilbert *et al.* 1988). Thus, a gene with ommochrome-associated expression on the forewing of one nymphalid butterfly is also perfectly linked to an ommochrome on/off switch locus on the forewing of another nymphalid butterfly. Together, these observations suggest that *wingless* may be the *Heliconius* forewing colour locus (Kronforst *et al.* 2006b). If this is true, the hybrid-origin hypothesis predicts *Heliconius pachinus wingless* haplotypes to have originated from *H. melpomene* haplotypes rather than

H. cydno haplotypes. In contrast, our analyses strongly favour an evolutionary scenario in which *wingless* haplotypes from *H. pachinus* group with those from *H. cydno* and thus support a non-hybrid origin for *H. pachinus*. However, it remains possible that *wingless* is the gene responsible for wing colour and the yellow colour of *H. pachinus* was in fact derived from *H. melpomene*, but recombination between *H. pachinus* and *H. cydno* obscured this phylogenetic signal along the entire gene except the actual sequence variation responsible for the colour difference. Future identification of the nucleotide variation actually responsible for the colour switch will truly resolve this question.

Figure 2. Neighbour-joining gene trees for two loci: (a) *Distal-less* and (b) *invected*. Each tree is based on DNA sequences from two putative hybrid species, *H. pachinus* and *H. heurippa*, two races of each of the suspected parental species, *H. cydno* and *H. melpomene*, and the closely related outgroup, *H. hecale*. Unlike *H. pachinus*, which exhibits a genetic association with *H. cydno* across loci, *H. heurippa* exhibits a discordant pattern in which the majority of *Dll* haplotypes group with those from *H. melpomene*, while the majority of *inv* haplotypes group with *H. cydno*.

It is also important to note that just as introgression has the potential to erase evidence of a hybrid origin, it also has the potential to produce patterns of shared genetic variation that can appear consistent with hybrid speciation. This is especially true with limited genetic data. Hence, the preliminary data available for *H. heurippa* should not be over-interpreted. The observations that *H. heurippa* shares genetic variation with both *H. cydno* and *H. melpomene* and that the two loci show a discordant pattern in which *H. heurippa* is more similar to *H. melpomene* at *Dll* but more similar to *H. cydno* at *inv* are consistent with a recent hybrid origin for *H. heurippa* (Mavarez *et al.* 2006). Alternatively, it is possible that *H. heurippa* originated without a genetic contribution from *H. melpomene* but has since experienced variable rates of interspecific gene flow at the two loci. It is also possible that the observed patterns are simply a result of ancestral genetic variation that is sorting out differently in different descendent lineages. However, it is perhaps telling that with data for just two loci, *H. heurippa* exhibits the discordance among loci predicted to result from hybrid speciation. *Heliconius pachinus*, on the other hand, does not exhibit any obvious discordance among loci even with data from 16 genes. While the *H. heurippa* genetic data are limited, the contrast with the substantial data available for *H. pachinus* strengthens the argument for a hybrid origin for *H. heurippa*. As with *H. pachinus*, analyses of the

evolutionary relationships among species at the actual genes that control the colour pattern elements believed to have originated from *H. cydno* and *H. melpomene* will be the real test of whether *H. heurippa* originated via hybridization.

We thank Ulrich Mueller, Joan Strassmann and Dave Queller for use of laboratory facilities, Laura Young and Lauren Blume for their assistance in the laboratory, Kenny Kronforst and Andres Vega for their assistance in the field, and Chris Jiggins, Andrew Brower and two anonymous reviewers for their comments on the manuscript. Butterflies were collected under permits provided by Costa Rica's Ministerio del Ambiente y Energia. This work was funded by Organization for Tropical Studies' Graduate Research Fellowship and National Science Foundation grants DEB 0206613 and DEB 0415718. M.L. and C.S. were funded by the Fondo Colombiano de Investigaciones Cientificas y Proyectos Especiales Francisco Jose de Caldas COLCIENCIAS grants 1204-05-11414 and 7155-CO, Banco de la Republica and private donations from Continautos S. A., Proficol El Carmen S. A., Didacol S. A. and F. Arango, Colombia.

REFERENCES

- Beltrán, M., Jiggins, C. D., Bull, V., Linares, M., Mallet, J., McMillan, W. O. & Bermingham, E. 2002 Phylogenetic discordance at the species boundary: comparative gene genealogies among *Heliconius* butterflies. *Mol. Biol. Evol.* **19**, 2176–2190.

