Refining mimicry: phenotypic variation tracking the local optimum

Claire Mérot, Yann Le Poul, Marc Thery, Mathieu Joron

Mimicry refinement: Phenotypic variations tracking the local optimum

Mérot C.[1*], Le Poul Y.[1], Théry M.[2], Joron M.[1,3*]

[1] Institut de Systématique Evolution et Biodiversité, UMR 7205 CNRS – MNHN – UPMC – EPHE, Muséum National d'Histoire Naturelle, 45 rue Buffon, 75005 Paris, France.

[2] Mécanismes Adaptatifs et Evolution, UMR 7179 CNRS, Museum National d'Histoire Naturelle, 1 avenue du petit château, Brunoy, France

[3] Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175 CNRS – Université de Montpellier – Université Paul Valéry Montpellier – EPHE, 1919 route de Mende, 34293 Montpellier 5, France

* corresponding author : mathieu.joron@cefe.cnrs.fr; merot@mnhn.fr

Running title: Mimicry refinement in butterflies

Key words: Colour pattern – adaptation - Fitness peak - Gene flow – Hybridization - Lepidoptera - Morphometrics – Müllerian mimicry – Perfect mimicry – Geographic variation

Summary

1. Müllerian mimicry between chemically defended preys is a textbook example of natural selection favouring phenotypic convergence onto a shared warning signal. Studies of mimicry have concentrated on deciphering the ecological and genetic underpinnings of dramatic switches in mimicry association, producing a well-known

mosaic distribution of mimicry patterns across geography. However, little is known about the accuracy of resemblance between natural co-mimics when the local phenotypic optimum varies.

2. In this study, using analyses of wing shape, pattern and hue, we quantify multimodal phenotypic similarity between butterfly co-mimics sharing the so-called postman pattern in different localities with varying species composition.
3. We show that subtle but consistent variation between populations of the localised species, *Heliconius timareta thelxinoe,* enhance resemblance to the abundant co-mimics which drive the mimicry in each locality.
4. Those results suggest that rarer co-mimics track the changes in the phenotypic optimum caused by gradual changes in the composition of the mimicry community, providing insights into the process by which intra-specific diversity of mimetic pattern may arise. Furthermore, our results suggest a multimodal evolution of similarity, with coordinated convergence in different features of the phenotype such as wing outline, pattern and hue.
5. Finally, multilocus genotyping allows estimating local hybridization rates between *H. timareta* and co-mimic *H. melpomene* in different populations, raising the hypothesis that mimicry refinement between closely-related co-mimics may be enhanced by adaptive introgression at loci modifying the accuracy of resemblance.

Introduction

Chemically-defended animal species often show striking convergence in their colour patterns with other prey coexisting in the same habitat. This convergence may be explained, since the work of Müller (1879), by natural selection favouring superficial resemblance, and operated by visual predators, a phenomenon called Müllerian mimicry. Theory proposes that predators learn upon experience the association of prey distastefulness and prey visual appearance, generally distinctive warning patterns. Mimicking a locally abundant warning signal, well known by local predators, constitutes a benefit to a defended prey species by decreasing predation risk. Mimicry benefits associated with a given warning signal depend on the relative numbers of prey sampled vs. available to learning predators, and are usually driven by the most toxic or the most abundant prey species (Mallet & Joron 1999).

Müller's general principle has been largely supported by theoretical models and empirical experiments (Turner 1977; Sheppard *et al.* 1985; Turner 1987; Ruxton, Sherratt & Speed 2004). Field transplant experiments have confirmed selection favouring local patterns (Mallet & Barton 1989; Kapan 2001; Chouteau & Angers 2011; Merrill *et al.* 2012) and the strong frequency-dependent selection acting on warning signals in diverse communities (Chouteau, Arias & Joron 2016). Natural selection for local mimicry explains local polymorphism of distinct colour patterns (Kapan 2001; Joron & Iwasa 2005), as well as the maintenance of geographical races with sharply distinct patterns, for instance the so-called postman vs. rayed patterns of *Heliconius erato* and *H. melpomene* (Mallet & Barton 1989). Most of those studies investigated the mimicry benefits associated with alternative warning strategies using completely distinct patterns, corresponding to mimetic optima and describing an adaptive landscape with adaptive peaks separated by valleys of low fitness (Leimar & Mallet 2012). However, for a given morph, variations may be found between individuals, between sexes and

between localities. Fewer studies have addressed the significance of mimicry variation around a given adaptive peak, or the underpinnings of precise resemblance within a given mimicry ring. Mimetic communities often involve assemblages of species which differ between localities, and species indeed vary in the level of mimicry precision to others (Penney *et al.* 2012), but it is unclear what determines the level of mimicry accuracy in those communities or to which extent the mimicry optimum may vary through space or time.

The ability of predators to generalize the signals of defended prey is an important determinant of selection on resemblance in a mimicry system (Rowe, Lindstrom & Lyytinen 2004; Ihalainen *et al.* 2012). Sharing key components of a warning signal with co-mimics can sometimes enhance protection and may allow crossing a valley of low fitness (Beatty, Beirinckx & Sherratt 2004; Balogh & Leimar 2010). For instance, jacamars trained to avoid butterflies with an orange patch were less likely to attack butterflies with a similar pattern painted red than those painted black (Langham 2004), suggesting that a red patch may be sufficient to reduce predation regardless of the differences in hue. Nevertheless, in that example red butterflies still received higher predation than orange controls (Langham 2004), suggesting that small deviations of colour hue are, to a certain extent, perceived by predators and translate into fitness differences. Increased predation against imperfect mimics has also been shown in lab experiments (Ihalainen *et al.* 2008), meaning that even if coarse resemblance is attained, selection may still favour the improvement of mimicry. Those findings support the classical scenario for the evolution of mimicry, first through a major phenotypic change allowing coarse resemblance for certain key warning feature, followed by the gradual improvement of mimicry under selection by narrowly generalizing predators (Sheppard *et al.* 1985; Turner 1987; Franks & Sherratt 2007; Balogh & Leimar 2010; Leimar & Mallet 2012).

The strength of selection for resemblance is affected by the complexity and diversity of the prey community. For instance, discrimination against imperfect mimics is less accurate when the community of prey is complex, i.e. composed of several distinct warning signals (Ihalainen *et al.* 2012). However, whether variations of pattern within a mimicry ring itself affect the intensity of selection for resemblance has rarely been tested, since experimental predators usually are only trained on a single prey type. Quantifying variations in resemblance of wild individuals within and between species of a given mimicry ring allows investigating if and how the accuracy of resemblance evolves.

Phenotypic similarity is influenced not only by natural selection favouring accurate mimicry but also by the genetic architecture underlying variation in phenotype. For instance, phenotypic variation remaining within the generalization range of predators might be little influenced by mimicry selection and more by genetic correlations or environmental trade-offs. Drift, mutations, environmental plasticity or hybridization are different sources of variation which can affect the accuracy of similarity and dissimilarity of phenotypes within and between species of the mimicry ring. In addition, specific genetic architectures can favour convergence and a good level of resemblance. In *Heliconius* butterflies, variations in mimetic colour patterns are largely controlled by a few Mendelian loci of large effect, often coined the "colour pattern toolbox", an architecture which may facilitate secondary convergence in wing patterns (Reed *et al.* 2011). The most striking example of mimicry achieved through a shared architecture is when alleles at colour-patterning loci are shared among co-mimics via adaptive introgression. This was documented, for instance, in the pair of species *H. timareta* and *H. melpomene* (Heliconius Genome Consortium 2012; Pardo-Diaz *et al.* 2012). In this case, colour pattern resemblance reflects the shared origin of adaptive alleles in both species, but

selection for mimicry may also play a role, first by facilitating the invasion of introgressed alleles in the receiving population, and second by favouring further refinement of resemblance in the new genome. Describing variations in resemblance between species is therefore required to disentangle the relative importance of shared genetics and mimicry selection in the evolution of accurate resemblance.

In this study, we investigate selection for mimicry perfection by quantifying phenotypic similarity among multiple species forming the so-called "postman" mimicry ring in Northern Peru.

The "postman" wing pattern is a very strong warning signal working as a major (and possibly ancient) mimicry attractor in this region (Hines *et al.* 2011). However, subtle geographic variation in colour pattern was reported for one of the co-mimics (*Heliconius timareta thelxinoe*), between localities separated by about 175 km (Mérot *et al.* 2013). The two localities display some differences in the assemblage of species participating in the postman mimicry. In the Alto Mayo, *H. timareta thelxinoe* co-occurs mostly with *H. telesiphe* while, in the Cordillera Escalera, the 'postman' community is dominated by *H. erato favorinus* and *H. melpomene amaryllis* (Fig.1). Theoretical simulations suggest that the most abundant species of a mimicry ring generally drives the evolution of phenotypic resemblance in other species (Turner 1977; Mallet 1999; Franks & Sherratt 2007; Ruxton *et al.* 2008). We therefore hypothesise that the subtle variation in colour pattern found in certain species participating in the mimicry ring might be the footprint of selection for different mimetic optima, reflecting spatial changes in the phenotypic composition of the different communities.

To test whether the local mimetic community may influence the strength and the nature of selection for mimicry, we investigate whether geographic variations for various modalities of the warning signals (colour pattern, reflectance of the colour patch, wing shape) within participating species appear to track variations in composition of the mimetic community. We use Colour Pattern Modelling (Le Poul *et al.* 2014), geometric morphometrics, and spectral colour measurements to quantify phenotypic similarity between co-mimics, and analyse it in the light of neutral molecular variation, hybridization rates, and species composition in the distinct communities.

Methods

Species studied, specimen collection and density.

The four species of the "postman" mimicry ring (Fig.1, *H. melpomene amaryllis* Feder & Feder, *H. timareta thelxinoe* Lamas & Mérot*, H. erato favorinus* Hopffer and *H. telesiphe* Doubleday*,* further abbreviated with their species name only) were sampled in two tropical montane areas separated by 175km, the Escalera and the Alto Mayo (San Martín, Peru). Sampling localities were chosen along an altitudinal continuum, which ranges from 400m to 1300m in the Escalera ("E", 06°27′28¨S; 76°17′53¨W) and from 1100m to 1800m in the Alto Mayo ("A", 05°39′58¨S; 77°44′35¨W).

To estimate the relative frequencies of each species, we used collection data corrected by the number of collecting days. On collecting days, all butterflies from the four species encountered were caught with entomological nets. A subset of this sample was used for genetic and phenotypic analysis (Table.S1&S2). We considered two populations of *H. melpomene* in the Escalera ("Low E", below 1000m; and "Escalera", above 1000m, sympatric with *H. timareta*). We also included an additional population of *H. melpomene* from

169 Moyobamba (06°05´13¨S; 76°59´36¨W, Peru) to investigate geographic variation in *H.*
170 *melpomene* between area of sympatry and allopatry with *H. timareta*.

171 ## *Phenotypic description and analyses*

172 ### Data acquisition

173 Images of ventral (v) and dorsal (d) forewings (FW) and hindwings (HW) were captured in
174 normalized light conditions (CIE Standard Illuminant D50) using a Nikon D90 digital camera
175 with a Nikon micro 105/2.8GEDVR lens.

176

177 Measurements of wing reflectance were done with a spectrometer (AvaSpec-3648, Avantes)
178 and a deuterium-halogen light source (DH-200, Avantes) connected to a 1.5mm diameter
179 sensor (FCR-7UV200-2-1.5x100, Avantes) inserted in a miniature black chamber (an opaque
180 black plastic tube surrounding the reflectance probe to exclude ambient light from the
181 measurement). Reflectance spectra were taken at 90° incidence relative to a 99% reflectance
182 standard (300-700nm spectralon) and to dark current. Spectra were recorded with the software
183 Avasoft 7.0 using an average of 5 measures with an integration time of 23 ms. On all wing
184 surface (FWv, FWd, HWv, HWd), we recorded reflectance of colour patches and black area at
185 the same location for each specimen.

186 ### Wing colour pattern

187 Colour pattern was analysed using Colour Pattern Modelling (Le Poul *et al.* 2014) which
188 allows quantifying colour pattern variation and similarity across the entire wing. Detailed
189 methods are given in supplementary material. Briefly, wing outline was extracted individually
190 from the background. Within this area, the RGB colours are categorized into three colour
191 classes (black, red or yellow). We called "yellow" the bar on all species hindwing so that
192 pattern itself would be comparable, despite some hue variation between species and between

individuals of different ages or wear. All individual wings were aligned by rotation, translation, scaling and normalization based on an iterative process. Variations in patch boundaries can then be fully described and compared between individuals.

Each wing surface (FWd, FWv, HWd, HWv), was characterized by a set of pixels with homologous position across specimens. Each pixel was associated with three presence/absence binary values for black, red and yellow, which allows colours to be treated separately. Variations in patch boundaries were then analysed for each surface separately with a principal component analysis applied to the set of pixels. The resulting components describe a morphological space used for subsequent analyses after the dimensionality reduction proposed by Baylac and Friess (2005).

Small colour elements located in the basal part of the wing were scored following Mérot *et al.* (2013) with an index describing variation at two qualitative characters: the red line on the ventral forewing (0=absent to 5=a well-marked line, Fig.S1a) and the extension and the number of basal red spots on the ventral hindwing (0=no dots to 5=large spots, Fig.S1b).

Wing outline

Wing outline, extracted and aligned through the first step CPM procedure, was further analysed with elliptical Fourier analysis using custom scripts developed in Matlab (Jones *et al.* 2013), following the directions of (Neto & Samal 2006). A PCA was applied to the first twenty Fourier harmonics and allowed describing a morphological space for wing outline.

Wing venation: analysis by geometric morphometrics

Wing venation was described using 15 (FW) and 14 (HW) landmarks, placed at vein intersections and vein termini on the ventral side, as described in Mérot *et al.* (2013). Standard tests of repeatability were done by taking the landmarks five times per wing on subsamples of five butterflies from a single species, population and sex. Landmark coordinates were digitalized using TpsDig2 (Rohlf 2010) and superimposed using a general Procrustes analysis (Bookstein 1991; Zelditch *et al.* 2004). For each set of landmarks, superposition includes all samples in a multidimensional Procrustes space whose tangent space is a Euclidian morphological space for wing venation. A principal component analysis was applied on this wing venation data followed by the dimensionality reduction proposed by Baylac and Friess (2005). Wing size was measured using log-transformed centroid size (Bookstein 1991).

Colour spectra and visual models

Colour spectra were analysed using Avicol V.6 (Gomez 2006). Colour spectra obtained between 300 and 700 nm were smoothed using a local Fourier correction at 650 nm (Fig.S5&S7). A mean colour spectrum was calculated for each colour patch of each individual by averaging measurements on the left and the right wing. Each kind of colour patch was further analysed separately.

