EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

Wing patterning genes and coevolution of Müllerian mimicry in Heliconius butterflies: support from phylogeography, co-phylogeny and divergence times

Journal:	*Evolution*
Manuscript ID	15-0530.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Hoyal Cuthill, Jennifer; University of Cambridge, Department of Earth Sciences Charleston, Michael; University of Tasmania, Mathematics & Physics
Keywords:	Adaptation, Coevolution, Morphological Evolution, Phylogenetics, Phylogeography

Wing patterning genes and coevolution of Müllerian mimicry in *Heliconius* butterflies: support from phylogeography, co-phylogeny and divergence times

Jennifer Hoyal Cuthill and Michael Charleston

Examples of long-term coevolution are rare among free-living organisms. Müllerian mimicry in *Heliconius* butterflies had been suggested as a key example of coevolution by early genetic studies. However, research over the last two decades has been dominated by the idea that the best studied co-mimics, *Heliconius erato* and *Heliconius melpomene*, did not coevolve at all. Recently sequenced genes associated with wing colour pattern phenotype offer a new opportunity to resolve this controversy. Here we test the hypothesis of coevolution between *H. erato* and *H. melpomene* using Bayesian multi-locus analysis of five colour pattern genes and five neutral genetic markers. We first explore the extent of phylogenetic agreement versus conflict between the different genes. Coevolution is then tested against three aspects of the mimicry diversifications: phylogenetic branching patterns, divergence times and, for the first time, phylogeographic histories. We show that all three lines of evidence are compatible with strict coevolution of the diverse mimicry wing patterns, contrary to some recent suggestions. Instead, these findings tally with a coevolutionary diversification driven primarily by the ecological force of Müllerian mimicry.

Coevolution is a powerful concept, describing a mechanism for reciprocal evolution between ecologically associated species (or populations) under mutualistic or competitive selection (Thompson 1989). Accordingly, coevolution is thought to underpin major

ecological and evolutionary processes including community structuring (Nuismer et al. 2012) and diversification. Indeed, it can be maintained that coevolution is the reason there are millions of uniquely specialised species, not thousands (Thompson 1994). In the broad sense, coevolution is an ecological and evolutionary association between different taxa. Arguably, during *strict* coevolution, selection and evolution should be reciprocal (Janzen 1980; Thompson 1989). A history of coevolution is often particularly evident in highly specialised interactions such as parasitism and endosymbiosis. However, the evolutionary importance of coevolution between free-living species is more contentious (Thompson 1989, 1994). One theoretical case of coevolution between free-living species is Müllerian mimicry (Müller 1879): convergent resemblance between species with protective adaptations, such as unpalatability and warning colours (Sherratt 2008). However, the real-world importance of coevolution in Müllerian mimicry has been highly controversial (Thompson 1994; Mallet 1999; Sherratt 2008). In the past it had been suggested that some of the best conditions for coevolution, and particularly strict coevolution, may occur between unpalatable Müllerian mimics (Joron and Mallet 1998), which receive mutual fitness benefits from mimicry (Müller 1879; Turner, 1981; Turner, 1983; Gilbert 1984; Sheppard et al. 1985; Turner 1987).

The spectacular Müllerian mimicry system of the Neotropical butterflies *Heliconius erato* and *Heliconius melpomene* (Fig. 1) is, therefore, an exceptional model for tests of coevolution (Gilbert 1984). Field studies have shown that the wing patterns of *Heliconius* butterflies are under strong purifying selection for mimicry, since hybrids with non-mimetic patterns are preferentially removed by predation (Kapan 2001). While these two species mimic each other with spectacular fidelity where they co-occur, each species is highly polymorphic and has diversified across South and Central America into

approximately thirty races, or "morphs", with different wing patterns (Fig. 1) (Quek et al. 2010). These morphs are partially reproductively isolated and may be in the process of speciation (Sheppard et al. 1985; Brower 1996). This prompts the question as to whether, and to what extent, these parallel diversifications involved coevolution (Brown, Sheppard and Turner 1974; Gilbert 1983; Sheppard et al. 1985).

During a long history of controversy over this question, several authors have suggested that the hypothesis of coevolution between *H. erato* and *H. melpomene* could be tested by comparing the phylogenetic branching patterns, divergence times and biogeographic histories of the co-mimic populations. Over the last two decades, several arguments against the possibility of coevolution between *H. erato* and *H. melpomene* have appeared incorporating one or more of these points (Brower 1996; Mallet et al. 1996; Flanagan et al. 2004; Quek et al. 2010). However, more recent findings have begun to bring this long held view into question again (see Hines et al. 2011; Hoyal Cuthill and Charleston 2012). Our study based on three neutral markers (Hoyal Cuthill and Charleston 2012) provided evidence that the first two arguments against parallel radiation (based on phylogenetic branching patterns and timing), at least, may be unfounded.

An important component of Müller's mimicry theory is the prediction that the fitness benefits of mimicry will be in inverse proportion to the square of the population size (Müller 1879; Mallet 1999). Overall, *H. erato* is approximately twice as abundant as *H. melpomene*, as supported by field observations (see Mallet, 1999) and effective population sizes estimated from genetic data (e.g. Hoyal Cuthill and Charleston 2012). With population size ratios of 2:1, Müllerian mimicry theory predicts fitness benefits in the ratio 1:4. In combination with related phylogenetic studies which suggested that *H. erato* radiated before *H. melpomene* (Brower 1994; Brower 1996), this overall difference in

population sizes led to the supposition that the evolution of mimicry between these species was relatively one-sided, with "advergence" of *H. melpomene* onto an earlier phenotypic diversification of *H. erato* (Mallet 1999).

However, differences in population size can bias phylogenetic estimates of divergence times (Wakely and Hey 1997). When population sizes (estimated from the genetic data) were taken into account (Hoyal Cuthill and Charleston 2012), using a phylogenetic coalescent model (Heled and Drummond 2010), the apparent time discrepancy was shown to be an artefact, since smaller populations may carry a smaller number of genetic differences, thereby appearing younger. Instead, our study of these neutral markers suggested that the diversifications of *H. erato* and *H. melpomene* occurred over approximately the same time period. Further to this, co-phylogeny analyses (based on the neutral genetic markers) suggested codivergence of major eastern and western biogeographic clades of co-mimic butterflies (see map in Fig. 1), compatible with a long history of coevolution between *H. erato* and *H. melpomene* (Brown, Sheppard and Turner 1974; Sheppard et al. 1985; Gilbert 1983).

There are also theoretical reasons why a difference in average population size between two species may not preclude the coevolution of Müllerian mimicry. In reality, we do not expect the course of biological evolution to be determined by computation of relative fitness, but simply by whether a new mutation produces a fitter warning pattern within the local population where that mutation occurs (Turner, 1981). Detailed field studies show that local population sizes for the two species are highly variable (Gilbert 1984) and the local abundance of *H. melpomene* may be lower than, similar to, or greater than, that of *H. erato* (Gilbert 1984), raising the possibility of hotspots for coevolution (Thompson and Cunningham 2002).

Until very recently (Hines et al. 2011), all molecular phylogenetic reconstructions for these butterflies (including those of our recent study) have been based on neutral markers (generally "housekeeping" genes and particularly the mitochondrial COI locus) (Brower 1996; Flanagan et al. 2004; Quek et al. 2010; Hoyal Cuthill and Charleston 2012). However, the latest studies have suggested that the evolutionary history of phenotypically distinct populations (such as the wing pattern morphs) may, instead, be most clearly shown by genes that are strongly linked to the phenotypic traits in question (such as the wing pattern genes themselves) (Hines et al. 2011; Supple et al. 2013). Consequently, five genes associated with wing colour pattern variation (that is, having shown significant genotype-by-phenotype associations), which have recently been sequenced across a range of co-mimic morphs (Hines et al. 2011), may offer an opportunity to resolve the controversies surrounding coevolution between *H. erato* and *H. melpomene*. These five genes include the homeobox transcription factor *optix* (Reed et al. 2011), which is strongly associated with red, yellow and white colour variation (Papa et al. 2013; Martin et al. 2014). The other genes, *kinesin*, *VanGogh* and *GPCR*, have also shown associations with colour pattern variation across hybrid zones (Counterman et al. 2010; Baxter et al. 2010), although weaker associations with wing pattern morphs in broader comparative analyses (Hines et al. 2011). These colour pattern genes appear to be mutually linked to some extent (Hines et al. 2011), although Supple et al. 2013 suggested that strong linkage disequilibrium may be restricted to comparatively narrow genomic regions asscoated with colour pattern phenotype. However, this new genetic data has so far been interpreted primarily within the non-coevolutionary paradigm dominant over the last two decades (Hines et al. 2011; Hill et al. 2013; Merrill et al. 2015) and quantitative predictions of coevolution, such as the extent of codivergence (Futuyma and Slatkin 1983; Gilbert 1984: Page 2003), have not yet been

tested.

One more factor which should be considered is the possibility of shared responses to environmental change. An early mixed abiotic-biotic model (proposed by Brown, Sheppard, Turner and colleagues e.g. see Brown, Sheppard and Turner 1974; Sheppard et al. 1985) suggested that diversification of mimicry had involved both coevolution in sympatry (between *H. erato* and *H. melpomene*) and allopatric isolation in glacial refugia (Haffer 1969) (promoting wing pattern diversification within each species). However, evidence from climate studies and the fossil pollen record is now thought to argue against major Neotropical forest fragmentation during the Quaternary (2.588 to 0.005 Mya) (Knapp and Mallet 2003). Further to this, recent computer modelling suggests that warning pattern diversity could be a stable outcome of Müllerian mimicry without vicariance (Sheratt 2006). This possibility was also acknowledged by the advocates of the refugium model, although these authors emphasised the greater likelihood of mimicry diversification if significant forest fragmentation did occur (Turner, 1981 p. 111; Sheppard et al. 1985 pp. 588-591).

Here, we test the hypothesis of coevolution between *H. erato* and *H. melpomene* using multi-locus coalescent phylogenetic reconstructions of five genes associated with wing colour patterns (as well as five neutral markers), incorporating co-phylogeny analyses, divergence time reconstruction and phylogeographic analysis. This test is based on the principle that coevolution will be best supported if the diversifications of the co-mimic colour morphs are compatible in each of their dimensions, that is, the phylogenetic pattern of diversification in phenotypes, time *and* space. This analysis sheds fresh light on the spectacular mimicry radiation of these butterflies and provides new support for coevolution driven by Müllerian mimicry.