- Brower, A. V. Z. 1994 Phylogeny of *Heliconius* butterflies inferred from mitochondrial DNA sequences (Lepidoptera: Nymphalidae). *Mol. Phylogenet. Evol.* **3**, 159–174. (doi:10.1006/mpev.1994.1018)
- Brower, A. V. Z. 1996a A new mimetic species of *Heliconius* (Lepidoptera: Nymphalidae), from southeastern Colombia, revealed by cladistic analysis of mitochondrial DNA sequences. *Zool. J. Linn. Soc.* **116**, 317–332. (doi:10.1006/zjls.1996.0022)
- Brower, A. V. Z. 1996b Parallel race formation and the evolution of mimicry in *Heliconius* butterflies: a phylogenetic hypothesis from mitochondrial DNA sequences. *Evolution* **50**, 195–221. (doi:10.2307/2410794)
- Brower, A. V. Z. & Egan, M. G. 1997 Cladistic analysis of *Heliconius* butterflies and relatives (Nymphalidae: Heliconiini): a revised phylogenetic position for *Eueides* based on sequences from mtDNA and a nuclear gene. *Proc. R. Soc. B* **264**, 969–977. (doi:10.1098/rspb.1997.0134)
- Buerkle, C. A., Morris, R. J., Asmussen, M. A. & Rieseberg, L. H. 2000 The likelihood of homoploid hybrid speciation. *Heredity* **84**, 441–451. (doi:10.1046/j.1365-2540.2000.00680.x)
- Bull, V., Beltran, M., Jiggins, C. D., McMillan, W. O., Bermingham, E. & Mallet, J. 2006 Polyphyly and gene flow between non-sibling *Heliconius* species. *BMC Biol.* **4**, 11. (doi:10.1186/1741-7007-4-11)
- Carroll, S. B., Gates, J., Keys, D. N., Paddock, S. W., Panganiban, G. E. F., Seale, J. E. & Williams, J. A. 1994 Pattern formation and eyespot determination in butterfly wings. *Science* **265**, 109–114. (doi:10.1126/science.7912449)
- Coyne, J. A. & Orr, H. A. 2004 *Speciation*. Sunderland, MA: Sinauer Associates.
- Dowling, T. E. & Secor, C. L. 1997 The role of hybridization and introgression in the diversification of animals. *Annu. Rev. Ecol. Syst.* **28**, 593–619. (doi:10.1146/annurev.ecolsys.28.1.593)
- Falush, D., Stephens, M. & Pritchard, J. K. 2003 Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587.
- Flanagan, N. S., Tobler, A., Davison, A., Pybus, O. G., Kapan, D. D., Planas, S., Linares, M., Heckle, D. & McMillan, W. O. 2004 Historical demography of Müllerian mimicry in the neotropical *Heliconius* butterflies. *Proc. Natl Acad. Sci. USA* **101**, 9704–9709. (doi:10.1073/pnas.0306243101)
- Gilbert, L. E. 2003 Adaptive novelty through introgression in *Heliconius* wing patterns: evidence for shared genetic “tool box” from synthetic hybrid zones and a theory of diversification. In *Ecology and evolution taking flight: butterflies as model systems* (eds C. L. Boggs, W. B. Watt & P. R. Ehrlich), pp. 281–318. Chicago, IL: University of Chicago Press.
- Gilbert, L. E., Forrest, H. S., Schultz, T. D. & Harvey, D. J. 1988 Correlations of ultrastructure and pigmentation suggest how genes control development of wing scales in *Heliconius* butterflies. *J. Res. Lepid.* **26**, 141–160.
- Gompert, Z., Fordyce, J. A., Forister, M. L., Shapiro, A. M. & Nice, C. C. 2006 Homoploid hybrid speciation in an extreme habitat. *Science* **314**, 1923–1925. (doi:10.1126/science.1135875)
- Hey, J. & Nielsen, R. 2004 Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167**, 747–760. (doi:10.1534/genetics.103.024182)
- Hintz, J. 2001 *NCSS and PASS*. Kaysville, UT: Number Cruncher Statistical Systems.
- Huelsenbeck, J. P. & Ronquist, F. 2001 MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755. (doi:10.1093/bioinformatics/17.8.754)
- Kass, R. E. & Raftery, A. E. 1995 Bayes factors. *J. Am. Stat. Assoc.* **90**, 773–795. (doi:10.2307/2291091)
- Kronforst, M. R. 2005 Primers for the amplification of nuclear introns in *Heliconius* butterflies. *Mol. Ecol. Notes* **5**, 158–162. (doi:10.1111/j.1471-8286.2004.00873.x)