Analyses of spectra were carried out using models of animal vision, which take into account the observer's vision and the illuminating light. Each photoreceptor is characterized by a sensitivity function which determines the wavelength of reflected light perceived by the eye. Analyses in the main text are with a tetrachromatic V-type bird visual system (Peafowl, *Pavo cristatus*, (Hart 2002)) and a light environment corresponding to large sunny gaps in a tropical

241 forest (Théry, Pincebourde & Feer 2008). Analysis with other visual system and other incident
242 lights gave consistent results and are presented in supplementary materials (Fig.S9). A
243 physiological model of Endler and Mielke (2005) was applied to visualize colour distribution
244 in an unconstrained space, a tetrahedron whose vertices corresponds to the four
245 photoreceptors (Fig.S6). Within this colour space, the relative location of each individual's
246 patch colour can be compared and the overlap between two clouds of points (Table.S6) can be
247 calculated. Relative excitation of each photoreceptor were treated as multivariate data and
248 analysed with PCA analyses.
249

250 ## Statistical analysis

251 Within each species, differences in colour pattern, colour spectra, wing outline and venation
252 between geographic populations were tested by a one-way MANOVA on each subset of PCs
253 with geographic origin as a factor. Then, within each species and for each trait, we tested
254 discrimination between groups defined by geographical populations with a linear discriminant
255 analysis. We compared the cross-validation values to discrimination between simulated
256 populations of similar size following the methods described in Evin *et al.* (2013).
257

258 Phenotypic similarity between each species and each population of *H. timareta* was quantified
259 with several methods within each morphological space. First, to measure the magnitude of
260 phenotypic similarity between individuals, Euclidian distances in the morphological space
261 were computed between all pairs of specimens. For colour spectra, we used perceptual
262 chromatic distances between all pairs of specimen, expressed in JNDs (just noticeable
263 differences) in a perception model of Vorobyev and Osorio (1998). This model takes into
264 account the noise due to errors in photoreceptor response and estimates more accurately
265 whether the discriminability between colours could really be perceived by predators.

Perceptual and Euclidian distances between a mimic and each population of *H. timareta* (Alto Mayo *vs.* Escalera) were analysed with a mixed model with geographic population as factor and identity of the compared specimen as random factors. Second, to compute an indicator of similarity between groups, Mahalanobis distances were calculated between the three species and the two populations of *H. timareta*. Contrary to Euclidian distances, this measure applies in a transformed space and is scaled by intra-group variation. Third, to confirm our similarity estimates, we calculated an index of cross-classification by performing a linear discriminant analysis on a subset of two groups (for example, *H. melpomene/H. timareta* (A); *H. melpomene/H. timerata* (E)). Since the latter two analyses are sensitive to unbalanced sample sizes (Kovarovic 2011), we performed them on subsets of similar sizes, by randomly drawing 30 specimens from each group, repeating the procedure 1000 times and averaging over the repetitions.

All analyses were performed in R 3.0.3. (R Core Team 2014) using the packages *ade4 , nlme , Mass* and *Rmorph* (Baylac 2012).

Multilocus microsatellite analysis

The multilocus microsatellite genotype, performed on *H. melpomene* and *H. timareta,* is based on 11 loci, and is an extension of the genotyping performed in Mérot et al (2013) with the addition of 31 *H. melpomene amaryllis* and 69 *H. timareta thelxinoe*. Detailed methods can be found in supplementary materials; genetic diversity and statistics in Table.S3. F-statistics were calculated using GENETIX 4.05 (Belkhir *et al.* 1996-2004). We used STRUCTURE (Pritchard, Stevens & Donnelly 2000), a multilocus Bayesian clustering method, to determine population structure, to confirm the assignment of individuals to species, and to detect admixed genotypes. STRUCTURE was run with 500,000 updates of the Markov chain after

an initial burn-in of 50,000 updates, to achieve chain convergence for a set of models with different numbers of clusters (K=1-6). Since *H. melpomene* and *H. timareta* can hybridize and disperse, we used the 'admixture model' and 'correlated allele frequencies'. The most likely *K* was chosen using the maximum likelihood value in STRUCTURE and the *ΔK*-method (Evanno, Regnaut & Goudet 2005) and then evaluated following recommendations from STRUCTURE documentation. Over the entire set of specimens included, the likelihood reaches a maximum plateau between *K=2* and *K*=3. When *K=2,* the two clusters correspond to the two species (identified phenotypically), so *K=2* bears a clear biological meaning and was retained. *K*=3 further splits *H. timareta* into two clusters with no obvious association with biological variables (such as geographic populations), nor with any identifiable experimental bias. Further STRUCTURE analyses run on each species separately support the absence of intraspecific clustering. Posterior probabilities of being member of a cluster were estimated and allowed detecting potential hybrids. Hybrid detection was also run separately with NewHybrids 1.0 (Anderson & Thompson 2002). Relying on the results from simulated hybrids (Mérot et al. 2013) and hybrids from controlled crosses (Mérot *et al.* 2015), individuals were considered as "pure" if the posterior probability was above 0.9 in STRUCTURE and above 0.7 in NewHybrids.

Results

Geographical variation within species

Populations of *H. timareta* from the Escalera and the Alto Mayo exhibit significant differences in colour spectra, pattern, wing outline and venation (Table.1). This result was consistent whether males and females were treated together or separately. A linear discriminant analysis on *H. timareta* allows discriminating the two populations, with a low level of misclassifications, better than the range of mis-classification obtained from randomly

316 simulated population (Table.1). Centroid sizes did not exhibit significant differences between
317 the two populations of *H. timareta* (ANOVA FW p=0.12, HW p=0.77).
318
319 Within *H. melpomene* and within *H. erato,* neither colour spectra, wing outline nor size
320 displayed any significant geographical phenotypic variations (Table.S4&S5). For pattern and
321 venation, slightly significant differences were observed between populations of *H. melpomene*
322 and *H. erato*. However, discrimination between geographic populations reaches 41-71% of
323 misclassifications, which is within the misclassification range from randomly simulated
324 populations of similar sample size.
325

326 ### *Relative frequency of postman co-mimics*

327 The Escalera and the Alto Mayo areas display differences in their geography and topography,
328 their community composition and the altitudinal ranges of local species (Fig.1C). The
329 Escalera is a relatively thin, mid-elevation Eastern cordillera jutting from adjacent Amazonian
330 lowlands. On its slopes below 1000m, *H. erato* and *H. melpomene* are the only species with a
331 postman pattern and are generally very abundant, at our collection places and elsewhere in the
332 foothills (>200 specimen of each species in 65 days, approximately 1:1 ratio, Table.S2). In the
333 highest areas, between 1000 and 1300m, *H. erato* becomes less abundant while *H. melpomene*
334 co-occurs with *H. timareta.* The Escalera is the southernmost tip of the known distribution of
335 *H. timareta* and the latter can be as abundant as *H. melpomene,* but only locally, as observed
336 at our collection places (1:1 ratio, Table.S2). At a finer scale, *H. timareta* gradually increases
337 in frequency relative to *H. melpomene* (Fig.1C). Despite intensive collection (83 days,
338 Table.S2), only two specimens of *H. telesiphe* were found in the Escalera, and only at the
339 highest collection station (1300m).
340

The Alto Mayo represents a wider, higher and more continuous area of montane forest surrounded by Andean areas. There, *H. melpomene* and *H. timareta* display less overlap since the transition is more abrupt and occurs at higher elevation (1300m, Fig.1C). *H. timareta* was never found below 1200m and was locally abundant at one collection station (1400-1600m). Over 1300m, *H. telesiphe* is an abundant species in the Alto Mayo and the commonest postman-patterned species. *H. erato* was not found in our standardized collection stations of the Alto Mayo although it is occasional up to 1200m, and thus parapatric with *H. timareta.*

Variation in wing phenotype in the postman mimetic community

Wing colour pattern

For forewing colour pattern, the first axis of the PCA (Fig.2a) displays variation from the large rounded red patch of *H. erato* and *H. melpomene,* to the dislocated, slender patches *H. telesiphe*. *H. timareta* is intermediate on this axis with the Escalera population closer to *H. melpomene* and *H. erato*. *H. telesiphe* is sitting in a quite distant position overall, because it exhibits a small additional subapical red patch on the forewing, missing in all other species. Nevertheless, the Alto Mayo population of *H. timareta* is the closest population to *H. telesiphe* in the colour pattern space because of its somewhat dislocated, zigzagging red patch shape, reminiscent of the main red patch of *H. telesiphe*.

On the hindwing, all four species display a yellow/white pattern barring both sides of the hindwing longitudinally. The first PC shows variation in the width of the bar. *H. melpomene* and *H. erato* have a rather wide and long yellow bar while *H. telesiphe* has a thinner and shorter bar (Fig.2b). Again, *H. timareta* appears in intermediate position along this direction of variation, with the Escalera population having a wide yellow bar like *H. melpomene*, while Alto Mayo populations show a thinner bar like *H. telesiphe.* The second PC shows a

366 continuous gradient from a long and pointed bar seen in *H. erato* to the thin bar characteristic
367 of *H. timareta* from the Alto Mayo, straight and wide in the proximal part and curved
368 posteriorly in the distal part. *H. melpomene* and *H. timareta* from the Escalera sit in an
369 intermediate position.
370
371 Results were consistent on the ventral colour pattern (Fig.S2) and whether males and females
372 were treated together or separately (Fig.S3).
373

374 ## Small elements of pattern (Fig.S1)

375 All specimens of *H. telesiphe* exhibit bright and large red spots on the ventral hindwing and a
376 well-marked red line on the ventral forewing (index=5). By contrast, *H. erato* nearly never
377 displays red dots or lines except in a few specimens (4-18%, depending on population) that
378 possess small faded spots or a slight line (index=0-1). Along this gradient between *H. erato*
379 and *H. telesiphe,* most *H. melpomene* samples are closer to the *H. erato* phenotype, with no
380 red line (89-100% depending on the population) and small faded spots (89-96%). Then, 45 to
381 47% of *H. timareta* samples from the Escalera exhibit small red spots (index=2-3) and no or a
382 tiny red line (index=0-3), making them closer to the *H. erato* and *H. melpomene* phenotypes.
383 On the contrary, a well-marked red line and large red dots (index=4-5) are observed in more
384 than 93% of *H. timareta* samples from the Alto Mayo (*vs.* only 53-55% from the Escalera
385 population).
386

387 ## Wing venation and outline

388 For forewing shape, results were congruent whether described by venation or by wing outline.
389 For both venation and outline, the first PCs display continuous variations from the rather
390 rounded forewings (Fig.3ac) of *H. melpomene,* shared by *H. timareta* from the Escalera,

391 towards the thinner and more elongated wings of *H. telesiphe*. By contrast, *H. timareta* from
392 the Alto Mayo display resemblance to *H. telesiphe* through its more elongated wing shape. In
393 the venation analysis, which takes into account the discal cell and the distal veins, the second
394 PC recapitulates phylogenetic relationships, separating *H. erato* and *H. telesiphe* from *H.*
395 *melpomene* and *H. timareta.*
396
397 For the hindwing shape, the first PCs of both the venation and outline analysis are mostly
398 associated with sexual dimorphism (presence/absence of male androconia), but the second
399 PCs display the same variation as the forewing, from the rounded wings of *H. melpomene* to
400 the thin, elongated wings of *H. telesiphe* (Fig.3bd). Both *H. timareta* populations are in
401 intermediate position.
402

403 ## Colour spectra

404 For both forewing and hindwing, the reflectance of colour patches also follows a continuum
405 between the four species from *H. erato* and *H. melpomene* to *H. telesiphe* (Fig.4). The two
406 populations of *H. timareta* occupy different positions along this continuum in the colour
407 space. On the one hand, the Escalera population displays similar colours to *H. melpomene*
408 and *H. erato* (red-orange and bright yellow), with which it overlaps in the three-dimensional
409 colour space (Fig.S6, Table.S6). On the other hand, the position of *H. timareta* from the Alto
410 Mayo comes closer to *H. telesiphe* for the red forewing patches (with a deeper red) and, to a
411 lesser extent, for the whitish hindwing patch.
412
413 It is notable that colours are generally lighter on the ventral side than on the dorsal side and
414 exhibit more intra- and inter-specific variability. On the ventral side, both red and yellow
415 patches of *H. timareta* and *H. telesiphe* reflect more short-wavelengths and UV than *H. erato*

and *H. melpomene*. The higher variability in colour on the ventral side is also shown by the analyses of perceptual distances, since discrimination was higher for the ventral side than for the dorsal side. On the ventral side, visual contrast within species was smaller than between species, for both red and yellow and reaches 2 to 6 JND (Fig.5ij). On the dorsal side, within species and between species visual distances spanned a similar range (2 JND, Fig.5gh), except between *H. telesiphe* and all others for the white dorsal hindwing patch.

Similarities in wing phenotype in the postman mimetic community

For forewing and hindwing colour patterns, forewing shape and hue of all patches, resemblance indicators suggest that the Escalera population of *H. timareta* is more similar to the most abundant local co-mimics of the Escalera (*H. erato, H. melpomene*) than the Alto Mayo population is. Euclidian and perceptual distances as well as Mahalanobis distances between *H. timareta* and *H. erato* or *H. melpomene* are smaller for the Escalera population than for the Alto Mayo population (Table.S7, Fig.5). The reverse was found for distances with *H. telesiphe*, suggesting that the Alto Mayo population of *H. timareta* is more similar to *H. telesiphe* than the Escalera population is. Those results are supported by the differences in cross-classification rates from the discriminant analysis (Table.S8). Results were consistent for males and females taken independently and together.

For hindwing shape, similarity estimates do not show any consistent trend differentiating the two geographic populations of *H. timareta*.

Genetic divergence and population structure

No genetic differentiation was found between the two geographic populations of *H. timareta.* The F_{ST} value over all loci was estimated at 0.025 and the two populations do not split into

441 different genetic clusters. For *H. melpomene,* no genetic differentiation was found either, F_{ST}
442 reaches only 0.009 between the four populations and they do not display any genetic
443 clustering.
444
445 Genetic differentiation between *H. melpomene* and *H. timareta* was significant in the two
446 geographic areas and of similar magnitude (A: F_{ST}=0.148; E: F_{ST}=0.157).
447

448 ### *Hybrid detection*

449 The level of admixture follows the expected trend, with slightly higher rates in the Escalera,
450 where *H. timareta* and *H. melpomene* populations show more altitudinal overlap and are the
451 most phenotypically similar. However, this difference was not significant, possibly due to our
452 limited samples sizes (test of equal proportions: p=0.67).
453
454 Bayesian clustering analyses with STRUCTURE detected 4.2% of admixed individuals in the
455 Escalera (Fig.S11, 2 back-crosses to *melpomene,* 2 back-crosses to *timareta* and 3 F1 hybrids
456 out of 167 specimens), and 3.2% in the Alto Mayo (1 back-cross to *melpomene,* 1 F1 or back-
457 cross to *timareta* out of 63 specimens). Surprisingly, given that *H. timareta* does not occur
458 below 1000m, one admixed specimen was detected in the low Escalera population at about
459 700m of elevation. Assignment tests with New Hybrids1.0 found the same admixed
460 individuals except in one case, but also pointed out six more putative back-crosses, raising the
461 proportion of admixed individuals to 8.7% in the Escalera and 6.3% in the Alto Mayo.
462

Discussion

Tracking variations in the local prey environment

Our results show that despite an absence of genetic differentiation at neutral markers, geographic populations in one of the species, *H. timareta,* display subtle but consistent variations in wing phenotype, associated with enhanced similarity to locally abundant co-mimics. Our results therefore denote a geographic shift in mimicry association consistent with quantitative changes in the composition of the mimicry community.