Materials and Methods

Molecular data

Analyses were performed on a DNA sequence dataset of ten genes downloaded from Genbank (accession numbers and sampling locations, Table S1 of (Hines et al. 2011)). This dataset includes five genes associated with *Heliconius* wing colour patterns (*optix*, *bves*, *kinesin*, *GPCR* and *VanGogh*) (Hines et al. 2011) and five neutral markers (a mitochondrial, mt, locus covering cytochrome oxidase COI and COII, plus four nuclear genes: *SUMO*, *Suz12*, *2654* and *CAT*).

This dataset includes six *Heliconius* species: *H. erato*, *H. himera*, *H. clysonymus*, *H. melpomene*, *H. ismenius* and *H. numata*. Sequences of each gene were aligned using Muscle (Edgar 2004). The program Gblocks (Castresana 2000) was then used to remove gaps and poorly aligned regions, a process which has been shown to improve phylogenetic reconstruction (Talavera and Castresana 2007). In total, these sequences cover 127 butterfly specimens and 5898 DNA base pairs (concatenated alignment, Table S1).

Phylogenetic analyses

Preliminary analyses were conducted to determine the best parameters for Bayesian phylogenetic reconstruction. These used one Markov Chain Monte Carlo (MCMC) chain of length 100,000,000 with a 25% burn-in, while the main analyses were based on combined output from two independent chains of length 400,000,000 each with a burn-in of 50%. Bayes factor comparisons indicated very strong support for substitution model partitioning by gene, over mt/ nuclear/colour gene partitions (*2 loge B* = 178),

neutral/colour gene partitions (*2 loge B* = 335), or a single partition (*2 loge B* = 392). Therefore all population-level phylogenetic analyses were partitioned by gene. The program jModelTest (Posada 2008) was used to select the best fitting DNA substitution model under the Bayesian information criterion. For individual genes, these were: mt GTR+G, *SUMO* K80+G, *Suz12* SYM+I+G, *2654* SYM *CAT* HKY+I+G*, optix* K80+G, *bves* K80, *kinesin* HKY+G, *GPCR* SYM+I+G, and *VanGogh* GTR+I.

Phylogenetic constraint tests were used to establish the appropriate population unit for phylogenetic analysis by testing the monophyly of the morphs *H. e. hydara* and *H. m. melpomene*, for which eastern (Amazonian) and western (Caribbean) populations (sampled from French Guiana versus Colombia, Panama and Trinidad, respectively (see Hines et al. 2011)) have previously been found to be non-monophyletic (Brower 1996). This method, implemented in MrBayes (Ronquist and et al. 2012), compares estimated marginal likelihoods between phylogenies that meet a positive monophyly constraint and those that do not, using their Bayes factor (Kaas and Raftery 1995). Estimated marginal likelihoods were based on two MCMC chains of length 1,000,000, sampled every 100 generations, with a 25% burn-in.

Population phylogenies were reconstructed using the Bayesian multilocus coalescent method, implemented in *BEAST (Heled and Drummond 2010). Output was examined using the program Tracer (http://beast.bio.ed.ac.uk/Tracer) and phylogenies were visualised using FigTree (http://tree.bio.ed.ac.uk/software/figtree/). The tree topology was linked for the colour pattern genes, as appropriate for non-recombining loci, and the appropriate ploidy level (mitochondrial or nuclear) was specified. A Yule prior was used for the population tree branching process.

Phylogenetic signals from the different gene partitions were compared using

incongruence length difference (ILD) tests implemented in PAUP* (Swofford 2000) and phylogenetic network construction using SplitsTree 4 (Huson and Bryant 2006). Distributions of shared polymorphisms were identified using the program SITES (Hey and Wakely 1997).

The number of implied derivations of major wing pattern characteristics (red forewing band, yellow hindwing band, rays and blue iridescence) was calculated for each phylogenetic estimate using the retention index (RI) (Farris 1989). RI equals one if there is only one implied derivation of a characteristic and decreases towards zero as the implied number of derivations increases relative to the maximum.

Divergence dating

Bayes factor comparisons indicated very strong support for a relaxed, uncorrelated lognormal molecular clock, which allows substitution rates to vary between branches, as opposed to a strict molecular clock ($2 \log_e B = 370$), therefore this was used in the main analyses. Phylogenetic divergence times were calibrated using a mitochondrial DNA substitution rate of 0.01909 substitutions per site per million years, found to be the average across seven (mitochondrial and nuclear) genes in a recent fossil-calibrated analysis of the butterfly subfamily Papilioninae (Simonsen et al. 2011). For comparison, we also performed an analysis using Brower's rate of 0.0115 substitutions/site/My (Brower 1994).

For the multilocus analysis of all ten genes, the substitution rate was set for the mitochondrial reference locus and specified as a prior for the other loci (following Heled and Drummond 2010) allowing their rates to deviate from the reference rate according to the uncorrelated lognormal molecular clock. As genetic variation in the colour pattern

genes may be under selection, mutations may not accumulate in a clock-like manner. Comparing the standard deviation of relaxed molecular clock rates (Heled and Drummond 2010) estimated for all ten genes suggested that the neutral markers did indeed behave in a more clock-like manner (average st. dev. = 0.371) than the colour pattern genes (average st. dev. = 1.439). Therefore, we based our divergence date estimates on the analyses of all ten genes with the substitution rate set for the mitochondrial reference locus (rather than the colour pattern genes alone, for example). Substitutions/site/my estimated for the colour pattern genes were: *optix* 0.00525, *bves* 0.00285, *kinesin* 0.01344, *GPCR* 0.01635, and *VanGogh* 0.00536. The Bayesian tree set for this analysis (excluding the burn-in) was visualised using the program DensiTree (Bouckaert 2010) (Fig. 4). Since no fossil-based calibration points were available for the taxon set used in this study, node calibration was not possible.

Phylogeographic analyses

Phylogeographic analyses were conducted separately for *H. erato* and *H. melpomene* on sequences of all ten genes using the programs BEAST (Drummond and Rambaut 2007; Lemey et al. 2009; Lemey et al. 2010) and SPREAD (Bielejec et al. 2011) and visualised using Google Earth version 7.1.1.1580 (beta) (Google Inc.). This reconstruction method takes as input the gene sequences and geographic locations of sampled individuals (from Table S1 of Hines et al. 2011, and Quek et al. 2010) and simultaneously reconstructs both the phylogeny and a history of phylogeographic diffusion using a Bayesian statistical approach (Lemey et al. 2010). Each analysis used two independent MCMC chains of length 100,000,000 with a burn-in of 50%, a relaxed, uncorrelated lognormal molecular clock, and a coalescent tree prior.

Co-phylogeny analyses

First, TreeMap 3 (Charleston and Robertson 2002) was used to test for statistically significant correlations in pairwise phylogenetic distances between associated subtrees in co-mimic phylogenies (Hoyal Cuthill and Charleston 2012). This general test was suitable for both our population-level phylogenies (with one-to-one mimicry associations) and our specimen-level phylogeographic trees, which include multiple individuals from each mimicry complex (giving many-to-many associations). For the phylogeographic trees, this test was conducted first with all co-mimic specimens and, second, after collapsing monophyletic clades of each morph (after Hoyal Cuthill and Charleston 2012).

Second, co-phylogeny mapping in Jane 4 (Conow et al. 2010) was used to reconstruct a minimal-cost history for associated phylogenies, based on a cost regime for modelled coevolutionary events. In the context of mimicry, the recoverable events are codivergence (joint divergence of co-mimic lineages), duplication (divergence of a lineage on one "mimic" phylogeny without divergence of the associated lineage), model switch (a change in mimicry association) and loss (cessation of a mimicry association) (Hoyal Cuthill and Charleston 2012). We used the Jane 4 default cost regime (codivergence, zero; duplication, one; model switch, one; loss, one) and tested whether the number of codivergences was significantly higher than expected based on 1000 randomisations of the mimic phylogeny (Conow et al. 2010). Co-phylogeny mapping is asymmetric: treating one phylogeny (the "mimic" phylogeny) as dependent on the other (the "model" phylogeny). Therefore, for comparison we repeated all cophylogeny analyses with a reversed model-mimic relationship (Hoyal Cuthill and Charleston 2012).

Results

Phylogenetic signals from the colour pattern and neutral genes.

In general, the phylogenetic analyses suggest that the colour pattern genes (*optix*, *bves*, *kinesin*, *GPCR* and *VanGogh*) greatly increase our ability to recover clades that correspond to broad wing pattern types (Table 1), relative to the neutral markers (mt COI and COII, nuclear *SUMO*, *Suz12*, *2654* and *CAT*), as suggested previously by Hines et al. (2011). In particular these analyses indicate only a single evolutionary origin within each species for the red forewing band, the rayed phenotype (Sheppard et al. 1985; Jiggins and McMillan 1997; Hines et al. 2011) and blue iridescence (except Fig. 2B in which the number of derivations of iridescence is ambiguous), see Table 1; Fig. 2. We note, however, that the genetic loci associated with wing pattern do not simply group all similar phenotypes, regardless of species or biogeographic region. For example all phylogenetic analyses of the colour pattern loci were able to correctly group the morphs of *H. melpomene* and those of *H. erato*, in relation to each other and their outgroup species (respectively *H. ismenius* and *H. numata*, versus *H. clysonymus* (Hoyal Cuthill and Charleston 2012)) (e.g. Fig. S10). All of our phylogenetic analyses (based on colour, neutral and combined gene partitions), also recovered sub-clades that correspond to biogeographic subregion (Amazonian, Caribbean or Chaco-Parana (Morrone 2006)) (Figs. 1, 2, S1, S2, S10, S11). For the colour pattern gene phylogenies, the high retention indices inferred for both wing pattern characters (mean RI = 0.75) and biogeographic region (RI = 1) demonstrate that these large-scale biogeographic clades are not in major conflict with the phenotypic groupings (Table 1). However, the neutral genes do show a reduced ability to recover phenotypic groupings

within these biogeographic clades (Table 1, Fig. S1). These results support the recent suggestion that genes involved with colour pattern determination provide improved markers for the phylogenetic history of the wing pattern morphs (Hines et al. 2011; Supple et al. 2013). However, they also suggest that the wing pattern diversification occurred alongside a biogeographic radiation, with additional phenotypic signal (beyond the broad-scale biogeographic groupings) present in the colour pattern gene sequences (Table 1). These major biogeographic groupings include splits between eastern (Amazonian and possibly Chaco-Parana) and western (Caribbean) clades which has been suggested by a number of previous phylogenetic studies (e.g. Brower, 1994,1996; Hoyal Cuthill & Charleston, 2012; Nadeau et al. 2015) (compare Figs. 2, S10, S11).