- Kronforst, M. R., Young, L. E., Blume, L. M. & Gilbert, L. E. 2006a Multilocus analyses of admixture and introgression among hybridizing *Heliconius* butterflies. *Evolution* **60**, 1254–1268. (doi:10.1554/06-005.1)
- Kronforst, M. R., Young, L. G., Kapan, D. D., McNeely, C., O'Neill, R. J. & Gilbert, L. E. 2006b Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. *Proc. Natl Acad. Sci. USA* **103**, 6575–6580. (doi:10.1073/pnas.0509685103)
- Kumar, S., Tamura, K., Jakobsen, I. B. & Nei, M. 2001 *MEGA2: molecular evolutionary genetics analysis software*. Tempe, AZ: Arizona State University.
- Levin, D. A. 2002 *The role of chromosomal change in plant evolution*. New York, NY: Oxford University Press.
- Linares, M. 1989 Adaptive microevolution through hybridization and biotic destruction in the neotropics. Ph.D. thesis. University of Texas, Austin.
- Mallet, J., McMillan, W. O. & Jiggins, C. D. 1998 Mimicry and warning color at the boundary between races and species. In *Endless forms: species and speciation* (eds D. J. Howard & S. H. Berlocher), pp. 390–403. New York, NY: Oxford University Press.
- Mallet, J., Neukirchen, W. & Linares, M. 2006 Wild-caught hybrids among *Heliconius* and *Eueides* species: a database. See <http://www.ucl.ac.uk/taxome/hyb/hybtabs.html>.
- Mavarez, J., Salazar, C. A., Bermingham, E., Salcedo, C., Jiggins, C. D. & Linares, M. 2006 Speciation by hybridization in *Heliconius* butterflies. *Nature* **441**, 868–871. (doi:10.1038/nature04738)
- Meyer, A., Salzburger, W. & Scharl, M. 2006 Hybrid origin of a swordtail species (Teleostei: *Xiphophorus clemenciae*) driven by sexual selection. *Mol. Ecol.* **15**, 721–730. (doi:10.1111/j.1365-294X.2006.02810.x)
- Nijhout, H. F. 1997 Ommochrome pigmentation of the *linea* and *rosa* seasonal forms of *Precis coenia* (Lepidoptera: Nymphalidae). *Arch. Insect Biochem. Physiol.* **36**, 215–222. (doi:10.1002/(SICI)1520-6327(1997)36:3<215::AID-ARCH5>3.0.CO;2-S)
- Nylander, J. A., Ronquist, F., Huelsenbeck, J. P. & Nieves-Aldrey, J. A. 2004 Bayesian phylogenetic analysis of combined data. *Syst. Biol.* **53**, 47–67. (doi:10.1080/10635150490264699)
- Pritchard, J. K., Stephens, M. & Donnelly, P. 2000 Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.
- Rieseberg, L. H. 1997 Hybrid origins of plant species. *Annu. Rev. Ecol. Syst.* **28**, 359–389. (doi:10.1146/annurev.ecolsys.28.1.359)
- Salazar, C., Jiggins, C. D., Arias, C. F., Tobler, A., Bermingham, E. & Linares, M. 2005 Hybrid incompatibility is consistent with a hybrid origin of *Heliconius heurippa* Hewitson from its close relatives, *Heliconius cydno* Doubleday and *Heliconius melpomene* Linnaeus. *J. Evol. Biol.* **18**, 247–256. (doi:10.1111/j.1420-9101.2004.00839.x)
- Salzburger, W., Baric, S. & Sturmbauer, C. 2002 Speciation via introgressive hybridization in east African cichlids? *Mol. Ecol.* **11**, 619–625. (doi:10.1046/j.0962-1083.2001.01438.x)
- Schneider, S., Roessli, D. & Excoffier, L. 2000 *ARLEQUIN: a software for population genetics data analysis*. Ver 2.0. Genetics and Biometry Lab, Department of Anthropology, University of Geneva.

- Schwarz, D., Matta, B. M., Shakir-Botteri, N. L. & McPherson, B. A. 2005 Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature* **436**, 546–549. (doi:10.1038/nature03800)
- Smith, P. F., Konings, A. & Kornfield, I. 2003 Hybrid origin of a cichlid population in Lake Malawi: implications for genetic variation and species diversity. *Mol. Ecol.* **12**, 2497–2504. (doi:10.1046/j.1365-294X.2003.01905.x)
- Vos, P. *et al.* 1995 AFLP: a new technique for DNA fingerprinting. *Nucl. Acids Res.* **23**, 4407–4414. (doi:10.1093/nar/23.21.4407)
- Wahlberg, N., Weingartner, E. & Nylin, S. 2003 Towards a better understanding of the higher systematics of Nymphalidae (Lepidoptera: Papilionoidea). *Mol. Phylogenet. Evol.* **28**, 473–484. (doi:10.1016/S1055-7903(03)00052-6)