Geographic mosaics of co-varying warning pattern are commonly observed in animal clades involved in mimicry associations, from butterflies to frogs (Thompson 2005; Twomey *et al.* 2013). An iconic case is the faithful co-variation of *Heliconius melpomene* and *H. erato* throughout their shared range, the two butterflies switching in concert from red postman pattern (in certain Andean valleys, south-eastern Brazil, and Carribean area) to orange-rayed Amazonian patterns (Sheppard *et al.* 1985; Reed *et al.* 2011). Our results suggest that co-mimetic species not only exhibit adaptive shifts of colour pattern but can also track one another geographically for more subtle variations within a type of pattern. Quantitative variations in pattern and hue appear adaptive and match local specificities of the surrounding communities of co-mimetic prey. Such partial phenotypic divergence represents an intermediate situation between undifferentiated populations and colour pattern races participating in disjoint mimicry assemblages, opening a window on the process by which diversity within species may evolve. This also brings support to the hypothesis that the variation in abundance of certain key species can influence the local mimicry optimum in a way that translates into changes in the selected phenotypes of other species, a process which was proposed to contribute to mimicry diversification in the face of selection for resemblance (Turner & Mallet 1996).

Geographic variations were not found, at our scale, within *H. melpomene amaryllis* nor *H. erato favorinus,* two species which are generally abundant and widespread compared to *H. timareta.* This observation fits the prediction that the commoner species should be less influenced by selection for perfect resemblance in a Müllerian association (Ruxton, Sherratt & Speed 2004; Ruxton *et al.* 2008). However, in other cases where *H. melpomene* is not the dominant species, such as in the Amazon lowlands, subtle geographic variations are reported where they match the local butterfly communities. For instance, among the orange-rayed Amazonian forms, local populations from the Marañon valley exhibit a partial lack of orange rays, which presumably enhance their resemblance to the commonest local species, *H. himera* (J. Mallet, pers. com.). Similarly, the two parapatric races named *H. melpomene aglaope* and *H. m. malleti* represent a geographic continuum of populations whose wing pattern varies quantitatively in the width and shape of the yellow forewing patch, matching similar variations in the patterns of local *H. erato* populations and likely other co-mimetic species. Together with those examples, our results suggest that coordinated quantitative co-variations of the elements of the warning signal are adaptive and, despite receiving little quantitative attention, reflect the tracking of a geographically changing optimum defined by a community of multiple mimetic species.

The multimodality of traits involved in mimicry

Resemblance between populations of *H. timareta* and the local co-mimetic species appear congruent over the different traits measured in our study: colour pattern, hue and wing shape, suggesting a multimodal and coordinated evolution of resemblance. This is consistent with recent studies showing that, although colour might have a greater influence on predator

512 decision, pattern features also play a significant role in signalling (Finkbeiner, Briscoe &
513 Reed 2014).
514
515 Whether wing shape is also under selection for mimicry is unclear. Wing venation itself
516 cannot presumably be seen by predators but it may be indirectly under selection since wing
517 motion and wing outline are associated with mimicry (Srygley 1994; Jones *et al.* 2013). Long
518 and narrow forewings (high aspect ratio) are generally associated with fast and extended flight
519 while broader wings may be better suited for slow and more fluttery flights (Betts & Wootton
520 1988), and *H. telesiphe* displays a notably faster and more elusive flight than *H. erato* and *H.*
521 *melpomene* (pers. obs.). Flight behaviours may be perceived by bird predators as informative
522 characters for prey identification and may participate in the definition of the local mimicry
523 optimum (Srygley & Ellington 1999), so the similarity in shape of *H. timareta* to *H. telesiphe*
524 in the Alto Mayo and to *H. melpomene* in the Escalera may represent an adaptive response to
525 selection on flight patterns.
526
527 A striking result from the spectral analysis is that, although both wing sides show a signal of
528 local mimicry adaptation, hue on the dorsal side is less variable (within and between species)
529 than on the ventral side. In butterflies, ventral and dorsal patterns can be genetically correlated
530 as shown by experimental evolution lines of *Bicyclus anynana* (Beldade, Koops & Brakefield.
531 2002) but also strikingly different, as in *Morpho* for instance, which exhibit a bright blue
532 dorsal side and a camouflaged ventral side. Pattern differences between the ventral and dorsal
533 sides then generally result from different selective pressures. Here, the lower variance
534 measured on dorsal hue presumably reflects the fact that dorsal patterns in *Heliconius* are
535 more frequently visible to predators in full light during sunny hours, when the butterflies are
536 flying or basking with wings spread out, translating into stronger natural selection on hue

resemblance. By contrast, ventral patterns are exhibited when resting in the shadow or roosting at night or at dusk or dawn (Mallet 1986; Finkbeiner, Briscoe & Reed 2012), in a light environment where colour differences are less noticeable than in sunlight (Théry, Pincebourde & Feer 2008). Inter-specific differences in ventral hue, especially in UV reflectance, might also be associated with intra-specific social and sexual communication (Silberglied & Taylor 1978).

The evolution of accurate mimicry

Most variations in resemblance quantified between co-mimics here are quite subtle. All populations belong to the general "postman" wing pattern but the precision of mimicry is enhanced between taxa from the same locality, as expected if predators select for mimicry refinement. A long-standing debate is whether evolution takes place by small gradual or large changes (Punnett 1915; Fisher 1927; Nicholson 1927; Fisher 1930). Mimicry evolution is classically suggested to occur first via a large phenotypic change which allows coarse resemblance, then by a phase of more subtle refinement (Sheppard *et al.* 1985; Turner 1987; Franks & Sherratt 2007; Balogh & Leimar 2010; Leimar & Mallet 2012). Introgression of alleles of the *optix* locus from *H. melpomene,* controlling the presence of the red forewing patch, is thought to have been a major step allowing *H. timareta* to join the postman mimicry ring (Heliconius Genome Consortium 2012; Pardo-Diaz *et al.* 2012). Our results describe the second phase of mimicry evolution in which small variations in shape, colour and pattern improve the match to the local optimum. Those mimicry improvements may have evolved through the selection of allelic variants at *optix* and other unlinked loci modifying wing morphology and pattern. In *H. melpomene,* the dislocated *vs.* rounded shape variation of the red forewing patch is associated with a quantitative trait locus (QTL) unlinked to *optix* (Baxter, Johnston & Jiggins 2008). Differential selection towards different mimetic optima

562 determined by changes in the local species communities would then be predicted to affect
563 such "secondary" mimicry-refining loci which act epistatically with the major mimicry
564 switches already identified.
565
566 ### *A role for gene flow?*
567 The process of gradual evolution is particularly efficient in populations with a rich supply of
568 genetic variation for traits contributing to the warning signal (Ruxton *et al.* 2008). Apart from
569 mutation and intra-specific migration, gene flow from closely-related species through
570 hybridization may be an important source of variation (Grant & Grant 1994). Hybridization
571 with local species may bring new, locally adapted alleles and our data show that a certain
572 percentage of F1 hybrids and admixed individuals between *H. melpomene* and *H. timareta* are
573 found in natural populations. Genome-wide signatures of past and ongoing gene flow are
574 documented between *H. melpomene* and *H. timareta*, and introgressed alleles of the red-
575 patterning gene *optix* from *H. melpomene* determine the general postman wing pattern in *H.*
576 *timareta* of Northern Peru (Heliconius Genome Consortium 2012; Pardo-Diaz *et al.* 2012;
577 Martin *et al.* 2013). The occurrence and timing of the introgression need to be clarified more
578 finely to assess its role in the differentiation we report between the two locations. For
579 instance, one hypothesis is that resemblance may have had more time to evolve in one
580 population if the postman pattern is older there. Similarly, recurrent interspecific gene flow
581 for adaptive alleles at *optix* and other loci readily favoured by selection may participate in the
582 refinement of mimicry in places where *H. melpomene* determines the position of the optimum
583 and overlaps with *H. timareta*.
584
585 One intriguing result here is that traits which might not be readily perceived by predators
586 (small red spots, costal line, forewing venation) also display a higher similarity between *H.*

587 *timareta* and *H. melpomene* in the Escalera than in the Alto Mayo. If those traits are indeed
588 subject to weaker selection by predators (cf. the stronger resemblance on dorsal phenotypic
589 traits), their similarity might reflect local gene flow between hybridizing *H. melpomene* and
590 *H. timareta.* Our genetic analysis does not support higher gene flow in the Escalera than in the
591 Alto Mayo, but the resolution is limited by the small number of markers. We suggest it would
592 be worth exploring the contribution to resemblance of inter-specific gene flow in this and
593 similar systems.
594

595 ## Conclusion

596 Here we showed that subtle variations in wing pattern across short geographic distances allow
597 populations of co-mimics to track gradual changes in the mimicry optimum, which suggest
598 that predators do not generalise widely and that selection is operating to maximise mimicry
599 efficiency even between relatively similar warning signals. This demonstrates that fitness
600 peaks in the morphological (signalling) space are sharp and that the quantitative movements
601 of the position of the optimum are of strong adaptive significance for certain species of the
602 community. Effectively, this amounts to *H. timareta* switching mimetic association and to our
603 knowledge, this is the first quantification of such a mimicry switch in a community context.
604 All traits quantified here display coordinated variations, suggesting multimodal and gradual
605 evolution of multiple traits improving the level of resemblance and describing the second
606 phase of colonisation of a new fitness peak. Our study also raises the possibility that the
607 observed geographic pattern of adaptation and the accuracy of mimicry might partly be
608 enhanced by interspecific gene flow providing adaptive alleles readily favoured by selection.
609

Acknowledgments

We thank V. Llaurens and D. Gomez for their help with colour analysis, A. Evin and V. Debat for their help with geometric morphometrics and people involved in fieldwork. We are grateful to Jim Mallet for introducing us to this study system and for helpful comments on this manuscript. We are grateful to Evan Twomey and one anonymous reviewer for their valuable comments on this manuscript. We thank the Peruvian Ministerio de la Agricultura (SERFOR), the SERNANP-BPAM and PEHCBM-ACR-CE and the Museo de Historia Natural for research and export permits in Peru. This work was supported by ERC Starting Grant MimEvol and ANR-JCJC HybEvol to MJ.

Data accessibility

Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.h4j6c

References

Anderson, E.C. & Thompson, A. (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics,* **160,** 1217-1229.

Balogh, A.C.V., Gamberale-Stille, G., Tullberg, B. S. & Leimar, O. (2010) Feature Theory and the Two-step Hypothesis of Müllerian Mimicry Evolution. *Evolution,* **64,** 810-822.

Baxter, S.W., Johnston, S.E. & Jiggins, C.D. (2008) Butterfly speciation and the distribution of gene effect sizes fixed during adaptation. *Heredity***,** 1-9.

Baylac, M. (2012) Rmorph: a R geometric and multivariate morphometrics library. *Available from the author: baylac@mnhn.fr*.

Baylac, M. & Friess, M. (2005) Fourier descriptors, procrustes superimposition, and data dimensionality: an example of cranial shape analysis in modern human populations. In: Slice DE, ed. Modern morphometrics in physical anthropology,part 1 theory and methods. New York, NY: Kluwer Academic/Plenum Publishers, 142-165.

Beatty, C.D., Beirinckx, K. & Sherratt, T.N. (2004) The evolution of mullerian mimicry in multispecies communities. *Nature,* **431,** 63-67.

Beldade, P., Koops, K. & Brakefield., P. (2002) Developmental constraints versus flexibility in morphological evolution. *Nature,* **416,** 844-847.

Belkhir, K., Borsa, P., Chikhi, L., Raufatse, N. & Bonhomme, F. (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

644 Betts, C.R. & Wootton, R.J. (1988) Wing shape and flight behaviour in butterflies
645 (Lepidoptera: Papilionoidea and Hesperioidea): a preliminary analysis. *Journal of*
646 *Experimental Biology,* **138(1),** 271-288.
647 Bookstein, F. (1991) Morphometrics tools for landmark data: geometry and biology. New
648 York, NY: Cambridge University Press.
649 Chouteau, M. & Angers, B. (2011) The Role of Predators in Maintaining the Geographic
650 Organization of Aposematic Signals. *American Naturalist,* **178,** 810-817.
651 Chouteau, M., Arias, M. & Joron, M. (2016) Warning signals are under positive frequency-
652 dependent selection in nature. *Proceedings of the national Academy of Sciences,* **113,**
653 2164-2169.
654 Endler, J.A. & Mielke, P.W. (2005) Comparing entire colour patterns as birds see them.
655 *Biological Journal of the Linnean Society,* **86,** 405-431.
656 Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals
657 using the software STRUCTURE: a simulation study. *Molecular Ecology,* **14,** 2611-
658 2620.
659 Evin, A., Cucchi, T., Cardini, A., Vidarsdottir, U.S., Larson, G. & Dobney, K. (2013) The long
660 and winding road: identifying pig domestication through molar size and shape.
661 *Journal of Archaeological Science,* **40,** 735-743.
662 Finkbeiner, S.D., Briscoe, A.D. & Reed, R.D. (2012) The benefit of being a social butterfly:
663 communal roosting deters predation. *Proceedings of the Royal Society B-Biological*
664 *Sciences,* **279,** 2769-2776.
665 Finkbeiner, S.D., Briscoe, A.D. & Reed, R.D. (2014) Warning signals are seductive: Relative
666 contributions of color and pattern to predator avoidance and mate attraction in
667 Heliconius butterflies. . *Evolution,* **68,** 3410-3420.
668 Fisher, R. (1930) The genetical theory of natural selection. . A complete variorum edition.
669 Oxford Univ. Press, Oxford, UK.

670 Fisher, R.A. (1927) On some objections to mimicry theory; statistical and genetic. *Trans. R.*
671 *Entomol. Soc.,* **75,** 269-278.
672 Franks, D.W. & Sherratt, T.N. (2007) The evolution of multicomponent mimicry. *Journal of*
673 *Theoretical Biology,* **244,** 631-639.
674 Gomez, D. (2006) AVICOL, a program to analyse spectrometric data. Last update october
675 2011. Free executable available at http://sites.google.com/site/avicolprogram/ or from
676 the author at dodogomez@yahoo.fr.
677 Grant, P.R. & Grant, B.R. (1994) Phenotypic and genetic effects of hybridization in Darwin's
678 finches. *Evolution,* **48,** 297-316.
679 Hart, N.S. (2002) Vision in the peafowl (Aves : *Pavo cristatus*). *Journal of Experimental*
680 *Biology,* **205,** 3925-3935.
681 Heliconius Genome Consortium (2012) Butterfly genome reveals promiscuous exchange of
682 mimicry adaptations among species. *Nature,* **487,** 94-98.
683 Hines, H.M., Counterman, B.A., Papa, R., de Moura, P.A., Cardoso, M.Z., Linares, M.,
684 Mallet, J., Reed, R.D., Jiggins, C.D., Kronforst, M.R. & McMillan, W.O. (2011) Wing
685 patterning gene redefines the mimetic history of Heliconius butterflies. *Proceedings of*
686 *the National Academy of Sciences, USA,* **108,** 19666-19671.
687 Ihalainen, E., Lindstrom, L., Mappes, J. & Puolakkainen, S. (2008) Can experienced birds
688 select for Mullerian mimicry? *Behavioral Ecology,* **19,** 362-368.
689 Ihalainen, E., Rowland, H.M., Speed, M.P., Ruxton, G.D. & Mappes, J. (2012) Prey
690 community structure affects how predators select for Mullerian mimicry. *Proceedings*
691 *of the Royal Society B-Biological Sciences,* **279,** 2099-2105.
692 Jones, R., Poul, Y.L., Whibley, A., Mérot, C., FFrench-Constant, R. & Joron, M. (2013) Wing
693 shape variation associated with mimicry in butterflies. *Evolution,* **67(8),,** 2323-2334.