While the population-level and morph-level coalescent phylogenies (Fig. 3, Figs. S1-S2) as well as our specimen-level phylogeographic reconstructions (Figs. S5 and S8) consistently recovered similar phenotypic and biogeographic clades, there was conflict regarding the first divergences within each species. In particular, these phylogenetic estimates (Fig. 3) differ as to whether the initial divergence split off a morph from the Chaco-Parana biogeographic subregion (followed by subsequent Amazonian and Caribbean divergences) or split the eastern morphs (Chaco-Parana plus Amazonian) from the Caribbean clade, the topology supported by the phylogeographic reconstructions (Figs. S5 and S8). This topological uncertainty has also been found by previous phylogenetic analyses (Quek et al. 2010) suggesting conflict inherent to the sampled genetic data, which is observed across different phylogenetic methodologies. Visualisation of the points of agreement versus conflict in the genetic data, using a splits network (Fig. 2) and a density representation of the regional population-level Bayesian

tree set (Fig. 4), also suggests that the position of the Chaco-Parana morphs remains somewhat uncertain, with alternative phylogenetic placements (either within an eastern clade or sister to all other morphs) each receiving some support (as discussed below).

All ten studied genes had some polymorphisms shared between two or more morphs. The average number of shared polymorphisms was relatively similar for the neutral genes from the mitochondrial (0.60) and nuclear (0.52) regions, but lower for the colour pattern genes (0.14). Shared polymorphisms may result from the incomplete sorting of ancestral polymorphisms, or from gene flow (Wang, Wakely and Hey 1997). However, the morphs with shared polymorphisms were often widely separated geographically (from different biogeographic subregions), suggesting incomplete lineage sorting (see also Brower 1994).

However, incongruence length difference (ILD) tests indicated statistically significant phylogenetic conflict between neutral and colour genes ($p = 0.01$) and for pairs of genes within and across these partitions (39 out of the 45 possible pairs that could be chosen from the 10 genes showed statistically significant incongruence, with $p < 0.004$ in each case). This indicates that phylogenetic histories differ somewhat between genes, both within and between the colour and neutral partitions, as expected based on standard coalescent models (22) (Fig. 2).

Incongruence length difference (ILD) tests indicated that phylogenetic histories differ somewhat between genes, both within and between the partitions of colour pattern versus neutral markers (after a global test confirmed significance, 39 out of the 45 possible pairwise comparisons, among the 10 genes, indicated statistically significant incongruence with $p < 0.004$ in each case) as expected based on standard coalescent models (Rosenberg and Nordborg 2002). These results support our use of a phylogenetic

multi-locus coalescence model (Heled and Drummond 2010) that explicitly accounts for individual gene histories as well as incomplete lineage sorting, and can accommodate low levels of gene flow (Zhang et al. 2011).

There was support for monophyly of Amazonian and Caribbean populations of *H. m. melpomene* from the colour genes (*2 log e Bayes factor* (*B*) = 161), though not the neutral genes (*2 log e B* = 473) or combined gene partitions (*2 log e B* = 395). This therefore represents a comparatively strong point of conflict between the colour and neutral gene partitions. All gene partitions were against monophyly of Amazonian and Caribbean *H. e. hydara* (*2 loge B*: colour 354, neutral 374, combined 705). These results could be interpreted as an indication that the Caribbean and Amazonian populations of this morph are genuinely non-monophyletic, a possibility which has been discussed previously (Brower 1994, 1996). However, the fact that additional genetic information (so far, from the colour pattern genes) has led to increased support for the monophyly of *H. m. melpomene* suggests that the phylogenetic positions of Caribbean and Amazonian populations of both morphs (*H. e. hydara* and *H. m. melpomene*) deserve further attention.

Congruence in phylogenetic branching patterns, timing and biogeography

Co-phylogeny correlation tests indicated statistically significant congruence in phylogenetic branching patterns for all estimated phylogeny pairs (Figs. 3, S1, S2, S5 and S6, Table 2). In particular, significant correlations were found for the root of every phylogenetic estimate tested, indicating significant congruence for the phylogenies overall, root to tips ($p \leq 0.001$ in each case, Table 2). This means that when we sample a monophyletic clade from one phylogeny, the co-mimics in the corresponding phylogeny

tend to be significantly more closely related than expected by chance. These congruent pairs of phylogenies include those based on the different gene partitions (colour, neutral and combined, as described above). This suggests that some possible differences between the phylogenies of *H. erato* and *H. melpomene* (such as the phylogenetic positions of the Chaco-Parana morphs) can be accommodated within the overall hypothesis of coevolution between the two species, contrary to some previous suggestions (Quek et al. 2010).

Co-phylogeny mapping indicated significantly more codivergence events than expected by chance for the population and morph phylogenies based on the colour pattern genes (with $p < 0.003$, Table 2) and for the majority of analyses based on the combined gene partitions (with $p < 0.037$, Table 2). For example, the population-level, colour pattern gene analysis with *H. erato* treated as the "model" phylogeny reconstructs eight codivergence events out of a possible total of eleven (as well as one mimicry "loss", and three "duplication" followed by "model switch" events) (Fig. S7B, with $p < 0.001$, Table 2). Where some phylogenetic estimates support contrasting phylogenetic positions for the Chaco-Parana morphs from the two species (e.g. Fig. 3A) this can still be incorporated within a significantly significant cophylogenetic association, for example with ancient associations between the eastern clades in *H. erato* and *H. melpomene*, with a subsequent mimicry switch establishing the Caribbean mimicry group (e.g. Fig. S7A). The combined analysis of all ten genes (that is, including all of the available genetic information) recovers identical branching patterns for the three major biogeographic clades (Chaco-Parana, Caribbean and Amazonian), Fig. 3C, which are compatible with ancient codivergence of these groups (Fig. S7K-L, $p < 0.037$, Table 2). However, the phenotypic groupings within these clades are

"scrambled" relative to the colour pattern gene phylogenies, requiring a number of mimicry loss and switch events if these were to be interpreted literally. Further to this, the phylogenies based solely on the neutral genetic markers did not show a statistically significant number of codivergences (Table 2). Analyses of the neutral genes show a reduced ability to recover broad phenotypic clades (as described above), diluting the signal of codivergence visible in the phylogenies of the colour pattern genes.

Overall, these results suggest that the colour pattern genes add new support for coevolution between *H. erato* and *H. melpomene*. We note that the details of co-phylogenetic mapping (Fig. S7) depend on the precise input topologies and will therefore be sensitive to phylogenetic uncertainty. However, the levels of codivergence suggested by the analyses of the colour pattern genes and combined genes are remarkably high, comparable to classic examples of codivergence such as the gopher-louse system (Hafner et al. 2003).

Our substitution rate-calibrated Bayesian phylogeny (Fig. 4), based on a population-level coalescent analysis of all available DNA sequences (from the colour and neutral genes), confirms that diversification of the wing pattern morphs occurred over similar time intervals in *H. erato* and *H. melpomene* (Hoyal Cuthill and Charleston 2012). Based on a DNA substitution rate of 0.01909 substitutions per site per million years per lineage, recently estimated for a closely related group (swallowtail butterflies, Simonsen et al. 2011), the ages for the MRCA of the studied morphs were estimated as follows: *H. erato*, 0.4024 mya with a 95% Bayesian credible interval (BCI) of [0.2555 to 0.5537]; *H. melpomene*, 0.4497 [0.1851, 0.7157] mya. This analysis therefore estimates overlapping diversifications of these species, running from approximately 450,000 years ago to the recent past (Fig. 4). These time estimates are similar to those from our

previous Bayesian coalescent phylogenies of the genus *Heliconius* conducted with comparable substitution rate priors (Hoyal Cuthill and Charleston 2012), which was based on a smaller gene sequence dataset (of three neutral markers) but a larger taxon sample of twenty-four *Heliconius* species (with rate-calibrated ages estimated at 0.35 Ma for *H. erato* and 0.49 Ma for *H. melpomene*). However, the addition of a secondary node calibration, which is not possible here given the included species, did increase estimated divergence ages in that study priors (Hoyal Cuthill and Charleston 2012). The incorporation of recently published mitochondrial sequences (Hill et al. 2013) for two Central American postman morphs, *H. erato demophoon* and *H. erato cruentus*, which extend beyond the geographic range of *H. melpomene* (e.g. see Sheppard et al. 1985), increases the estimated age of the split between eastern (Amazonian and Chaco-Parana) and western (Caribbean and Central American) populations of *H. erato* to 0.4966 [0.486, 0.5] mya (Fig. S8).

Taxon sampling across the colour pattern genes included in our dataset (Hines et al. 2011) was focussed on the key mimics *H. erato* and *H. melpomene* and covers a relatively small sample of outgroups (with 6 specimens from four additional *Heliconius* species, shown in Fig. 4). Consequently, this study is not best placed to estimate the timing of the deeper, inter-species splits within the *Heliconius*. That said, our estimated age for the divergence of *H. erato* and *H. clysonymus* at approximately 3 Ma (Fig. 4), for which a direct comparison is possible given the taxon overlap between studies, appears very similar to that for a recent, independent phylogeny of the wider *Heliconius*, reconstructed using a comparable coalescent method with *BEAST (Kozak et al. Fig. S3).

An additional analysis, conducted here, using Brower's much cited rate of 0.0115

subs/site/my/lineage (corresponding to 2.3% pairwise divergence/my) (Brower 1994; Papadopoulou et al. 2010) inferred older ages of origin for the mimetic morphs from approximately 700,000 years ago, though these were also similar for both species, with overlapping Bayesian credible intervals, (*H. erato* 0.6354 [0.4362, 0.8543] ma; *H. melpomene* 0.7218 [0.3027, 1.1634] ma). With this slower molecular clock rate, the age of the divergence between *H. erato* and *H. clysonymus* was estimated at 5.4574 [3.6012, 7.6953] ma.

Ages of phylogenetic divergence will depend on the selection of molecular data, methodologies and for Bayesian methods, the choice of priors (e.g. see Arias et al. 2014; Kozak et al, 2015). Consequently, estimates of absolute divergence time are likely to differ between analyses. However, the important point, in relation to the potential for coevolution between *H. erato* and *H. melpomene* (which is the question motivating this study), is that analyses incorporating new genetic data (Figs 3-4; see also Kozak et al. 2015) are compatible with parallel radiations of these co-mimics over the same time interval.