694 Joron, M. & Iwasa, Y. (2005) The evolution of a Mullerian mimic in a spatially distributed
695 community. *Journal of Theoretical Biology,* **237,** 87-103.
696 Kapan, D.D. (2001) Three-butterfly system provides a field test of müllerian mimicry. *Nature,*
697 **409,** 338-340.
698 Kovarovic, K., Aiello, L.C., Cardini, A., Lockwood, C.A., (2011) Discriminant function
699 analyses in archaeology: Are classification rates too good to be true? . *Journal of*
700 *Archaeological Science,* **38**.
701 Langham, G.M. (2004) Specialized avian predators repeatedly attack novel color morphs of
702 Heliconius butterflies. *Evolution,* **58,** 2783-2787.
703 Le Poul, Y., Whibley, A., Chouteau, M., Prunier, F., Llaurens, V. & Joron, M. (2014)
704 Evolution of dominance mechanisms at a butterfly mimicry supergene. . *Nature*
705 *Communications,* **5**.
706 Leimar, O., Tullberg, B. S. & Mallet, J. (2012) The Adaptive Landscape in Evolutionary
707 Biology
708 Mallet, J. (1986) Gregarious roosting and home range in *Heliconius* butterflies. *National*
709 *Geographic Research,* **2,** 198-215.
710 Mallet, J. (1999) Causes and consequences of a lack of coevolution in mullerian mimicry.
711 *Evolutionary Ecology,* **13,** 777-806.
712 Mallet, J. & Barton, N.H. (1989) Strong Natural-Selection in a Warning-Color Hybrid Zone.
713 *Evolution,* **43,** 421-431.
714 Mallet, J. & Joron, M. (1999) Evolution of diversity in warning color and mimicry:
715 Polymorphisms, shifting balance, and speciation. *Annual Review of Ecology and*
716 *Systematics,* **30,** 201-233.
717 Martin, S.H., Dasmahapatra, K.K., Nadeau, N.J., Salazar, C., Walters, J.R., Simpson, F.,
718 Blaxter, M., Manica, A., Mallet, J. & Jiggins, C.D. (2013) Genome-wide evidence for
719 speciation with gene flow in *Heliconius* butterflies. *Genome Research,* **23,** 1817-1828.
720 Mérot, C., LePoul, Y., Théry, M., Joron, M. 2016. Data from: Mimicry refinement:
721 Phenotypic variations tracking the local optimum. Dryad Digital Repository.
722 doi:10.5061/dryad.h4j6c
723 Mérot, C., Frérot, B., Leppik, E. & Joron, M. (2015) Beyond magic traits: Multimodal mating
724 cues in *Heliconius* butterflies. . *Evolution,* **69,** 2891-2904.
725 Mérot, C., Mavarez, J., Evin, A., Dasmahapatra, K.K., Mallet, J., Lamas, G. & Joron, M.
726 (2013) Genetic differentiation without mimicry shift in a pair of hybridizing
727 *Heliconius* species (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean*
728 *Society,* **109,** 830-847.
729 Merrill, R.M., Wallbank, R.W.R., Bull, V., Salazar, P.C.A., Mallet, J., Stevens, M. & Jiggins,
730 C.D. (2012) Disruptive ecological selection on a mating cue. *Proceedings of the Royal*
731 *Society B-Biological Sciences,* **279,** 4907-4913.
732 Müller, F. (1879) *Ituna* and *Thyridia*: A remarkable case of mimicry in butterflies.
733 *Transactions of the entomological society. London.*
734 Neto, J.C., Meyer, G. E., Jones, D. D. & Samal, A.K. (2006) Plant species identification
735 using Elliptic Fourier leaf shape analysis.*Computers and Electronics in Agriculture,*
736 **50,** 121-134.
737 Nicholson, A.J. (1927) A new theory of mimicry in insects. . *Australian zoologist,* **5,** 10-104.
738 Pardo-Diaz, C., Salazar, C., Baxter, S.W., Merot, C., Figueiredo-Ready, W., Joron, M.,
739 McMillan, W.O. & Jiggins, C.D. (2012) Adaptive Introgression across Species
740 Boundaries in *Heliconius* Butterflies. *PLOS Genetics,* **8,** 13.
741 Penney, H.D., Hassall, C., Skevington, J.H., Abbott, K.R. & Sherratt, T.N. (2012) A
742 comparative analysis of the evolution of imperfect mimicry. *Nature,* **483,** 461-U110.

743 Pritchard, J.K., Stevens, M. & Donnelly, P. (2000) Inference of population structure using
744 multilocus genotype data. *Genetics,* **155,** 945-959.
745 Punnett, R.C. (1915) *Mimicry in butterflies.* . Cambridge Univ. Press, Cambridge.
746 R Core Team (2014) R: A language and environment for statistical computing. R Foundation
747 for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
748 Reed, R.D., Papa, R., Martin, A., Hines, H.M., Counterman, B.A., Pardo-Diaz, C., Jiggins,
749 C.D., Chamberlain, N.L., Kronforst, M.R., Chen, R., Halder, G., Nijhout, H.F. &
750 McMillan, W.O. (2011) *Optix* drives the repeated convergent evolution of butterfly
751 wing pattern mimicry. *Science 333: 1137-1141.*
752 Rohlf, F. (2010) *TPSDig 2.16. Department of Ecology and Evolution, State University of New*
753 *York at Stony Brook, Stony Brook, NY.*
754 Rowe, C., Lindstrom, L. & Lyytinen, A. (2004) The importance of pattern similarity between
755 Mullerian mimics in predator avoidance learning. *Proceedings of the Royal Society B-*
756 *Biological Sciences,* **271,** 407-413.
757 Ruxton, G.D., Franks, D.W., Balogh, A.C.V. & Leimar, O. (2008) Evolutionary Implications
758 of the Form of Predator Generalization for Aposematic Signals and Mimicry in Prey.
759 *Evolution,* **62,** 2913-2921.
760 Ruxton, G.D., Sherratt, T.N. & Speed, M.P. (2004) Avoiding attack: the evolutionary ecology
761 of crypsis, warning signals and mimicry. *Avoiding attack: the evolutionary ecology of*
762 *crypsis, warning signals and mimicry.*, pp. i-xii, 1-249.
763 Sheppard, P.M., Turner, J.R.G., Brown, K.S., Benson, W.W. & Singer, M.C. (1985) Genetics
764 and the evolution of muellerian mimicry in *Heliconius* butterflies. *Philosophical*
765 *Transactions of the Royal Society B-Biological Sciences,* **308,** 433-613.
766 Silberglied, R.E. & Taylor, O.R. (1978) Ultraviolet reflection and its behavioral role in
767 courtship of sulfur butterflies *Colias eurytheme* and *C. philodice* (Lepidoptera,
768 Pieridae). *Behavioral Ecology and Sociobiology,* **3,** 203-243.
769 Srygley, R.B. (1994) Locomotor mimicry in butterflies? The associations of positions of
770 centers of mass among groups of mimetic, unprofitable prey. *Philosophical*
771 *Transactions of the Royal Society B-Biological Sciences,* **343,** 145-155.
772 Srygley, R.B. & Ellington, C.P. (1999) Discrimination of flying mimetic, passion-vine
773 butterflies Heliconius. *Proceedings of the Royal Society B-Biological Sciences,* **266,**
774 2137-2140.
775 Théry, M., Pincebourde, S. & Feer, F. (2008) Dusk light environment optimizes visual
776 perception of conspecifics in a crepuscular horned beetle. *Behavioral Ecology,* **19,**
777 627-634.
778 Thompson, J.N. (2005) *The geographic mosaic of coevolution.* . University of Chicago Press.
779 Turner, J.R. (1987) The evolutionary dynamics of batesian and muellerian mimicry:
780 similarities and differences. pp. 81-95. Wiley Online Library.
781 Turner, J.R. & Mallet, J.L. (1996) Did forest islands drive the diversity of warningly coloured
782 butterflies? Biotic drift and the shifting balance. . *Philosophical Transactions of the*
783 *Royal Society B-Biological Sciences,* **351,** 835-845.
784 Turner, J.R.G. (1977) Butterfly mimicry - the genetical evolution of an adaptation. .
785 *Evolutionary Biology,* **10,** 163-206.
786 Twomey, E., Yeager, J., Brown, J.L., Morales, V., Cummings, M. & Summers, K. (2013)
787 Phenotypic and Genetic Divergence among Poison Frog Populations in a Mimetic
788 Radiation. *Plos One,* **8**.
789 Vorobyev, M. & Osorio, D. (1998) Receptor noise as a determinant of colour thresholds.
790 *Proceedings of the Royal Society B-Biological Sciences,* **265,** 351-358.
791 Zelditch, M.L., Swiderski, D.L., Sheets, H.D. & Fink, W.L. (2004) Geometric Morphometrics
792 for Biologists. A Primer. Elsevier Academic Press, San Diego.

793

794

Figure legends

Figure.1 Distribution and phenotype of the four species belonging to the postman mimicry ring in Northern Peru.

(a) Overview of the phylogenetic relationship between the four co-mimics (left: dorsal side, right: ventral side). Note that this phylogeny is simplified to present the four taxa of interest and the postman pattern is not shared through common ancestry. **(b)** Map of the study area with an example of phenotypic variation observed in *H. timareta* (T) between the two regions (Escalera and Alto Mayo). **(c)** Abundance of each species at different localities (AV=Aguas Verdes; Se=Serranoyacu; Af=Afluente; V=Venceremos; Sh=Shilcayo; RV=Rancho Vista; U=Urahuasha; T=Túnel; An=Antena, Table.S1) sorted by elevation, corrected by sampling effort.

Figure.2 Variation in colour pattern between and within species of the postman mimicry ring.

PCA based on colour pattern of the dorsal side and visualisation of the variations in the morphospace. **(a-b:**FW, **c-d:**HW**)**. Open dots are females and solid dots are males.

Figure.3 Variation in wing shape between and within species of the postman mimicry ring.

PCA based on wing venation **(a:**FW, **b:**HW**)** and wing outline **(c:**FW, **d:**HW**).** Variation in shape along the axis is represented with the red solid lines at the positive part of the axis and blue dotted lines at the negative value of the axis. Open dots are females and solid dots are males.

Figure.4 Colour variation between and within species of the postman mimicry ring.

PCA based on the excitation of the four photoreceptors in the physiological model of Endler & Mielke. The spot on each wing corresponds to the area where the reflectance measurement was taken **(a:**FWd, **b:**HWd, **c:**FWv, **d:**HWv**)**.

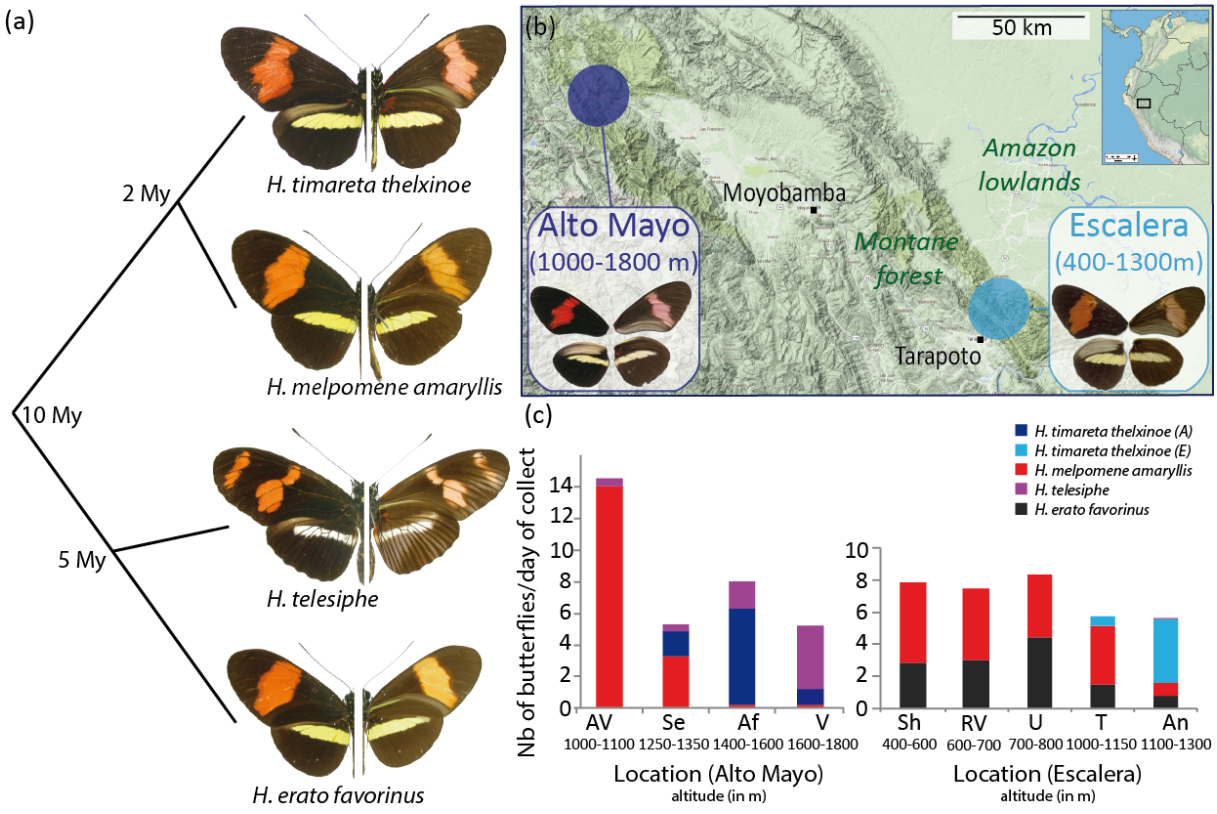
Figure.5 Pairwise phenotypic distances between groups (co-mimetic species and geographic populations of *H. timareta*)

Euclidian distances in each PCA morphospace **(a-f**). Perceptual distances in the physiological model of Vorobyev & Osario **(g-j**). Coloured boxes describe phenotypic distances between a given co-mimic (E= *H. erato,* M= *H. melpomene,* Te= *H. telesiphe*) and each population of *H. timareta* (dark blue: Alto Mayo *T(A)* - light blue: Escalera *T (E)*). Statistical differences between pairwise distances to a given mimic were tested using a mixed model with the geographic population as factor (A *vs.* E) and identity of the compared specimen as random factors (***:P≤0.001, **:P≤0.01, *:P≤0.05, ns:P≥0.05). White boxes provide distances between the remaining groups.

Table.1 MANOVA on phenotypic variation and mis-classification rate from the linear discriminant analysis between geographical populations of *H. timareta* (E *vs.* A)

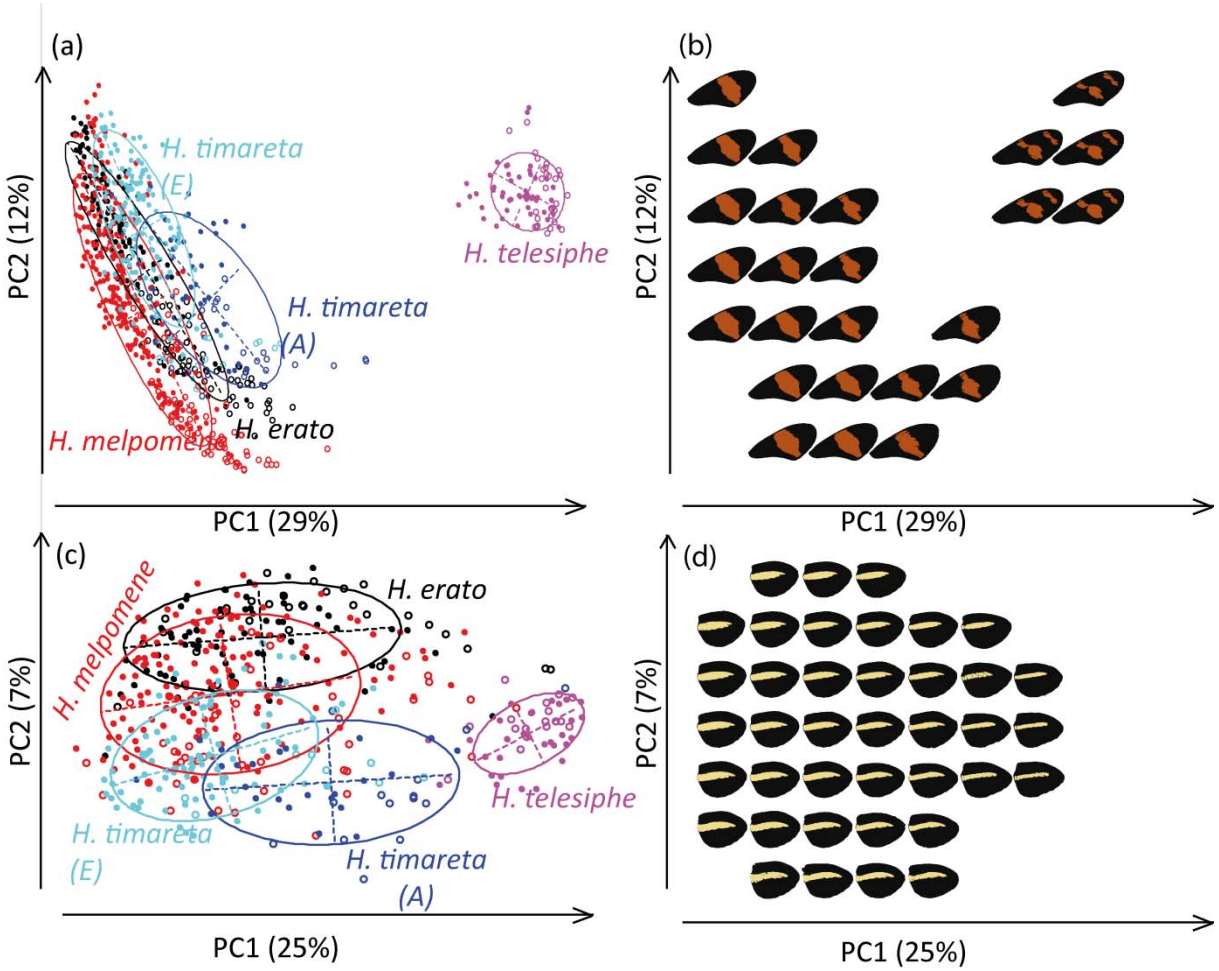
		dF	F	Pillai	P	mis-classification (LDA)
Colour spectra	dorsal red FW	1,132	13.5	0.24	<0.001	28%
	ventral red FW	1,133	8.9	0.17	<0.001	39%
	dorsal yellow HW	1,134	8.7	0.17	<0.001	29%
	ventral yellow HW	1,135	7.9	0.15	<0.001	34%
Colour pattern	dorsal red patch FW	1,136	16.5	0.66	<0.001	13%
	dorsal yellow bar HW	1,137	18.9	0.70	<0.001	11%
Wing venation	FW	1,138	7.1	0.68	<0.001	30%
	HW	1,139	7.0	0.69	<0.001	27%
Wing outline	FW	1,133	5.7	0.29	<0.001	23%
	HW	1,120	5.3	0.30	<0.001	21%

822 **Figure 1**



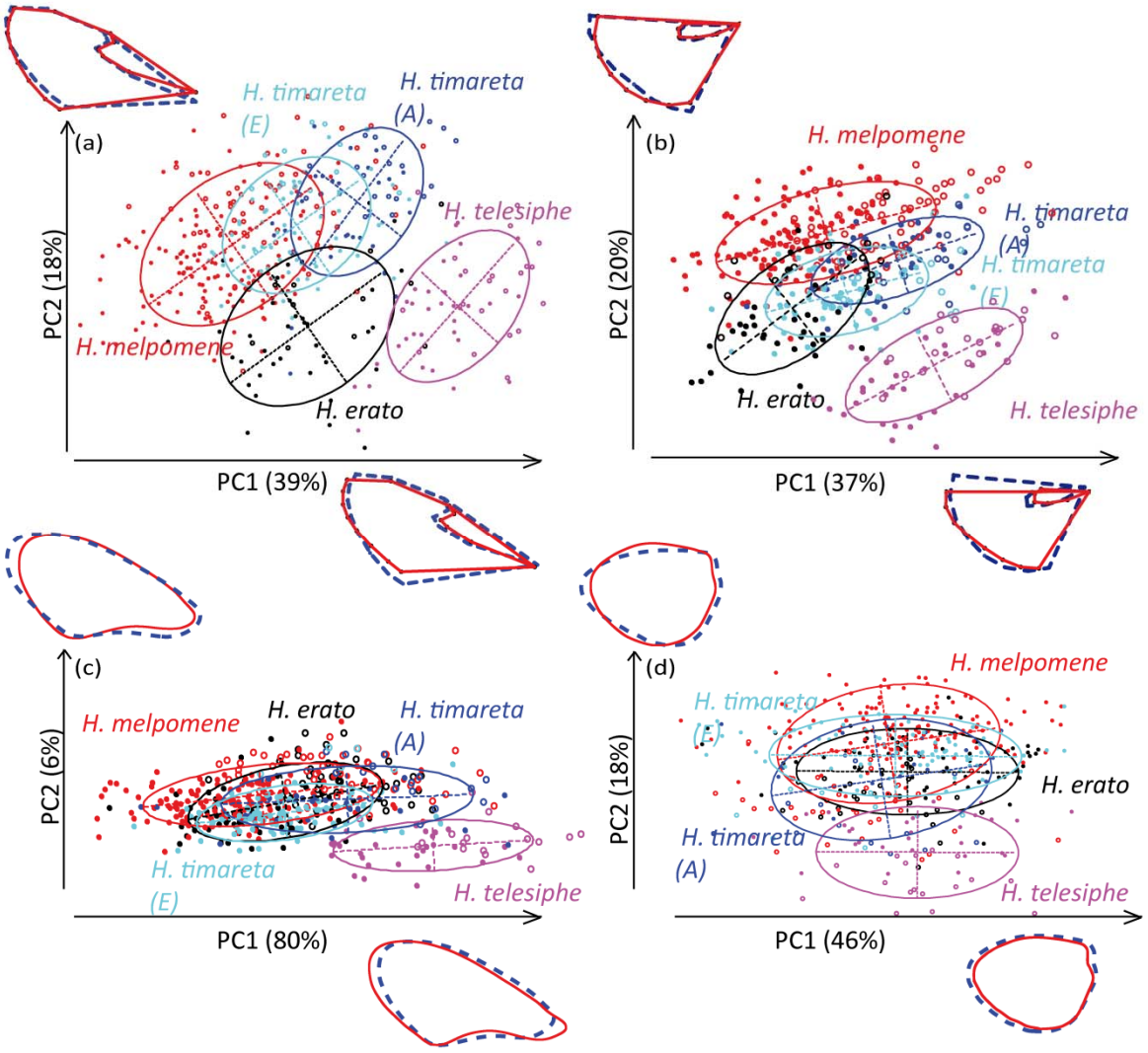
823

824 **Figure 2**



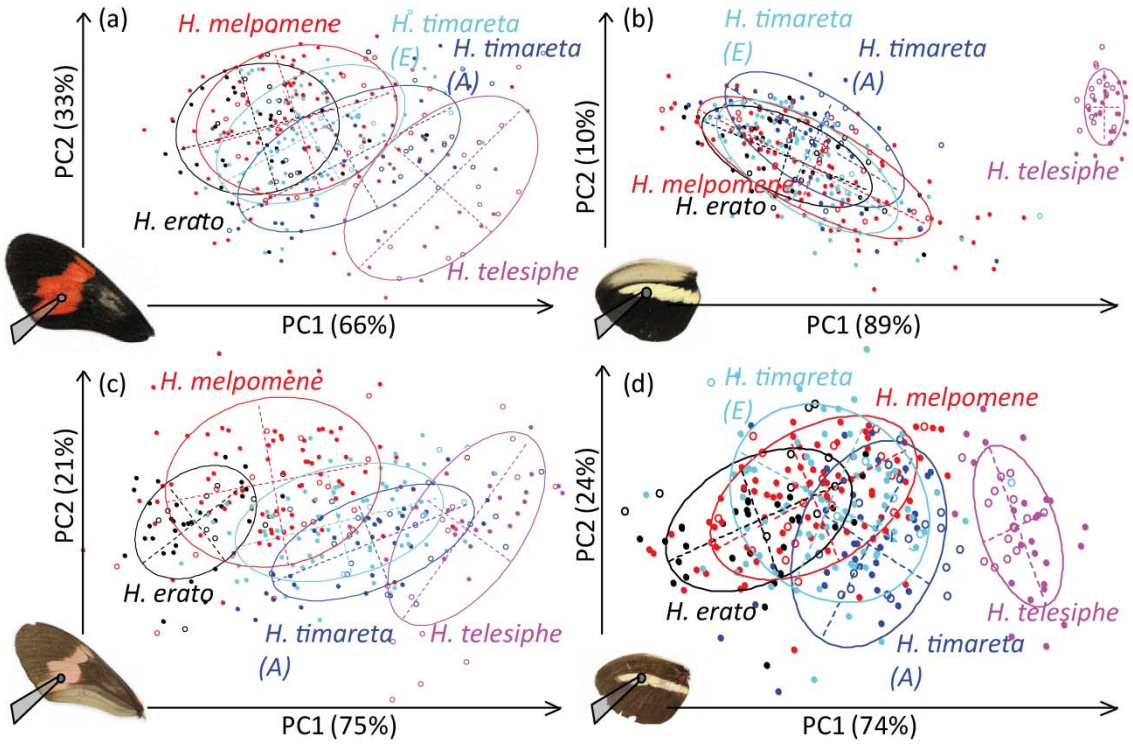
825

826 **Figure 3**



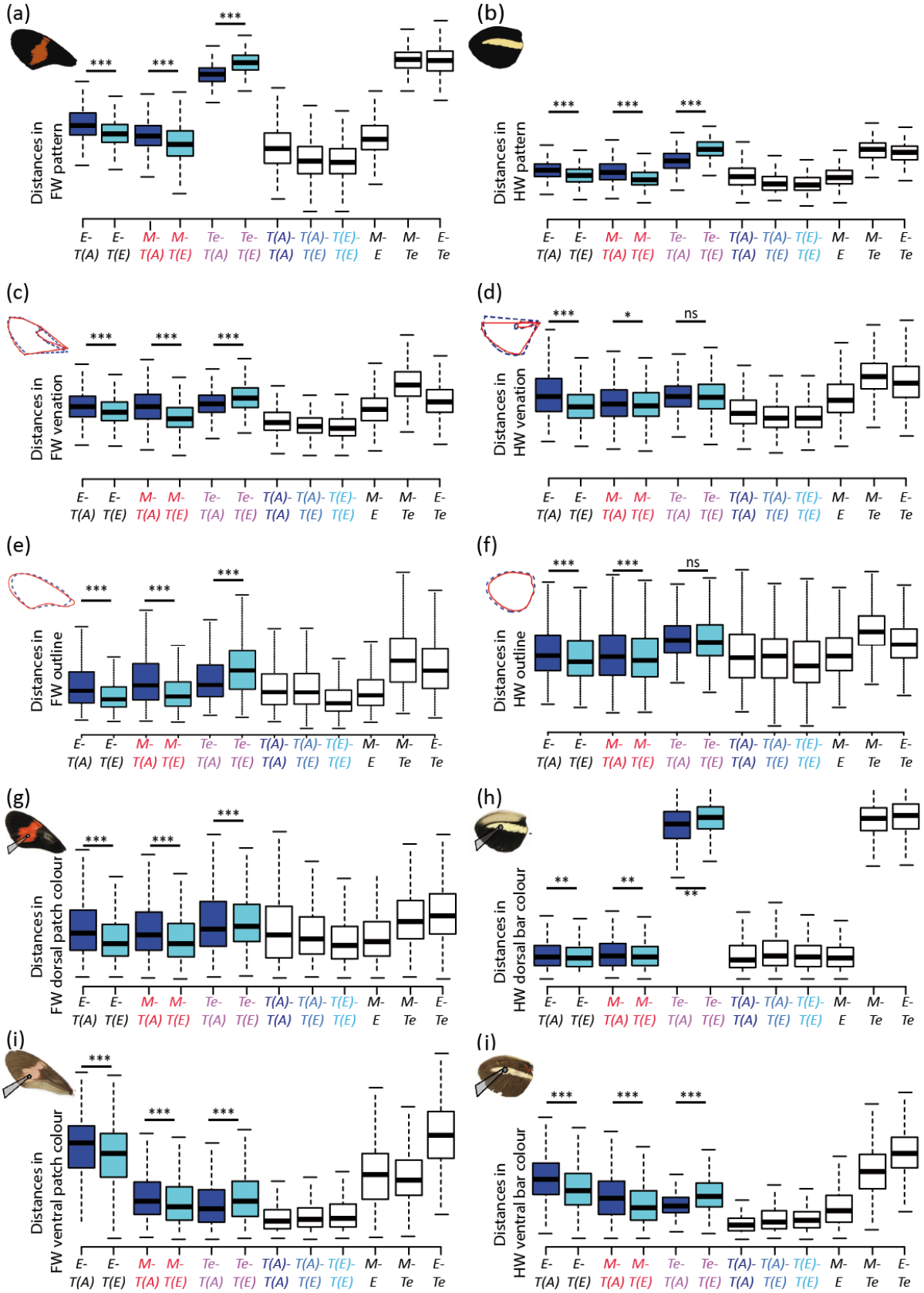
827

828 **Figure 4**



829

Figure 5


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
Distances in FW pattern
Distances in HW pattern
Distances in FW venation
Distances in HW venation
Distances in FW outline
Distances in HW outline
Distances in FW dorsal patch colour
Distances in HW dorsal bar colour
Distances in FW ventral patch colour
Distances in HW ventral bar colour
***
**
*
ns
E- T(A)
E- T(E)
M- T(A)
M- T(E)
Te- T(A)
Te- T(E)
T(A)- T(A)
T(A)- T(E)
T(E)- T(E)
M- E
M- Te
E- Te

Mimicry refinement: Phenotypic variation tracking the local optimum

Mérot C.[1*], Le Poul Y.[1], Théry M.[2], Joron M.[1,3*]

Supplementary materials

Multilocus microsatellite analysis (Mérot et al*, 2013)*

Multilocus genotypes were derived by examining variation at eleven microsatellite loci developed for *Heliconius*, using primers and PCR conditions adapted from (Flanagan *et al.*, 2002) and (Mavarez & Gonzalez, 2006). A preliminary set of specimens of 59 *H. m. amaryllis* and 28 *H. t. thelxinoe* was analysed using GeneMapper with the Genescan Rox-500 size standard for allele size determination (Applied Biosystems). A secondary set of specimens of 68 *H. m. amaryllis* and 128 *H. t. thelxinoe*, including one putative hybrid and four reference individuals per species from the first set, was analysed using GeneMarker 2.2.0 with the Genescan-500Liz size standard. Linkage disequilibrium and departure from Hardy-Weinberg within each population within each species were tested using exact tests implemented in GENEPOP 4.1.4 (Rousset, 2008). We used FSTAT 2.9.3 (Goudet, 2001) to survey within-species genetic diversity in terms of expected heterozygosity (H_E), observed heterezygosity (H_O) and allelic richness (A), estimated on the smallest sample size per locus per population (N=22). Allelic frequency and F-statistics (Weir & Cockerham, 1984) were calculated using GENETIX 4.05 (Belkhir *et al.*, 1996-2004).