Our Bayesian phylogeographic reconstructions, which use the locations of the extant populations to estimate an historical process of biogeographic diffusion (Bielejec et al. 2011), indicate broadly similar biogeographic histories for *H. erato* and *H. melpomene* (Fig. 5). Significant features of these reconstructions include origination points for the studied morphs of both species in Amazonia (the Amazon rainforest region), followed by broadly parallel radiations across the Neotropics: northwest into the Caribbean, east/west within the Amazonian subregion, and southeast into the Chaco-Parana. Both species also show parallel sub-radiations on the western (Caribbean) versus eastern (Amazonian) slopes of the Andes; with the Caribbean populations diffusing

southwards from the northern Andes, while the Amazonian populations radiate north and south from more southerly starting points. There are also some, relatively minor, differences between the phylogeographic reconstructions for *H. erato* and *H. melpomene*. For example, a secondary migration, from the central Amazonian basin, onto the eastern slopes of the Andes is reconstructed for *H. erato microclea*, while its co-mimic *H. melpomene xenoclea* falls within the main eastern Andean sub-radiation, as summarised above (Fig. 5).

Discussion

Our results support suggestions that five recently identified genes, including *optix* (Reed et al. 2011), which are strongly associated with wing pattern phenotype (Hines et al. 2011; Papa et al. 2013; Martin et al. 2014), add significantly to our understanding of the spectacular phenotypic diversifications of *H. erato* and *H. melpomene* (Fig. 1). However, our comparisons of phylogenetic branching patterns, timing and biogeography argue against the non-coevolutionary paradigm, which has dominated over the past two decades (Brower 1996; Flanagan et al. 2004; Mallet et al. 1996; Quek et al. 2010) and within which the new genotype-to-phenotype results have generally been interpreted (Hines et al. 2011; Hill et al. 2013; Merrill et al. 2015). Instead, the colour pattern genes reveal with new resolution that parallel clades of wing pattern co-mimics evolved in both species (see also discussion in Hines et al. 2011, although Hines et al. argued against parallel diversification of *H. erato* and *H. melpomene* beyond the derivation of the rayed Amazonian morphs). Further to this, our results demonstrate that these mimetic wing pattern morphs have phylogenetic branching histories that are significantly more similar than expected by chance, that they diversified during

overlapping time intervals and have broadly similar phylogeographic histories. All three lines of evidence are compatible with a shared history of mimicry, incorporating codivergence between co-mimic populations, as predicted by an older hypothesis of coevolution (Brown et al. 1974; Sheppard et al. 1985). Interestingly, some parallel features of the molecular phylogenies of *H. erato* and *H. melpomene* (Fig. 3 and see also Hines et al. 2011) were suggested previously (Turner, 1981; 1983) based on a most parsimonious history of colour pattern evolution under the assumption that recessive alleles will generally be ancestral to dominant ones (Haldane's sieve). These features include the secondary derivation of the rayed pattern type from an ancestral postman form (Turner, 1981; 1983), which is also the most parsimonious phenotypic character state transition on the molecular phylogenies reconstructed from the colour pattern gene sequences (Table 1).

A history of codivergence has often been viewed as strong evidence for coevolution (Futuyma and Slatkin 1983; Gilbert 1984; Page 2003) at least in the broad sense of a consistent ecological and evolutionary association between different species or populations. However, it does not in itself prove reciprocal evolutionary change (Gilbert 1984), that is, coevolution in the strict sense (Janzen 1980; Thompson 1989). The nature of strict coevolution in Müllerian mimicry has been controversial, but might constitute either alternating evolution, or simultaneous convergent evolution, of similar wing patterns in two co-mimics (reviewed in Thompson 1994 p. 32). Study of the *H. erato* - *H. melpomene* mimicry system provides an insight into coevolution at the level of the phenotypic and biogeographic population, at which this process likely operates (Thompson 1994). However, the unavoidable uncertainty of divergence date estimation for this recent diversification (e.g. Fig. 4) makes detailed comparisons of divergence

dates for co-mimics within *H. erato* and *H. melpomene* extremely difficult (Hoyal Cuthill and Charleston 2012). Consequently, it may not be possible to determine whether each of the mimetic morphs was derived first in *H. erato* or in *H. melpomene*. As a result, direct temporal evidence to test a hypothesis of alternating coevolution (which would involve some initial derivations in *H. erato* and some in *H. melpomene*) against a hypothesis of approximately simultaneous convergence, may not be recoverable. However, the overlapping timescales estimated for these parallel phenotypic radiations (described above) are compatible with the overall hypothesis of strict coevolution (Hoyal Cuthill and Charleston 2012) (via either alternating or simultaneous derivations of mimetic wing patterns). Thus, the three potential lines of evidence (from phylogeny, timing and biogeography) demonstrate coevolution in the broad sense at least, and are compatible with coevolution in the strict sense, contrary to previous suggestions (Brower 1996; Flanagan et al. 2004; Quek et al. 2010).

Contrary to some previous suggestions (Quek et al. 2010; Hill et al. 2013), the parallel phylogeographic radiations we reconstruct suggest that the studied morphs of both *H. erato* and *H. melpomene* originated in Amazonia, one of the most important generators of biodiversity (Hoorn et al. 2010). This illustrates that explicit modelling of a biogeographic radiation (here conducted using a phylogeographic diffusion model (Bielejec et al. 2011)) can lead to conclusions quite different to those based on the current location of certain extant populations (including those thought to be phylogenetically "basal" e.g. Brower 1996; Quek et al. 2010). Since extant species or populations are not necessarily representative of an ancestral location (or phenotype) (Crisp and Cook 2005) an evolutionary hypothesis (e.g. congruence among biogeographic histories) should be tested against an explicit evolutionary

reconstruction, such as the phylogeographic analysis presented here. These phylogeographic reconstructions argue against a Chaco-Parana origin of the studied morphs of *H. melpomene*, as had been previously suggested (Quek et al. 2010) based on the modern location of *H. m. nanna*, which may have diverged early in the *H. melpomene* radiation (Fig. 2 see also Quek et al. 2010). Instead when our phylogenetic (Figs. 2-4) and phylogeographic (Fig. 5) results are considered alongside the modern biogeographic distribution of phenotypes (Fig. 1), the consistent picture is one of a parallel phylogeographic and phenotypic mimicry radiation of *H. erato* and *H. melpomene*. These reconstructions suggest that ancestral forms of each species, possibly with a "postman" pattern (combining black, with red, yellow or both: Turner, 1981; Turner, 1983; Sheppard et al. 1985; Mallet 1993), originated in the central Amazon, but were then displaced to the north-west (Caribbean), and south and east (south Amazon and Chaco-Parana) by parallel clades of orange rayed morphs, which are now found across the Amazonian region. This suggestion had in fact been made in the earlier *Heliconius* literature, as the most parsimonious explanation for the disjunct biogeographic distribution of the peripheral postman morphs (see Sheppard et al. 1985 p 591; Mallet 1993 p. 250; Hines et al. 2011).

Two sources of Neotropical vicariance have previously been suggested as possible influences on mimicry between *H. erato* and *H. melpomene*: fragmentation of the forest into glacial refugia (Brown et al.1974; Sheppard et al. 1985) and mountain uplift in the Andes (Brower 1996). However, several lines of evidence suggest that the wing pattern diversification of *H. erato* and *H. melpomene* was driven primarily by their Müllerian mimicry, rather than a common history of allopatry. We estimate that the diversifications of mimetic *H. erato* and *H. melpomene* began no earlier than

720,0000 years ago, based on the arthropod mtDNA substitution rate estimated by Brower (1994), or 400,000 years ago based on a faster molecular-clock rate estimated for the swallowtail butterflies (Simonsen et al. 2011). These figures are broadly in line with those estimated by other studies using modern coalescent methods (e.g. see Kozak et al. 2015 Table 1 and Fig. S3). The Bayesian multilocus coalescent method (Heled and Drummond 2010), used to date these divergences, estimates and accounts for effective population size, which can bias divergence time estimates based on simple comparisons of genetic divergence (e.g. Wakely and Hey 1997; Arbogast et al. 2002; Hoyal Cuthill and Charleston 2012; though see also Flanagan et al. 2004). These estimated ages are well after the formation of the Andes (with active mountain uplift between approximately 23 and 2 mya (Gregory-Wodzicki 2000; Hoorn et al. 2010).

The diffusions on either side of the Andean range, visible in the phylogeographic reconstructions (Fig. 5), also suggest the presence of the Andes as a significant barrier to east-west migration (see also Miller et al. 2008). In particular, diversification to the west of the Andes (blue iridescent morphs of *H. erato* and *H. melpomene* in Colombia and Ecuador, Fig. 1), has the appearance of being channelled between the mountains and the Pacific ocean (Fig. 5). This suggests that the Andean range was already a major geographical feature during the mimicry radiation of *H. erato* and *H. melpomene* and its uplift did not split the eastern and western clades (as had been previously suggested for *H. erato* in the northern Andean region (Brower 1996)). Instead, our phylogeographic reconstructions suggest that the western Andean coast was colonised during the latter stages of parallel biogeographic radiations (well after Andean uplift is thought to have ceased (Gregory-Wodzicki 2000; Hoorn et al. 2010). This implies that the western Andean populations diversified southwards from

the North of the Andean range, where low mountain passes may have aided trans-Andean migration (e.g. Miller et al. 2008). The continuous phylogeographic diffusion model used in our reconstructions (Drummond and Rambaut 2007; Lemey et al. 2009; Lemey et al. 2010) does not directly incorporate any migration barriers. Therefore, these reconstructions suggest that this major landscape feature left a "footprint" (Lemey et al. 2010) in the biogeographic distributions of these butterflies, recoverable via phylogeographic analysis of their gene sequences.