Colour pattern modelling (Le Poul et al*, 2014)*

Normalized photographs. First, for each specimen, both sides of each forewing and hindwing were photographed in normalized light conditions (CIE Standard Illuminant D50), with a high colour rendering light source (Philips Master TL-D 90 Graphica pro). A scale indicator was included in each picture, and the white balance was normalized. A Nikon D90 digital camera with a Nikon micro 105/2.8G ED VR lens was used to capture high resolution images with accurate colour rendering.

Wing extraction. The first step of CPM relies in identifying and extracting outline of wings on pictures. Wings were automatically detected in the images using their colour difference with the homogenous white background, and were then precisely extracted using the marker-based watershed transformation (Meyer & Beucher, 1990) along the image colour gradient (Meyer, 1992). This segmentation method finds the maximum intensity of the colour transition between the marked wings and the marked background, which was then considered to be the wing outline.

Colour number reduction and colour attribution. Then, the pattern was modelled by considering explicitly the mosaic distribution of colour across the wings, which allows describing efficiently the variation in patches boundaries. A set of discrete colours characteristic to each wing was first identified using an algorithm based on colour histograms (Cheng & Sun, 2000; Kurugollu *et al*, 2001). After simplifying the spatial structures of wing images (Nikolaev & Nicolayev, 2004), pixels were then attributed to each colour (black, red and yellow) using a simple threshold.

Colours classes were extracted from the histogram by the following procedure. First, to smooth the colour distribution on the histogram and simplify colour histogram processing, the image was projected from the 3-dimensional RGB colour space to a 2-dimensional (2D) colour space, where

dimension 1 corresponds to luminance (the Y component in the Ycrbr colour space), and dimension 2 to the major colour variation axis using all wing images. This projection preserved about 97% of colour variance in the images. A 2D-histogram, representing the distribution of pixel colour, was computed in the same 2D-projected colour space for all wings. Each separate colour on the wing was defined as a local maximum on the 2D-histogram. These local maxima are always numerous because of the complexity of the natural image. To prevent over-segmentation, minor peaks were automatically removed by consideration of their proximity to and separation from neighbouring peaks. We performed a watershed transformation on the additive inverse of the 2D-histogram to partition the colour space among the major peaks (Shafarenko *et al,* 1998). At the end of this process, each wing could be associated with a set of characteristic colour partitioning the colour space and accounting for the colour variation actually present on the wing RGB image.

In order to preserve the patch structure of colour patterns, we also performed a routine to merge neighbouring pixels of homogenously-coloured regions in the images (Nikolaev & Nicolayev, 2004). The scale indicator within each image was first used to rescale images to an output length of around 512 pixels, leading to a mean spatial resolution of 10.9px/mm. Each reduced image was then transformed into a mosaic of homogeneously-coloured spatial zones. A watershed transformation of the image colour gradient was used to carry out the mosaicking (Meyer, 1992). Each homogenously-coloured region was then attributed a colour according to the classification given by the segmented histogram.

Finally, the attribution to the different colours (black, red, and yellow) was done automatically using a threshold on RGB values, followed by manual checking to correct errors, which were mostly due to minor damage to parts of the wings, resulting in the final segmented image.

Alignment. For the wing images to be efficiently comparable pixel by pixel, a proper homology of pixel positions was needed. This match was obtain by transforming each set of processed images into a common coordinate system which maximize similarity between each wing pattern and wing outline to a wing model (*i.e.* the ‘mean’ of all individuals), treating each wing surface separately. Similarity was measured by the Mattes implementation of mutual information metric (Thévenaz & Unser, 2000; Mattes *et al,* 2001; Mattes *et al,* 2003), which is minimal when colour patches and outlines are aligned in an optimal compromise. The one+one evolutionary optimizer (Styner & Gerig, 1997; Styner *et al,* 2000), implemented within the ITK free image proceeding library in C++ (Yoo *et al,* 2002; Ibañez, 2003; Martin & Hoffman, 2003); , was used to find the scale, rotation and translation parameter set that minimized this mutual information value. This procedure created an initial registration set based on wing shape, which allowed generation of the wing pattern model. Each wing was then recursively aligned to the model, until the variance of the metric stabilized (variance varying less 1% (Rohlf & Slice, 1990). At the completion of this process, all wings could be considered to be positioned in the same physical space, with pixel locations and colour values among wings being comparable among all individuals.

Table S2. Number of specimen used in each analysis

	Density estimates	***Density*** (butterflies/day of collection)	***Scored for genotype*** *(microsatellites)*	***Scored for phenotype*** *Band/ spots*	*Venation*	*Outline (FW)*	*Outline (HW)*	*Pattern (FW v)*	*Pattern (FW d)*	*Pattern (HW v)*	*Pattern (HW d)*	*Reflectance*
Alto Mayo	**180 (25 days)**		**59**	**200**	**151**	**118**	**105**	**121**	**124**	**109**	**112**	**118**
H. timareta	83	3.3	38	71	54	42	31	42	44	34	36	43
H. melpomene	54	2.2	20	60	39	27	25	28	28	25	25	31
H. erato	0	0.0		5	5	5	5	5	5	5	5	5
H. telesiphe	43	1.7		63	53	44	44	46	47	45	46	39
H. melpomene x H. timareta ?			1									
Escalera	**473 (83 days)**		**172**	**324**	**148**	**187**	**172**	**195**	**195**	**190**	**191**	**98**
H. timareta	203	2.4	115	167	70	91	89	97	97	98	97	73
H. melpomene	175	2.1	54	118	67	61	55	64	63	62	63	16
H. erato	93	1.1		34	10	33	26	32	33	28	29	8
H. telesiphe	2	0.0		2	1	2	2	2	2	2	2	1
H. melpomene x H. timareta ?			3									
Escalera_low	**517 (65 days)**		**32**	**163**	**78**	**128**	**119**	**132**	**132**	**121**	**124**	**44**
H. melpomene	293	4.5	32	103	55	77	72	81	80	73	73	22
H. erato	224	3.4		60	23	51	47	51	52	48	51	22
Moyobamba			**11**	**54**	**37**	**47**	**44**	**50**	**48**	**45**	**45**	**22**
H. melpomene			11	35	25	30	29	32	30	30	30	16
H. erato				19	12	17	15	18	18	15	15	6
Raised												
H. melpomene x H. timareta			11	11								9
Total			**285**	**752**	**414**	**480**	**440**	**498**	**499**	**465**	**472**	**291**

Table S3: Genetic polymorphism of the studied sample.

Allelic richness (A) is estimated for the smallest population (N=22). H_O represents the observed heterozygosity and H_E, the expected heterozygosity. Significant deviations from Hardy-Weinberg expectations are indicated by asterisks (P < 0.05 *; P < 0.01 **). Loci come from Flanagan et al (2002)[1], and Mavarez et al (2006)[2].

			H. timareta thelxinoe (Alto Mayo)							*H. timareta thelxinoe* (Escalera)						
Locus	Primer sequence (5'-3')	GenBank accession no.	No. of sample	No. of alleles	Allelic Richness	Size range (bp)	Ho	HW DEFICIT	He	No. of sample	No. of alleles	Allelic Richness	Size range (bp)	Ho	HW DEFICIT	He
Hel2 (1)	TCAAAATGTTGCAGACCGAG TGCACTTCATTGTAAGGCGT	AF481467	40	4	3.6	186-202	0.20	**	0.52	113	5	4.3	196-202	0.50	**	0.58
Hel4 (1)	CGTTGCCGCTTATACTTTCC GGAACGGAGTGCCCTAAAAC	AF481469	40	7	6	268-282	0.34	**	0.62	109	10	6.8	262-302	0.37	**	0.77
Hel5 (1)	TGCTGTCCATACCCAACTCA CGAACTCACAACCATCAGTCA	AF81470	40	3	2.6	314-330	0.24		0.32	115	5	2.9	314-330	0.21	**	0.22
Hm02 (2)	TATTTGCACGATGGAAACCC GCGAGGTGGAGACAAAAGAC	DQ020073	40	6	5.5	180-200	0.54	**	0.75	112	11	7.2	180-200	0.71	**	0.72
Hm03 (2)	GACGTCACAGCGGGGAAC AGAGGGGAACGGAGTGTCAT	DQ020074	40	9	7.5	292-352	0.49	*	0.65	110	10	6.8	312-348	0.59	*	0.72
Hm04 (2)	CCTGGCTTATCTACGACGACA ATGCAGCTTACTCGCTGGTT	DQ020075	40	10	8	410-458	0.66	**	0.79	111	7	5	406-452	0.49		0.60
Hm05 (2)	GCGGTAAGGTAAAACCGTGA CAGAAGAAAATGGTTGGATGG	DQ020076	40	8	7.2	284-302	0.76		0.71	112	12	8.3	274-300	0.71	**	0.82
Hm06 (2)	AAATAGTGTGCGGCGGAATA TGGAGTAGAAATGCGGGTTTA	DQ020077	40	4	4	240-246	0.49	**	0.73	111	5	4.9	238-246	0.59	**	0.74
Hm13 (2)	TCACTAGTTTTCGGCTTATCG AAGGCTAAATGATGCCTAAAG	DQ020083	40	6	5.9	168-188	0.78		0.82	112	9	6.5	168-190	0.65	**	0.77
Hm19 (2)	CGCTAATTCAAAGGAAAGAGGA AGTGCTGTCATGGCTAACGA	DQ020088	40	5	4	182-216	0.22		0.25	110	9	5	182-216	0.14	**	0.32
Hm22 (2)	CCTCGTCCAAACTCCAAAAC AACAATGTCACAACCATCGC	DQ020092	40	3	2.8	252-256	0.51		0.53	113	6	3.5	252-264	0.54		0.53

			H. melpomene amaryllis (Alto Mayo)							*H. melpomene amaryllis* (Escalera)						
Locus	Primer sequence (5'-3')	GenBank accession no.	No. of sample	No. of alleles	Allelic Richness	Size range (bp)	Ho	HW DEFICIT	He	No. of sample	No. of alleles	Allelic Richness	Size range (bp)	Ho	HW DEFICIT	He
Hel2 (1)	TCAAAATGTTGCAGACCGAG TGCACTTCATTGTAAGGCGT	AF481467	22	5	5	184-192	0.55	**	0.65	53	6	5	184-192	0.54	*	0.63
Hel4 (1)	CGTTGCCGCTTATACTTTCC GGAACGGAGTGCCCTAAAAC	AF481469	22	12	12	236-312	0.46	**	0.88	50	22	14.7	236-322	0.29	**	0.90
Hel5 (1)	TGCTGTCCATACCCAACTCA CGAACTCACAACCATCAGTCA	AF81470	22	11	11	310-358	0.82		0.85	53	21	14.9	296-358	0.77	*	0.90
Hm02 (2)	TATTTGCACGATGGAAACCC GCGAGGTGGAGACAAAAGAC	DQ020073	22	7	7	184-200	0.46	**	0.64	53	10	7.3	170-198	0.64	**	0.69
Hm03 (2)	GACGTCACAGCGGGGAAC AGAGGGGAACGGAGTGTCAT	DQ020074	22	7	7	314-352	0.46		0.55	53	111	8.2	302-348	0.56		0.67
Hm04 (2)	CCTGGCTTATCTACGACGACA ATGCAGCTTACTCGCTGGTT	DQ020075	22	11	11	402-444	0.64	*	0.78	53	14	10.9	396-428	0.71	**	0.80
Hm05 (2)	GCGGTAAGGTAAAACCGTGA CAGAAGAAAATGGTTGGATGG	DQ020076	22	13	13	274-320	0.77		0.85	52	14	11.5	274-306	0.75	*	0.87
Hm06 (2)	AAATAGTGTGCGGCGGAATA TGGAGTAGAAATGCGGGTTTA	DQ020077	22	4	4	240-246	0.77		0.76	53	6	5.5	238-248	0.64		0.73
Hm13 (2)	TCACTAGTTTTCGGCTTATCG AAGGCTAAATGATGCCTAAAG	DQ020083	22	12	12	144-186	0.73	*	0.88	53	13	10.2	170-198	0.56	**	0.76
Hm19 (2)	CGCTAATTCAAAGGAAAGAGGA AGTGCTGTCATGGCTAACGA	DQ020088	22	13	13	154-208	0.64	**	0.85	53	20	14.8	150-210	0.52	**	0.88
Hm22 (2)	CCTCGTCCAAACTCCAAAAC AACAATGTCACAACCATCGC	DQ020092	22	13	13	246-279	0.68	**	0.91	52	15	11.8	246-280	0.67	**	0.85

Geographic variation in H. melpomene *and* H. erato

Table S4. MANOVA on phenotypic variation and mis-classification rate from the linear discriminant analysis between four geographical populations of *H. melpomene* ("Low Escalera", "High Escalera", "Alto Mayo", "Moyobamba")

		dF	F	Pillai	P	mis-classification (CVA)
Colour spectra	dorsal red FW	3,95	2.2	0.20	0.02	65%
	ventral red FW	3,95	1.9	0.17	0.055	64%
	dorsal yellow HW	3,95	0.6	0.06	0.77	70%
	ventral yellow HW	3,95	1.4	0.13	0.19	71%
Colour pattern	dorsal red patch FW	3,194	1.7	0.36	0.004	58%
	dorsal yellow bar HW	3,198	2.6	0.52	<0.001	52%
Wing venation	FW	3,178	1.8	0.43	0.001	60%
	HW	3,178	2.7	0.59	<0.001	53%
Wing outline	FW	3,193	1.2	0.16	0.27	59%
	HW	3,177	1.8	0.26	0.009	59%