The extent to which Neotropical forests were affected by Quaternary glaciations has been controversial. It had been suggested that climate drying led to isolation of forest refugia by areas of savannah (Haffer 1969), which might have affected *Heliconius* mimicry dynamics (Brown, Sheppard and Turner 1974; Sheppard et al. 1985). However, pollen and climate studies are now thought to indicate relatively continuous Neotropical forest cover during the time interval estimated for the mimicry radiation of *H. erato* and *H. melpomene* (Knapp and Mallet 2003), though there may have been more forest fragmentation in peripheral areas (Fouquet et al. 2012). In support of this latter suggestion, both of the Chaco-Parana co-mimics *H. e. phyllis* and *H. m. nanna* (Fig. 1) show comparatively high levels of sequence divergence (Figs. 2, S10, S11), indicating relative genetic isolation. Overall, our phylogeographic reconstructions suggest that both species were able to diffuse widely, throughout the Neotropics, without signs of any major barrier to migration (other than the Andes, as discussed above). For example, the reconstructed radiations east of the Andes cross two previously proposed glacial refugia (Napo and Inambari e.g. see Noonan and Wray 2006). However, we note that the model used here to reconstruct phylogeographic histories (Fig. 5) is one of continuous geographic diffusion (Bielejec et al. 2011) and,

therfore, may not be suited to direct tests of refugial isolation and release. Furthermore, although an additional influence from glacial forest fragmentation remains a possibility, Müllerian mimicry itself must have been a primary driver of a coevolutionary phenotypic radiation by *H. erato* and *H. melpomene* (Turner, 1981; Sheppard et al. 1985). Given the spectacular correspondence between their phenotypes and geographic ranges (Fig. 1), it is likely that *H. erato* and *H. melpomene* have been major partners in mimicry throughout their evolutionary history. However, wider interactions with other butterfly species, such as *Heliconius timareta* and *Heliconius elevatus*, may also have had an influence this mimicry system (e.g. Turner, 1983; Arias et al. 2014).

Acknowledgements

Research by JHC and MC was conducted at the University of Sydney, funded by Australian Research Council grant number DP1094891 and by JHC at the University of Cambridge, during a Postdoctoral Research Associateship supervised by Simon Conway Morris and funded by Templeton Foundation World Charity grant number LBAG/143. We would like to thank the University Museum of Zoology, Cambridge for access to *Heliconius* specimens. We are very grateful to John Turner as well as two further anonymous reviewers for comments on the manuscript.

References

Arbogast, B. S., S. V. Edwards, J. Wakely, P. Beerli, and J. B. Slowinski. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic

timescales. Ann. Rev. Ecol. Syst. 33:707-740.

Arias, C. F., C. Salazar, C. Rosales, M. R. Kronforst, M. Linares, E. Bermingham, W. O. McMillan. 2014. Phylogeogrpahy of *Heliconius cydno* and its closest relatives: disentangling their origin and diversification. Mol. Ecol. 23: 4137-4152.

Baxter, S. W. et al. 2010. Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in the *Heliconius melpomene* clade. PLoS Genet. 6(2):e1000794.

Bielejec, F., A. Rambaut, M. A. Suchard, and P. Lemey. 2011. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics 27:2910-2912.

Brower, A. V. Z. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proc. Natl. Acad. Sci. USA 91:6491-6495.

Brower, A. V. Z. 1996. Parallel race formation and the evolution of mimicry in *Heliconius* butterflies: a phylogenetic hypothesis from mitochondrial DNA sequences. Evolution 50:195-221.

Brown, K. S., P. M. Sheppard, and J. R.G. Turner. 1974. Quaternary refugia in tropical America: evidence from race formation in *Heliconius* butterflies. Proc. Roy. Soc. Lond. B 187:369-378.

Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17:540-552.

Charleston, M. A., and D. L. Robertson. 2002. Preferential host switching by primate lentiviruses can account for phylogenetic similarity with the primate phylogeny. Syst. Biol. 51:528-535.

Conow, C., D. Fielder, Y. Ovadia and R. Libeskind-Hadas. 2010. Jane: a new tool for the cophylogeny reconstruction problem. Algorithm. Mol. Biol. 5:16.

Counterman et al. 2010. Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in *Heliconius erato*. PLoS Genet. 6(2):e1000796.

Crisp, M. D., and L. G. Cook. 2005. Do early branching lineages signify ancestral traits? Trends. Ecol. Evol. 20:122-128.

Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7:214.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792-1797.

Farris, J. S. 1989. The retention index and rescaled consistency index. Cladistics 5:417-419.

Flanagan, N. S., et al. 2004. Historical demography of Müllerian mimicry in the neotropical *Heliconius* butterflies. Proc. Natl. Acad. Sci. USA 101:9704-9709.

Fouquet, A., B. P. Noonan, M. T. Rodrigues, N. Pech, A. Gilles, and N. J. Gemmell. 2012. Multiple Quaternary refugia in the eastern Guiana Shield revealed by comparative phylogeography of 12 frog species. Syst. Biol. 61:461-489.

Futuyma, D. J., and M. Slatkin. 1983. Introduction. Pp 1-3 *in* D. M. Futuyma and M. Slatkin, eds. Coevolution. Sinauer Associates Inc., Sunderland.

Gilbert, L. E. 1983. Coevolution and mimicry. Pp 263-281 *in* D. M. Futuyma and M. Slatkin, eds. Coevolution. Sinauer Associates Inc., Sunderland.

Gilbert, L. E. 1984. The biology of butterfly communities. Pp 41-54 *in* R. Vane-Wright and P. Ackery, eds. The biology of butterflies. Academic Press, New York**.**

Gregory-Wodzicki, K. M. 2000. Uplift history of the Central and Northern Andes: a

review. Geol. Soc. Am. Bull. 112:1091-1105.

Haffer, J. 1969. Speciation in Amazonian forest birds. Science 165:131-137.

Hafner M. S., J. W. Demastes, T. A. Spradling, and D. L. Reed. 2003. Cophylogeny between pocket gophers and chewing lice. Pp 195-220 *in* R. D. M. Page, ed. Tangled trees: Phylogeny, cospeciation, and coevolution. The University of Chicago Press, Chicago.

Heikkilä, M., L. Kaila, M. Mutanen, C. Peña, and N. Wahlberg. 2011. Cretaceous origin and repeated tertiary diversification of the redefined butterflies. Proc. Roy. Soc. Lond. B. 279:1093-1099.

Heled, J., and A. J. Drummond. 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27:570-580.

Hey, J, and J. Wakely. 1997. A coalescent estimator of the population recombination rate. Genetics 145: 833-846.

Hill, R. I., L. E. Gilbert, and M. R. Kronforst. 2013. Cryptic genetic and wing pattern diversity in a mimetic *Heliconius* butterfly. Mol. Ecol. 22:2760-2770.

Hines, H. M., et al. 2011. Wing patterning gene redefines the mimetic history of *Heliconius* butterflies. Proc. Natl. Acad. Sci. USA 108(49):19666-19671.

Hoorn, C., et al. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science 330:927-931.

Hoyal Cuthill, J., and M. Charleston. 2012. Phylogentic codivergence supports coevolution of mimetic *Heliconius* butterflies. PLoS ONE 7:e36464.

Huson, D. H., and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23:254-267.

Janzen, D. H. 1980. When is it coevolution? Evolution 34:611-612.

Jiggins, C. D., and W. O. McMillan. 1997. The genetic basis of an adaptive radiation: warning colour in two *Heliconius* species. Proc. Roy. Soc. B 264:1167-1175.

Joron, M., and J. L. Mallet. 1998. Diversity in mimicry: paradox or paradigm? Trends Ecol. Evol. 13:461-466.

Kaas, R. E., and A. E. Raftery. 1995. Bayes factors. J. Am. Stat. Assoc. 90:773-795.

Kapan, D. D. 2001. Three-butterfly system provides a field test of Müllerian mimicry. Nature 409:338-340.

Knapp, S., and J. Mallet. 2003. Refuting refugia? Science 300:71-72.

Kozak, K. M., N. Wahlberg, A. F. E. Neild, K. K. Dasmahapatra, J. Mallet, and C. D. Jiggins. 2015. Multilocus species trees show the recent adaptive radiation of the mimetic *Heliconius* butterflies. Syst. Biol. 64:505.524.

Lemey, P., A. Rambaut, A. J. Drummond, and M. A. Suchard. 2009. Bayesian phylogeography finds its roots. PLoS Comput. Biol. 5:e1000520.

Lemey, P., A. Rambaut, J. J. Welch, and M. A. Suchard. 2010. Phylogeography takes a relaxed random walk in continuous space and time. Mol. Biol. Evol. 27:1877-1885.

Mallet, J. 1993. Speciation, raciation, and color pattern evolution in *Heliconius* butterflies: evidence from hybrid zones. Pp 226-260 *in* R. G. Harrison, ed. Hybrid zones and the eovlutionary process. Oxford University Press, New York.

Mallet, J. 1999. Causes and consequences of a lack of coevolution in Müllerian mimicry. Evol. Ecol. 13:777-806.

Mallet, J., C. D. Jiggins, and W. O. McMillan. 1996. Mimicry meets the mitochodrion. Curr. Biol. 6:937-940.

Martin, A., K. J. McCulloch, N. Patel, A. D. Briscoe, L. E. Gilbert, and R. D. Reed. 2014. Multiple recent co-options of Optix associated with novel traits in adaptive

butterfly wing radiations. EvoDevo 5:7.

Merrill, R. M. et al. 2015. The diversification of *Heliconius* butterflies: What have we learned in 150 years? J. Evol. Biol. doi: 10.1111/jeb.12672.

Miller, M. J., et al. 2008. Out of Amazonia again and again: episodic crossing of the Andes promotes diversification in a lowland forest flycatcher. Proc. Roy. Soc. B 275:1133-1142.

Morrone, J. J. 2006. Biogeographic areas and transition zones of Latin America and the Caribbean Islands based on panbiogeographic and cladistic analyses of the entomofauna. Annu. Rev. Entomol. 51: 467-494.

Müller, F. 1879. *Ituna* and *Thyridia*; a remarkable case of mimicry in butterflies. Trans. Entomol. Soc. Lond. 1879:xx-xxix.

Nadeau, N. J. et al. 2014. Population genomics of parallel hybrid zones in the mimetic butterflies, H. melpomene and H. erato. Genome Res. 24.8:1316-1333.

Noonan, B. P., and K. P. Wray. 2006. Neotropical diversification: the effects of a complex history on diversity within the poison frog genus *Dendrobates*. J. Biogeogr. 33:1007-1020.

Nuismer, S. L., P. Jordano, and J. Bascompte. 2012. Coevolution and the architecture of mutualistic networks. Evolution 67:338-354.

Page, R. D. M. 2003. Introduction. Pp 1-21 *in* R. D. M. Page, ed. Tangled trees: Phylogeny, cospeciation, and coevolution. The University of Chicago Press, Chicago.

Papa, R., et al. 2013. Multi-allelic major effect genes interact with minor effect QTLs to control adaptive color pattern variation in *Heliconius erato*. PLoS ONE 8:e57033.

Papadopoulou, A., I. Anastasiou, and A. P. Volger. 2010. Revisiting the insect

mitochondrial molecular clock: the mid-Aegean trench calibration. Mol. Biol. Evol. 27:1659-1672.

Posada, D. 2008. jModeltest: phylogenetic model averaging. Mol. Biol. Evol. 25:1253-1256.

Quek, S.-P., et al. 2010. Dissecting comimetic radiations in *Heliconius* reveals divergent histories of convergent butterflies. Proc. Natl. Acad. Sci. USA 107:7365–7370.

Reed, R. D., et al. 2011. *optix* drives the repeated convergent evolution of butterfly wing pattern mimicry. Science 333:1137-1141.

Ronquist, F et al. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61:539-542.

Rosenberg, N. A., and M. Nordborg. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nat. Rev. Genet. 3:380-390.

Sheppard, P. M., J. R. G. Turner, K. S. Brown, W. W. Benson, and M. C. Singer. 1985. Genetics and the evolution of Muellerian mimicry in *Heliconius* butterflies. Phil Trans. Roy. Soc. Lond. (Biol.) 308:433-613.

Sheratt, T. N. 2006. Spatial mosaic formation through frequency-dependent selection in Müllerian mimicry complexes. J. Theor. Biol. 240:165-174.

Sherratt, T. N. 2008. The evolution of Müllerian mimicry. Naturwissenschaften 95:681-695.

Simonsen, T. J., E. V. Zakharov, M. Djernaes, A. M. Cotton, R. I. Vane-Wright, and F. A. H. Sperling. 2011. Phylogenetics and divergence times of Papilioninae (Lepidoptera) with special reference to the enigmatic genera *Teinopalpus* and *Meandrusa*. Cladistics 27:113-137.

Supple, M. A., Hines, H. M., Dasmahapatra, K. K., Lewis, J. J., Nielsen, D. M.,

Lavoie, C., Ray, D. A., Salazar, C., McMillan, W. O., and Counterman, B. A. 2013. Genomic architecture of adaptive color pattern divergence and convergence in *Heliconius* butterflies. Genome Res. 150615.112.

Swofford, D. L. 2000. PAUP* 4. Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland.

Talavera, G., and J. Castresana. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequences. Syst. Biol. 56:564-577.

Thompson, J. N. 1989. Concepts of coevolution. Trends Ecol. Evol. 4:179-183.

Thompson, J. N. 1994. The coevolutionary process. University of Chicago Press, Chicago.

Thompson, J. N., and B. M. Cunningham. 2002. Geographic structure and dynamics of coevolutionary slection. Nature 417:735-738.

Turner, J. R.G. 1981. Adaptation and evolution in *Heliconius*. Ann. Rev. Ecol. Syst. 12:99-121.

Turner, J. R.G. 1983. Mimetic butterflies and punctuated equilibria: some old light on a new paradigm. Biol. J. Linn. Soc. 20:277-300.

Turner, J. R.G. 1987. The evolutionary dynamics of Batesian and Muellerian mimicry: similarities and differences. Ecol. Entomol. 12: 81-95.

Wakely, J., and J. Hey. 1997. Estimating ancestral population parameters. Genetics 145:847-855.

Wang, R. L., J. Wakely, and J. Hey. 1997. Gene flow and natural selection in the origin of *Drosophila pseudoobscura*. Genetics 147:1091-1106.

Zhang, C., D.-X. Zhang, T. Zhu, and Z. Yang. 2011. Evaluation of a Bayesian coalescent method of species delimitation. Syst. Biol. 60:747-761.

Figure Legends

Figure 1. Neotropical distributions of the Müllerian co-mimics *Heliconius erato* (left) and *Heliconius melpomene* (right) included in this study (redrawn after Brower 1996). Shaded areas indicate biogeographic subregions (Morrone 2006). Black lines indicate range boundaries (after Brower 1996). Coloured circles indicate the broad pattern type: postman with red forewing band (red circles), additional yellow hindwing band (yellow circles), or rayed (orange circles) (after Sheppard et al. 1985; Hines et al. 2011).

Figure 2. Splits network for the colour pattern and neutral gene partitions (including *optix*, *bves*, *kinesin*, *GPCR*, *VanGogh*, mt COI and COII, *SUMO*, *Suz12*, *2654* and *CAT*). Parallel edges split specimens into two groups (corresponding to a branch among neighbour joining phylogenies for the ten genes). Highlighting indicates supporting gene partition(s) (red, colour and neutral genes; yellow, colour genes only; black (no highlight), neutral genes only). Coloured circles indicate a specimens pattern type (as for Fig. 1) and biogeographic subregion (Caribbean, dark grey, solid border; Amazonian, light grey, dashed border; Chaco-Parana, light grey, solid border). The scale bar indicates DNA substitutions per site.

Figure 3. Bayesian multi-locus coalescent phylogenies reconstructed with *BEAST for co-mimic populations of *H. erato* (blue) and *H. melpomene* (yellow) based on **(A)** a population-level analysis of the colour pattern genes (*optix*, *bves*, *kinesin*, *GPCR* and *VanGogh*) **(B)** a morph-level analysis of the colour pattern genes **(C)** a combined

morph-level analysis of all genes (colour pattern and neutral: *optix*, *bves*, *kinesin*, *GPCR*, *VanGogh*, mt COI and COII, *SUMO*, *Suz12*, *2654* and *CAT*), corresponding to the complete phylogeny of all studied species and morphs shown in Fig. S9. Grey lines connect co-mimic populations. Colour coding as for Fig. 2. Additionally, red circles indicate phylogenetic subtrees showing statistically significant congruence in the co-phylogeny correlation test (figure key shows *p* values).

Figure 4. Visualisation of the Bayesian tree set for the population-level analysis of all ten studied genes reconstructed with *BEAST (*optix*, *bves*, *kinesin*, *GPCR*, *VanGogh*, mt COI and COII, *SUMO*, *Suz12*, *2654* and *CAT*). The axis shows divergence time (mya) on a log scale. Densely coloured regions (shown in blue) indicate phylogenetic topologies and branch lengths represented by larger numbers of trees (Bouckaert 2010). Consensus topologies and branch lengths are shown in purple.

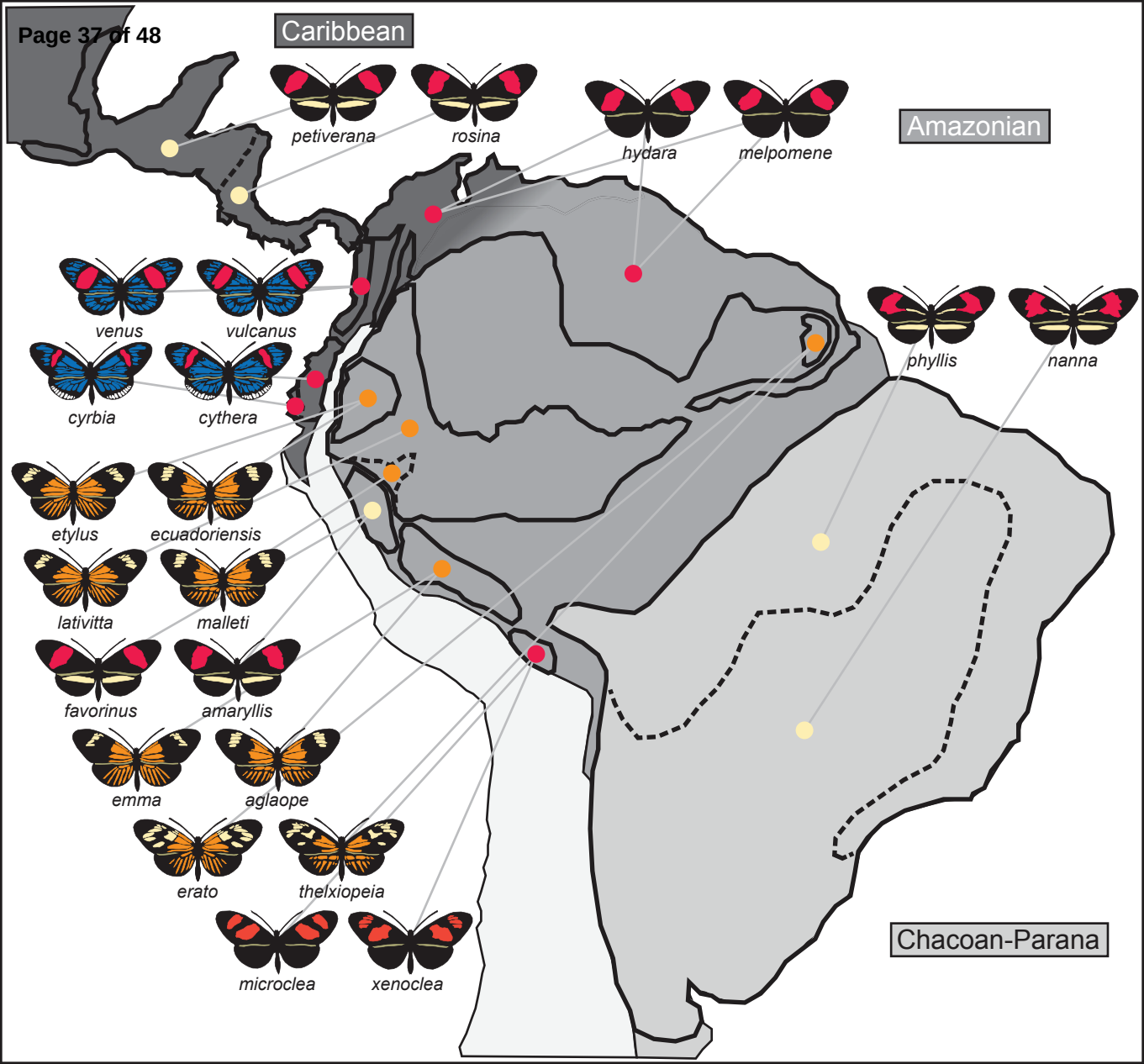
Figure 5. Bayesian Phylogeographic reconstructions for *H. erato* (blue) and *H. melpomene* (yellow) reconstructed with BEAST and SPREAD using all ten studied genes (*optix*, *bves*, *kinesin*, *GPCR*, *VanGogh*, mt COI and COII, *SUMO*, *Suz12*, *2654* and *CAT*). Branch colour gradient indicates relative tree height (white, oldest branches). Circles indicate reconstructed points of origin for each species. The blue star indicates the reconstructed ancestral reconstruction for a supplementary analysis (Fig. S8) incorporating additional mitochondrial sequences (mt COI and COII) for *H. erato cruentus* from Mexico and *H. erato demophoon* from Costa Rica (sourced from Hill et al. 2013). Dashed lines indicate major geographic features (green, Amazonia; red, the Andes). Satellite image courtesy of Google Earth, US Dept. of State Geography,