Table S5. MANOVA on phenotypic variation and mis-classification rate from the linear discriminant analysis between three geographical populations of *H. erato* ("Low Escalera", "High Escalera", "Moyobamba")

		dF	F	Pillai	P	mis-classification (CVA)
Colour spectra	dorsal red FW	2,34	1.2	0.20	0.33	43%
	ventral red FW	2,34	1.2	0.20	0.32	43%
	dorsal yellow HW	2,34	0.5	0.09	0.80	41%
	ventral yellow HW	2,34	0.3	0.06	0.91	41%
Colour pattern	dorsal red patch FW	2,98	1.3	0.38	0.14	52%
	dorsal yellow bar HW	2,96	1.1	0.33	0.36	54%
Wing venation	FW	2,34	1.4	0.98	0.18	57%
	HW	2,34	1.7	1.09	0.06	60%
Wing outline	FW	2,99	1.0	0.17	0.52	48%
	HW	2,87	1.1	0.22	0.36	46%

Colour pattern analyses

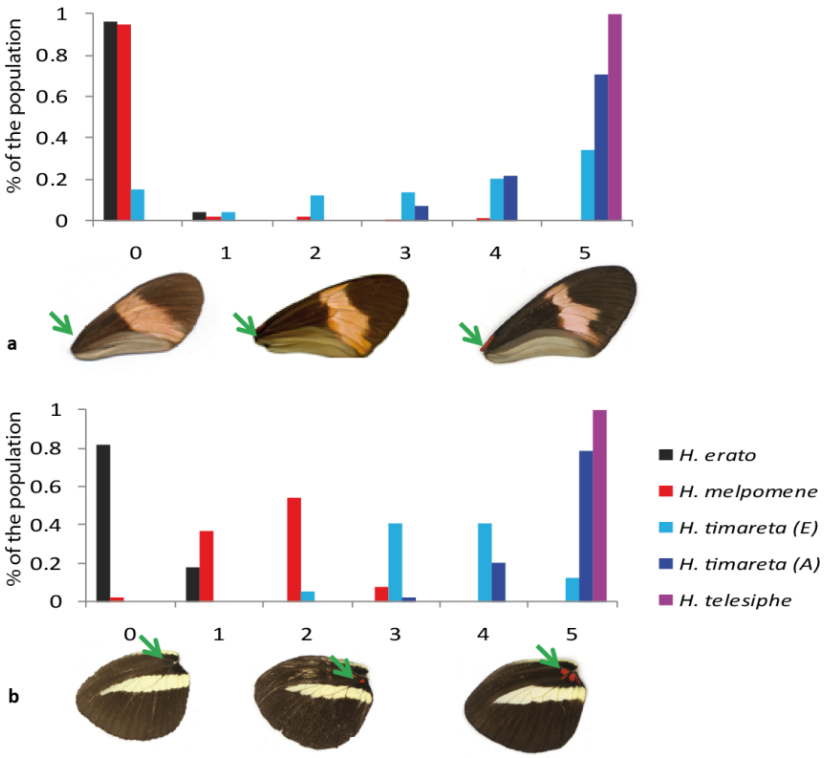


Figure S1: Distribution of the indices scoring the small ventral elements of pattern between and within species of the postman mimicry ring. (a) Presence and size of a red line at the base of the costal vein of the ventral forewing. The index ranges from 0 (absence of red line) to 5 (large and well-marked red line). (b) Presence and size of red dots at the base of the ventral hindwing. The index ranges from 0 (no dot) to 5 (presence of four large smudgy spots). 1 to 4 correspond to a gradation between one or two minute dots and more numerous or larger spots.

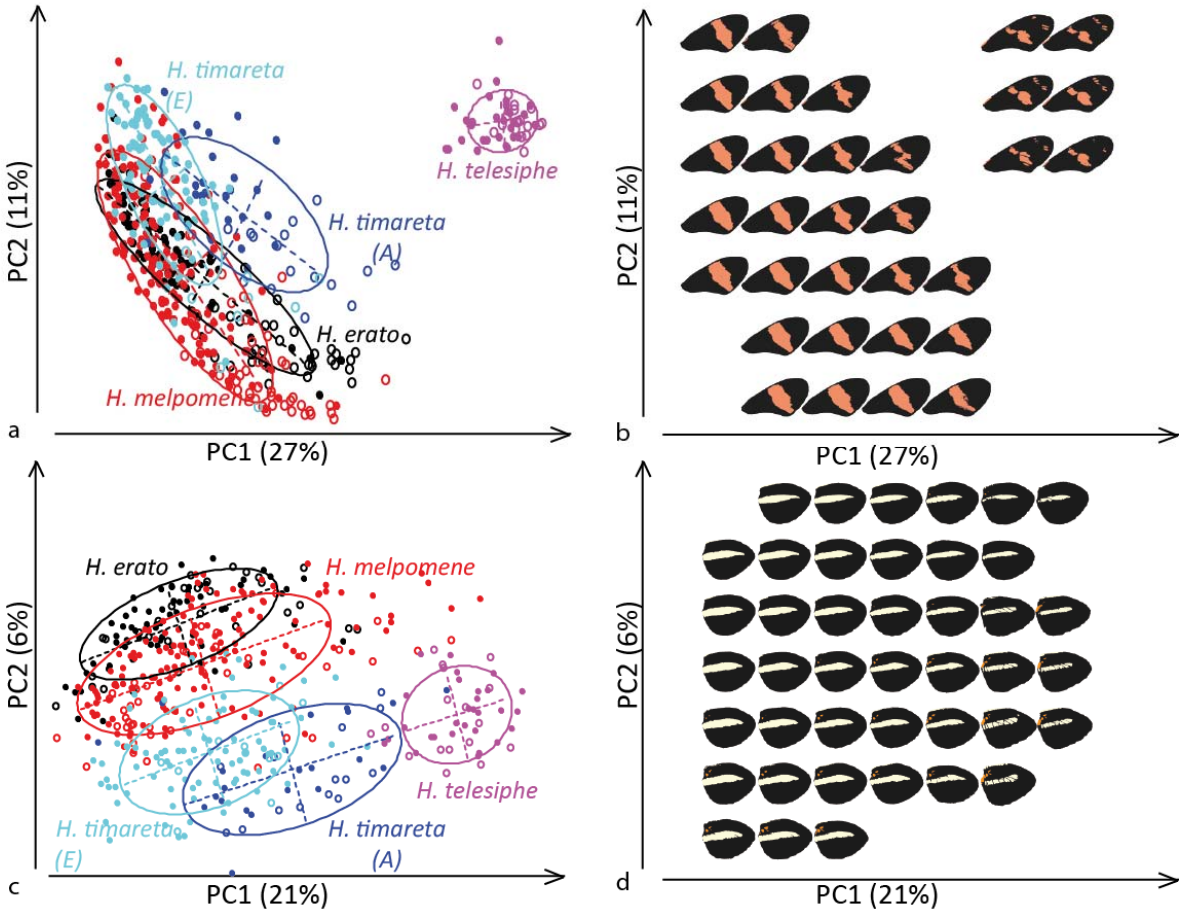


Figure S2: Colour pattern variation between and within species of the postman mimicry ring. Principal component analysis based on colour pattern and visualisation of variation in the morphospace (ventral FW (a-b), ventral HW (c-d))

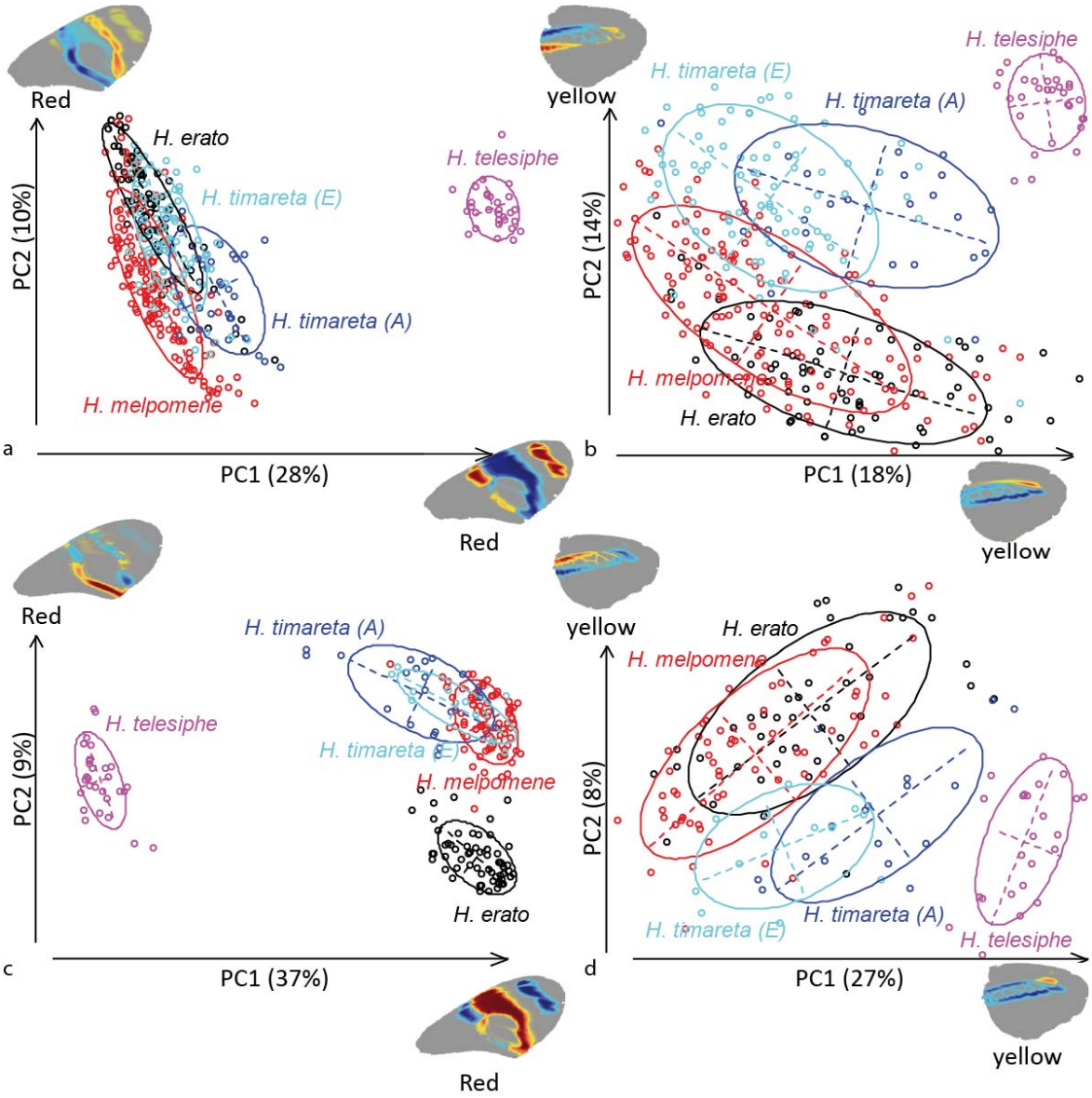


Figure S3: Wing colour pattern PCA within each sex.

Wing sketches show the pixels involved in the variation of colour pattern associated with each axis. A positive contribution (red) for one colour indicates that an increase in the score is associated with the appearance (presence) of this colour on the phenotype. On the contrary, a negative contribution (blue) is associated with a disappearance of this colour. (a) Red colour on male FW, (b) yellow colour on male HW, (c) red colour on female FW, (d) yellow colour on female HW.

Wing venation

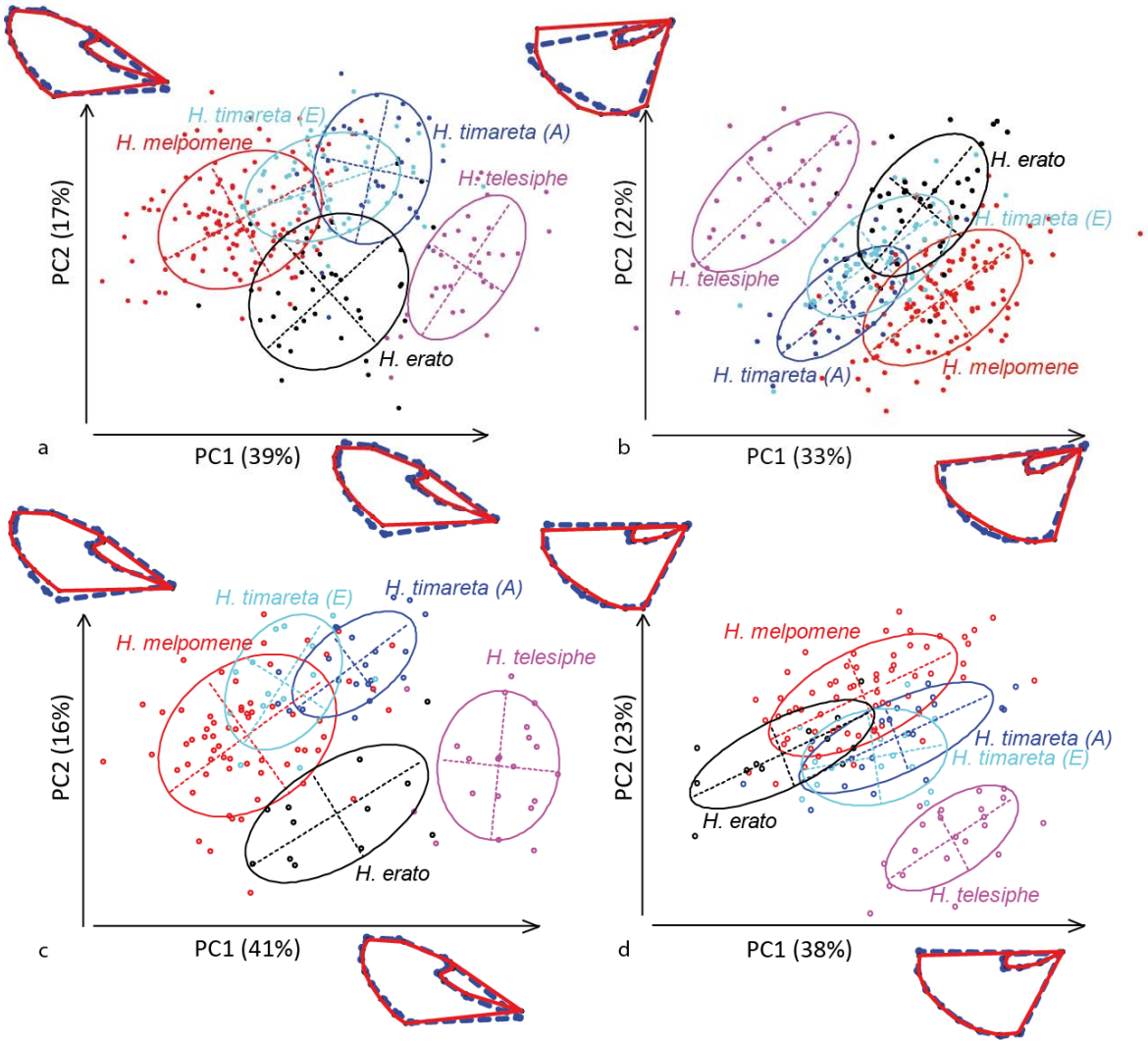


Figure S4: PCA on wing venation for each sex.
(a) FW for males, (b) HW for males, (c) FW for females, (d) HW for females.