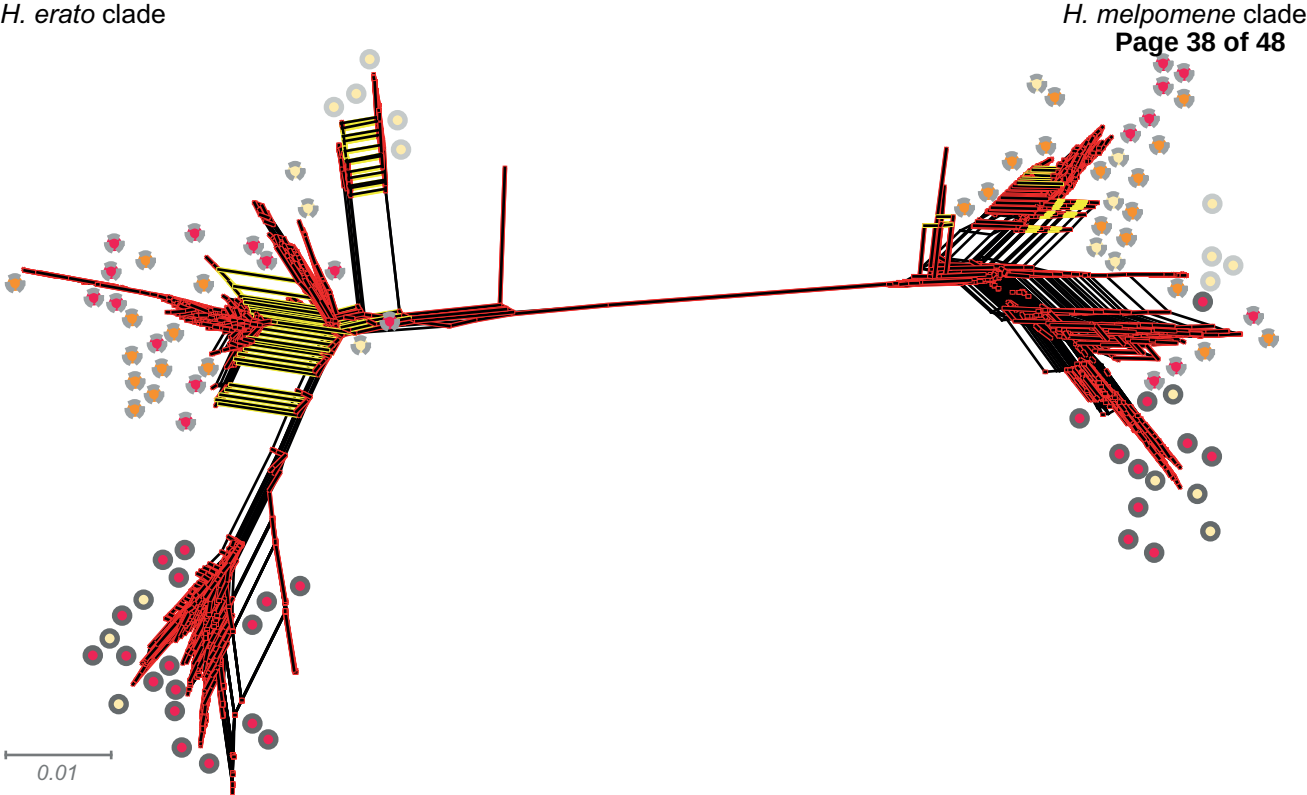
Inav/Geosistemas SRL, MapLink, Data SIO, NOAA, U.S. Navy, NGA, and GEBCO.

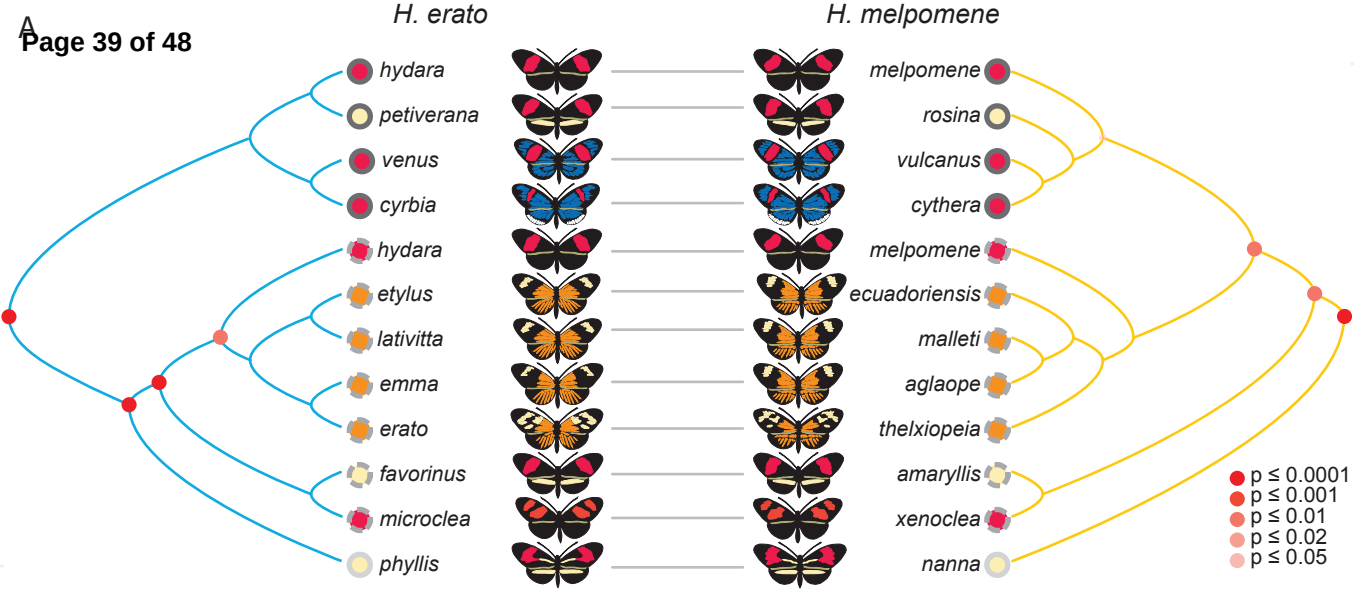
Table Legends

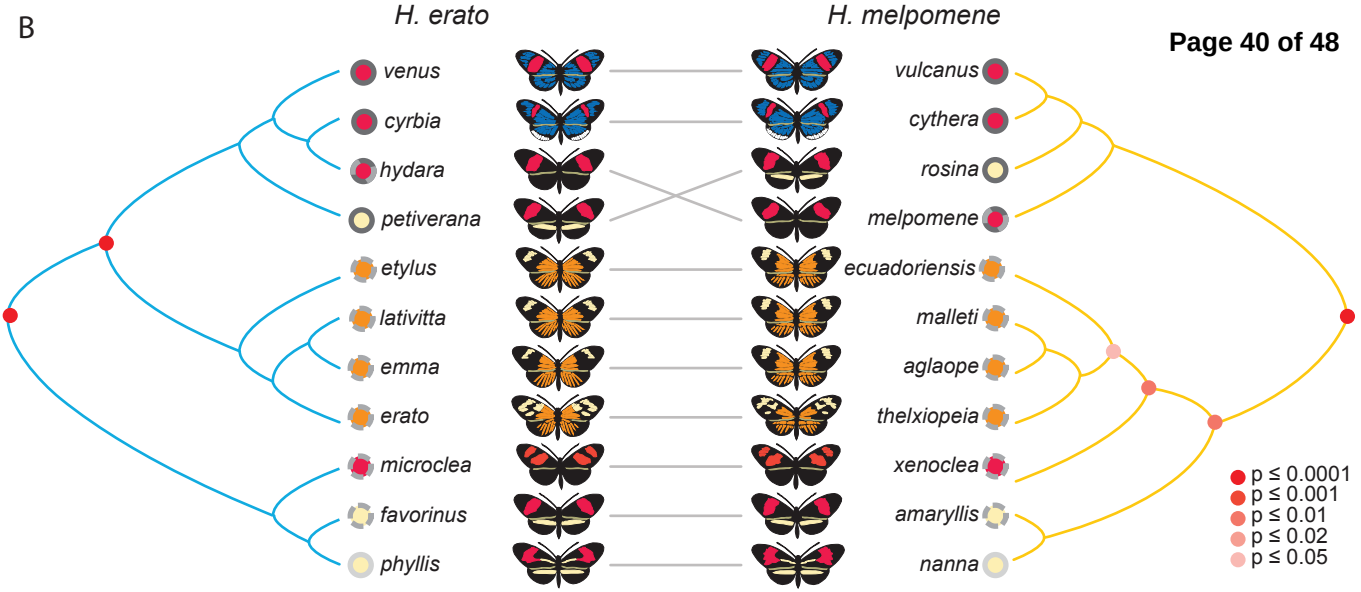
Table 1. Retention indices (RI = (G-S)/(G-M), where S is the most parsimonious number of evolutionary steps, and M and G are, respectively, the minimum and maximum number of steps (Farris 1989)) for the broad wing pattern types (after Sheppard et al. 1985; Hines et al. 2011) and biogeographic subregions (Morrone 2006) as reconstructed on phylogenetic estimates based on the colour, neutral and combined gene partitions.

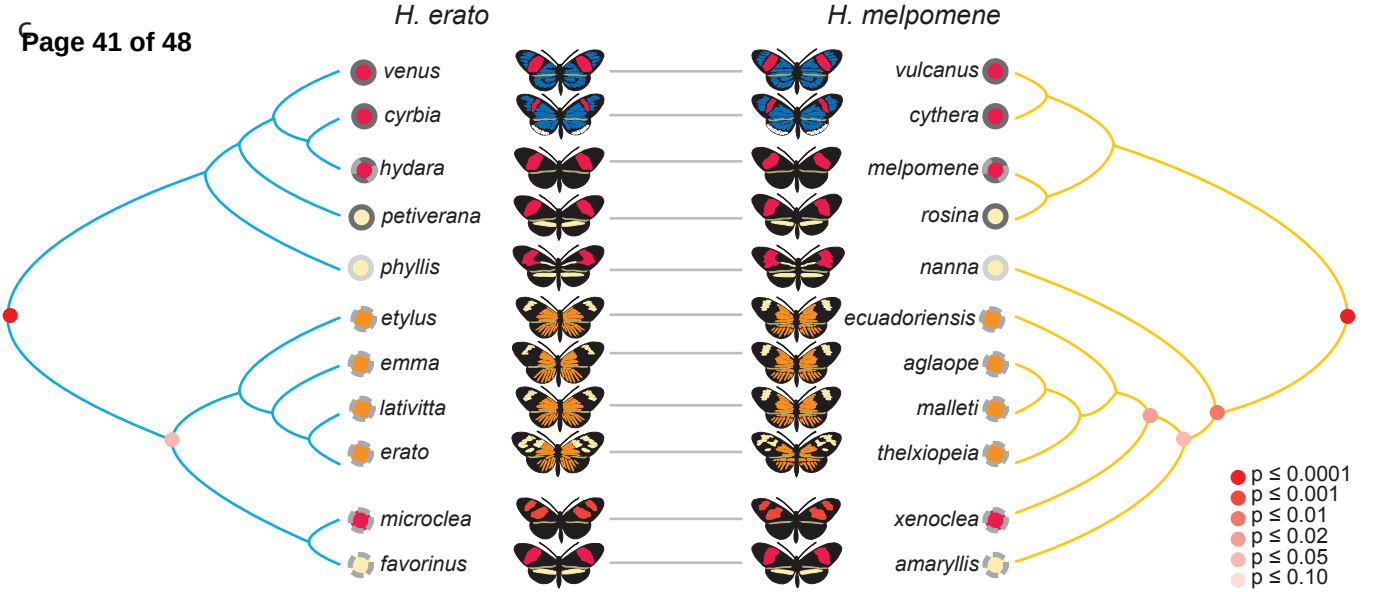
Table 2. Co-phylogeny correlation and co-phylogeny mapping results.

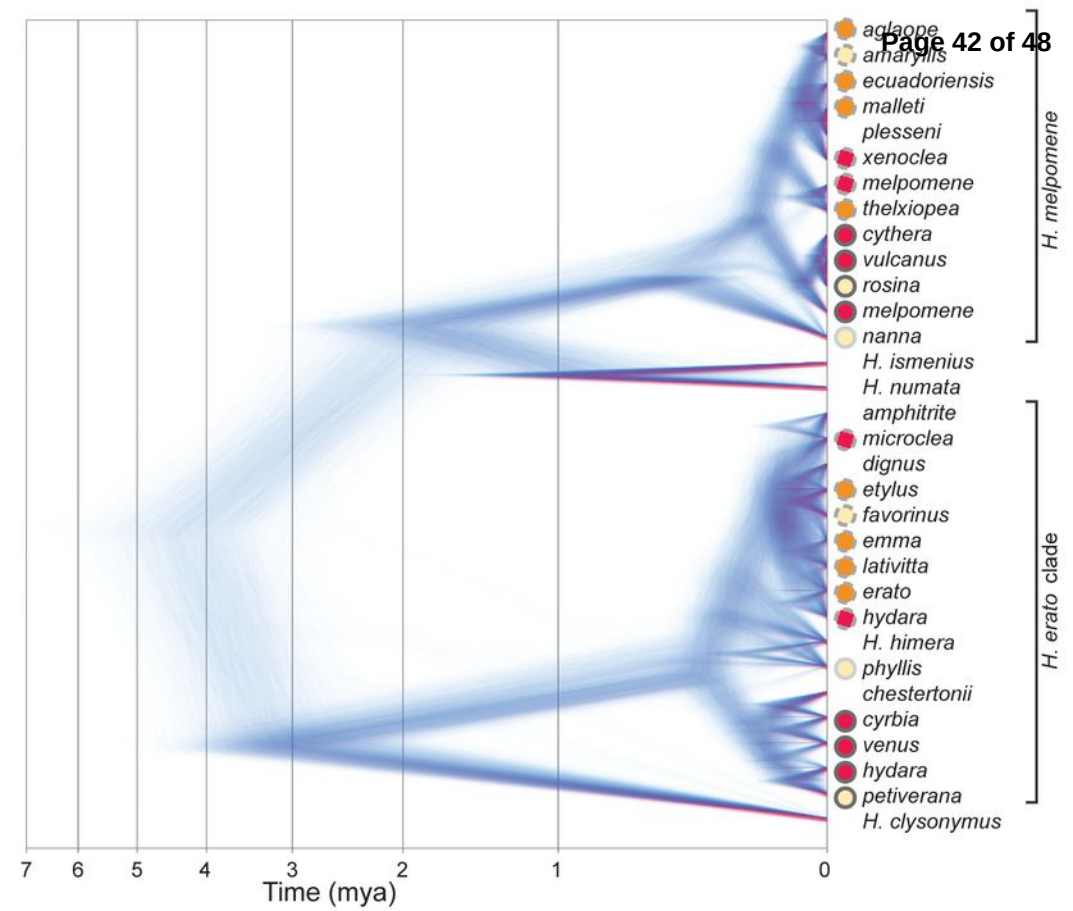

Caribbean
petiverana
rosina
hydara
melpomene
Amazonian
venus
vulcanus
cyrbia
cythera
phyllis
nanna
etylus
ecuadoriensis
lativitta
malleti
favorinus
amaryllis
emma
aglaope
erato
thelxiopeia
microclea
xenoclea
Chacoan-Parana


0.01


H. erato
H. melpomene
hydara
petiverana
venus
cyrbia
hydara
etylus
lativitta
emma
erato
favorinus
microclea
phyllis
melpomene
rosina
vulcanus
cythera
melpomene
ecuadoriensis
malleti
aglaope
thelxiopeia
amaryllis
xenoclea
nanna
p ≤ 0.0001
p ≤ 0.001
p ≤ 0.01
p ≤ 0.02
p ≤ 0.05


B
H. erato
H. melpomene
venus
cyrbia
hydara
petiverana
etylus
lativitta
emma
erato
microclea
favorinus
phyllis
vulcanus
cythera
rosina
melpomene
ecuadoriensis
malleti
aglaope
thelxiopeia
xenoclea
amaryllis
nanna
p ≤ 0.0001
p ≤ 0.001
p ≤ 0.01
p ≤ 0.02
p ≤ 0.05


H. erato
H. melpomene
venus
cyrbia
hydara
petiverana
phyllis
etylus
emma
lativitta
erato
microclea
favorinus
vulcanus
cythera
melpomene
rosina
nanna
ecuadoriensis
aglaope
malleti
thelxiopeia
xenoclea
amaryllis
p ≤ 0.0001
p ≤ 0.001
p ≤ 0.01
p ≤ 0.02
p ≤ 0.05
p ≤ 0.10


agdaope
amaryllis
ecuadoriensis
malleti
plesseni
xenoclea
melpomene
thelxiopea
cythera
vulcanus
rosina
melpomene
nanna
H. ismenius
H. numata
amphitrite
microclea
dignus
etylus
favorinus
emma
lativitta
erato
hydara
H. himera
phyllis
chestertonii
cyrbia
venus
hydara
petiverana
H. clysonymus
H. melpomene
H. erato clade
7
6
5
4
3
2
1
0
Time (mya)

Population level	Species	Genes	Character	RI	Mean	Species	Genes
regions	*H. erato*	colour	red forewing band	1		*H. melpomene*	colour
			yellow hindwing band	0			
			rayed	1			
			blue iridescence	1	0.75		
			biogeographic region	1			
morphs			red forewing band	1			
			yellow hindwing band	0.5			
			rayed	1			
			blue iridescence	0	0.63		
			biogeographic region	1			
regions		neutral	red forewing band	0			neutral
			yellow hindwing band	0			
			rayed	0			
			blue iridescence	0	0.00		
			biogeographic region	1			
		all	red forewing band	0.33			all
			yellow hindwing band	0			
			rayed	0.33			
			blue iridescence	1	0.42		
			biogeographic region	1			

Character	RI	Mean	Phylogenetic estimate
red forewing band	1		Fig. 2A
yellow hindwing band	0		
rayed	1		
blue iridescence	1	0.75	
biogeographic region	1		
red forewing band	1		Fig. 2B
yellow hindwing band	0.5		
rayed	1		
blue iridescence	1	0.88	
biogeographic region	1		
red forewing band	0.33		
yellow hindwing band	0		
rayed	0.33		
blue iridescence	0	0.17	
biogeographic region	1		Fig. S1
red forewing band	0.33		
yellow hindwing band	0		
rayed	0.33		
blue iridescence	0	0.17	
biogeographic region	1		Fig. S2

Genes	**Population level**	**Phylogenetic estimate**	**Mimicry model**	***p* cophylogeny correlation (root model)**
colour	regions	Fig. 3A	*H. erato*	**< 0.0001**
			H. melpomene	**< 0.0001**
	morphs	Fig. 3B	*H. erato*	**< 0.0001**
			H. melpomene	**< 0.0001**
neutral	regions	Fig. S1	*H. erato*	**< 0.0001**
			H. melpomene	**< 0.0001**
	morphs		*H. erato*	**< 0.0001**
			H. melpomene	**< 0.0001**
all	regions	Fig. S2	*H. erato*	**< 0.0001**
			H. melpomene	**< 0.0001**
	morphs		*H. erato*	**< 0.0001**
			H. melpomene	**< 0.0001**

p cophylogeny correlation (root mimic)	*p* cophylogeny mapping	Cospeciations	Duplication & host switch	Losses	Cost
< 0.0001	**<0.001**	8	3	1	7
< 0.0001	**0.003**	7	4	2	10
< 0.0001	**0.002**	6	4	0	8
< 0.0001	**<0.001**	7	3	0	6
< 0.0001	0.109	6	5	4	14
< 0.0001	0.063	6	5	3	13
0.005	0.593	3	7	1	15
0.003	0.098	4	6	0	12
< 0.0001	0.052	6	5	3	13
< 0.0001	**0.017**	6	5	2	12
< 0.0001	**0.037**	6	4	3	11
< 0.0001	**0.035**	6	4	3	11

Cophylogeny map

Fig. S7A
Fig. S7B
Fig. S7C
Fig. S7D
Fig. S7E
Fig. S7F
Fig. S7G
Fig. S7H
Fig. S7I
Fig. S7J
Fig. S7K
Fig. S7L