Colour analyses

The best-documented predators of *Heliconius* include jacamars (Chai & Srygley 1990; Langham 2006), which belong to the order Galbuliformes. Within this order, only *Nystalus maculates* visual system has been studied and would have a tetrachromatic violet-type visual system according to opsin sequence similarity (Odeen & Hastad 2003). However, the guild of bird predators on those butterflies may include other birds, which possibly differ in the spectral sensitivities of their short-wavelength-sensitive cone visual pigments. Therefore, we included two tetrachromatic visual systems in our analysis: a bird with a V-type visual system (Peafowl, *Pavo cristatus*,(Hart 2002)), a bird with a UV-type visual system (Blue tit, *Parus caeruleus*,(Hart *et al.* 2000)). Because the results were not qualitatively different when analysed with different visual systems, we chose to present the results in the V-type vision model. Results in the UV-type visual system displayed consistent results with the V-type

system (Fig. S9). All models were run considering two kinds of incident light: a light environment corresponding to large sunny gaps in a tropical forest, and a light environment corresponding to small light gaps. Those light spectra were measured by M. Théry in primary forest in French Guiana (Théry, Pincebourde & Feer 2008). Results were consistent with both types of light so we only present models with large gap incident light here, which corresponds better to the natural habitat of the species considered here.

Surrounding colours can affect perception and it is usually important to take background colour into account. Here, all patches are large making the colour well-visible. The patches are surrounded by the black part of the wing, for which we analyzed reflectance. Colour contrasts calculated with the model of Vorobyev and Osorio (1998) indicate that differences between and within species were not perceptible for observers (below the usual value of 1 JND), with mean values around 0.5 JND in the large gap conditions and 0.2 JND in the small gap conditions (Fig. S8) Therefore, we focused on the comparison of actual patches' colour between species and populations.

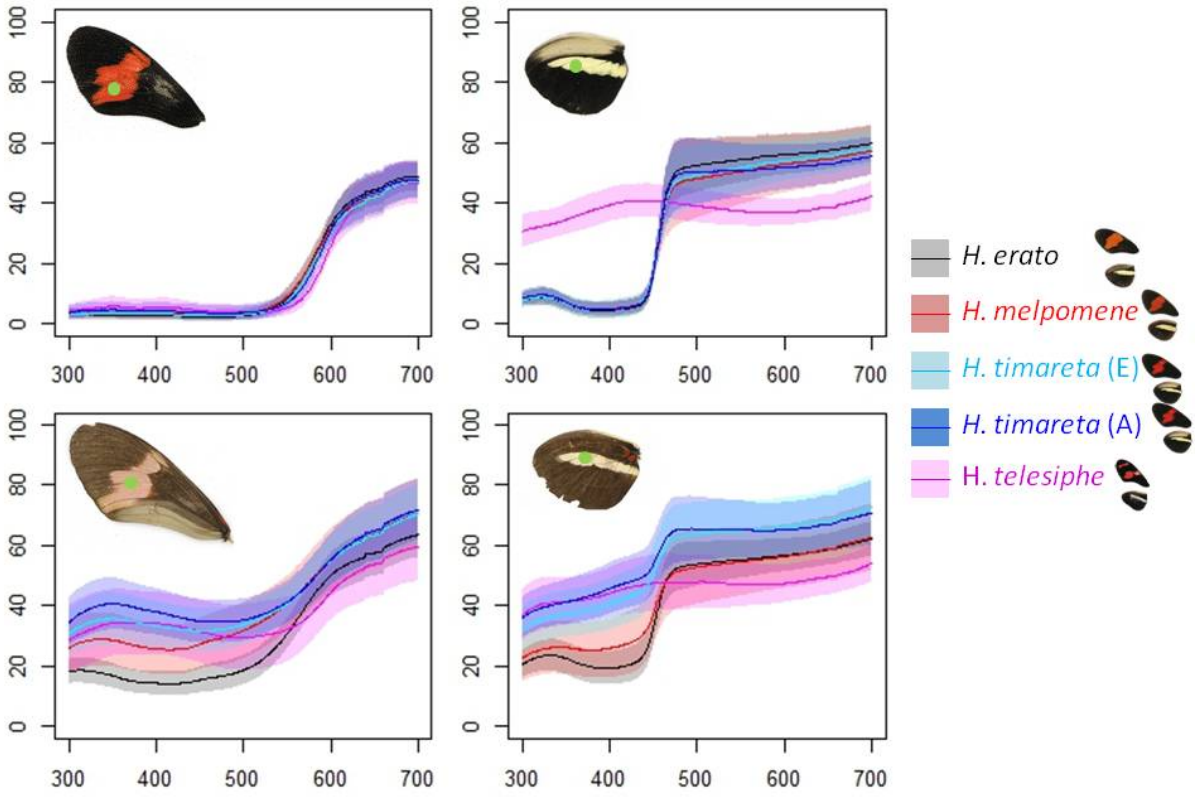


Figure S5: Mean raw spectra of colour patches for each taxon. The x-axis represents the wavelength from 300 to 700 nm and the y-axis the reflectance in percentage of the white control. The shadows represent standard deviation.

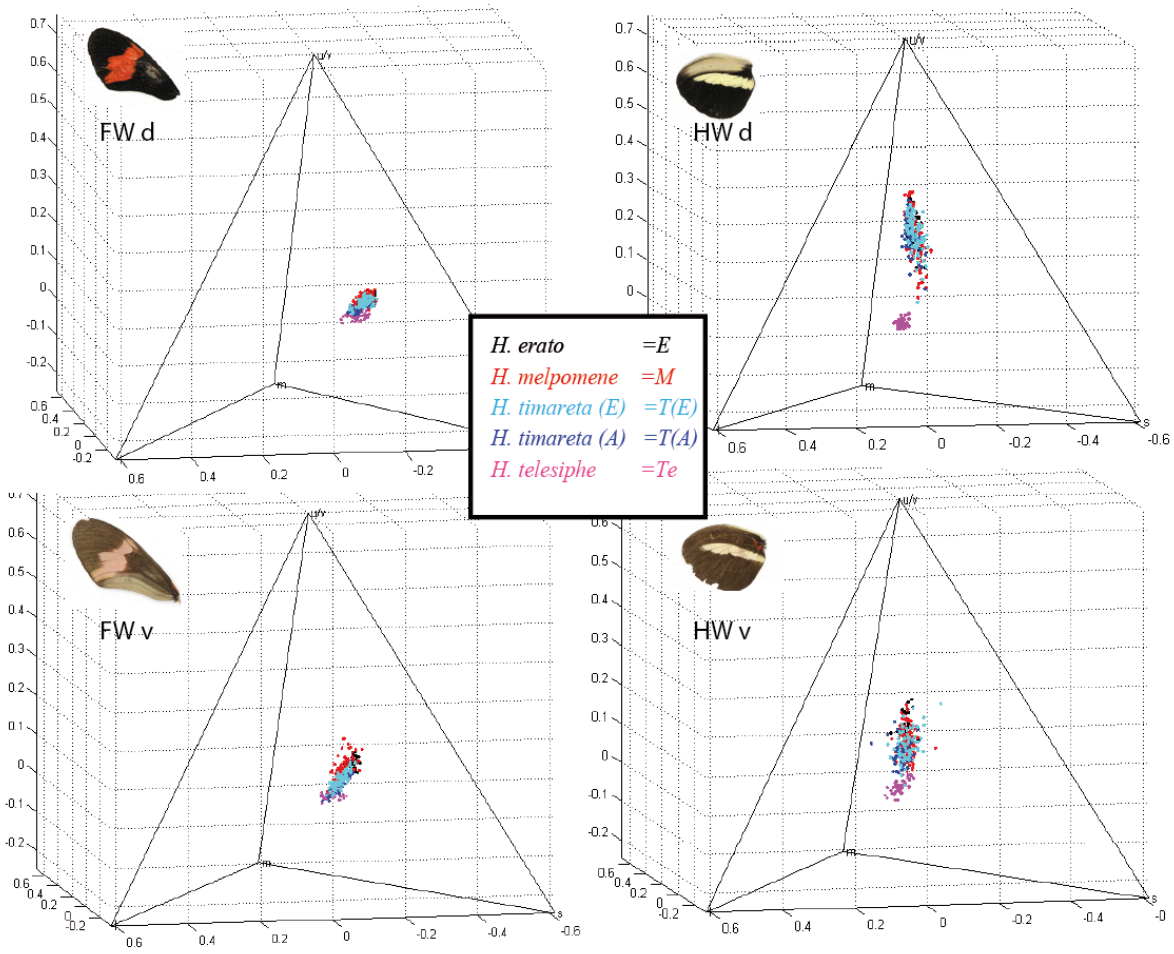


Figure S6: Location of wing colours in the tetrahedral chromatic visual space.
Based on a model Endler and Mielke (2005) with a bird V-type visual system (Peafowl) and a large gap light.

Table S6. Percentage of overlap in the tetrahedral colour space between spectral variation of each species and each population of *H. timareta*.
The last column gives the overlap between the two populations of *H. timareta*. This is based on a physiological model Endler and Mielke (2005) with a bird V-type system (Peafowl) and a large gap light.

	H. erato		*H. melpomene*		*H. telesiphe*		*H. timareta*
	H. timareta A	*H. timareta E*	*H. timareta A*	*H. timareta E*	*H. timareta A*	*H. timareta E*	*A vs. E*
Red FW d	0.22	0.36	0.43	0.58	0.09	0.05	0.58
Red FW v	0.00	0.08	0.08	0.22	0.20	0.11	0.39
Yellow HW d	0.31	0.32	0.25	0.29	0.00	0.00	0.42
Yellow HW v	0.09	0.12	0.26	0.43	0.00	0.00	0.34

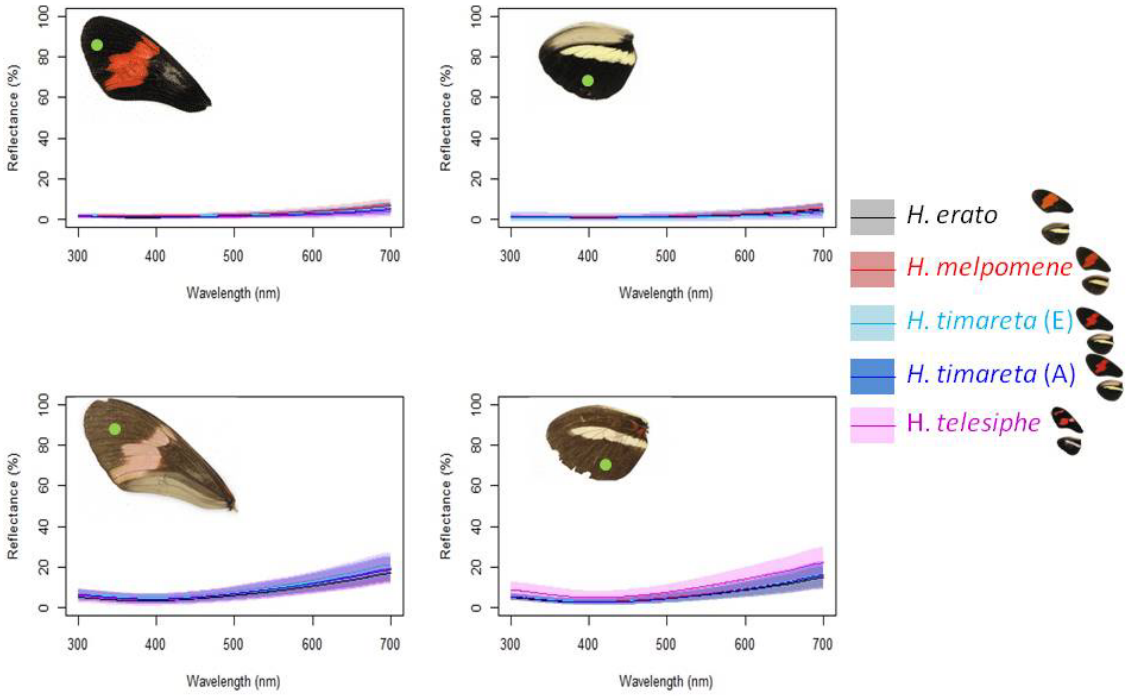


Figure S7: Mean raw spectra of the black area for each taxon. The x-axis represents the wavelength from 300 to 700 nm and the y-axis the reflectance in percentage of the white control. The shadows represent standard deviation.

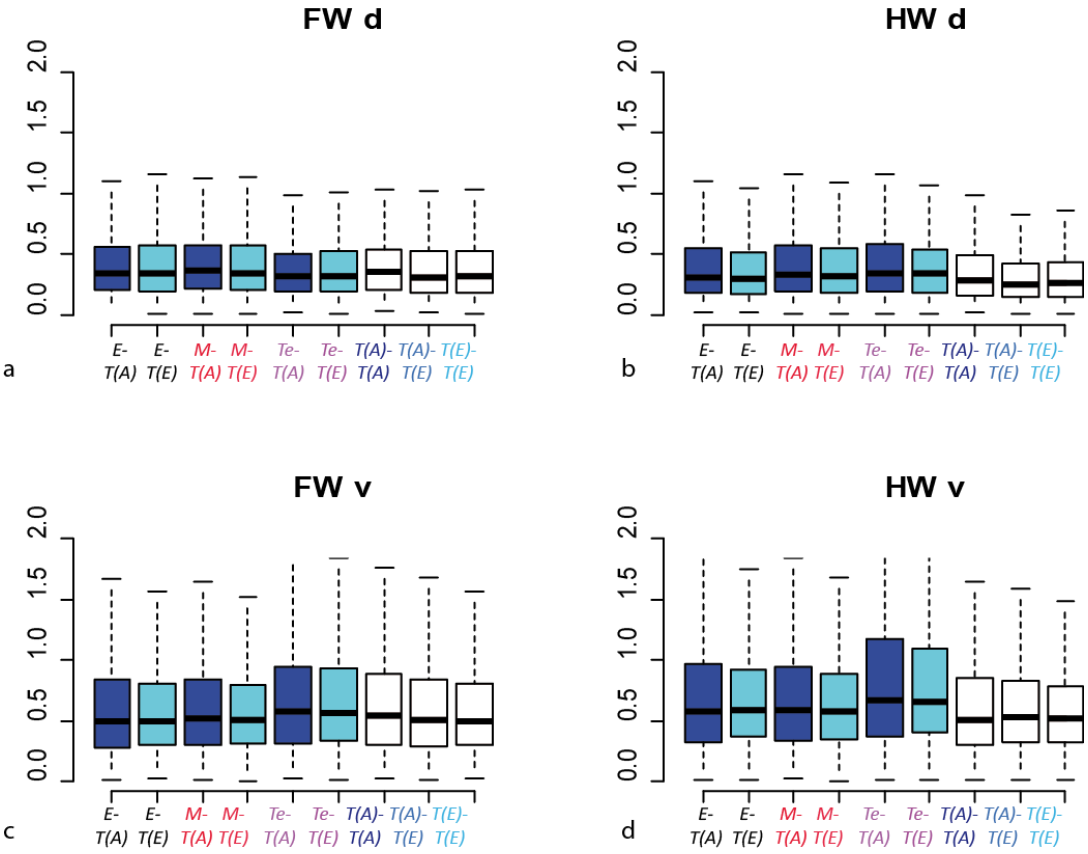


Figure S8: Controlling perceptual distances in the black.

Colour perceptual distances for the black part of each wing in the model of Vorobyev & Osorio (1998) expressed in unit of just noticeable differences (JND). Bird V-type system (Peafowl) and a large gap light. Contrast between species blackness is below 1 JND.

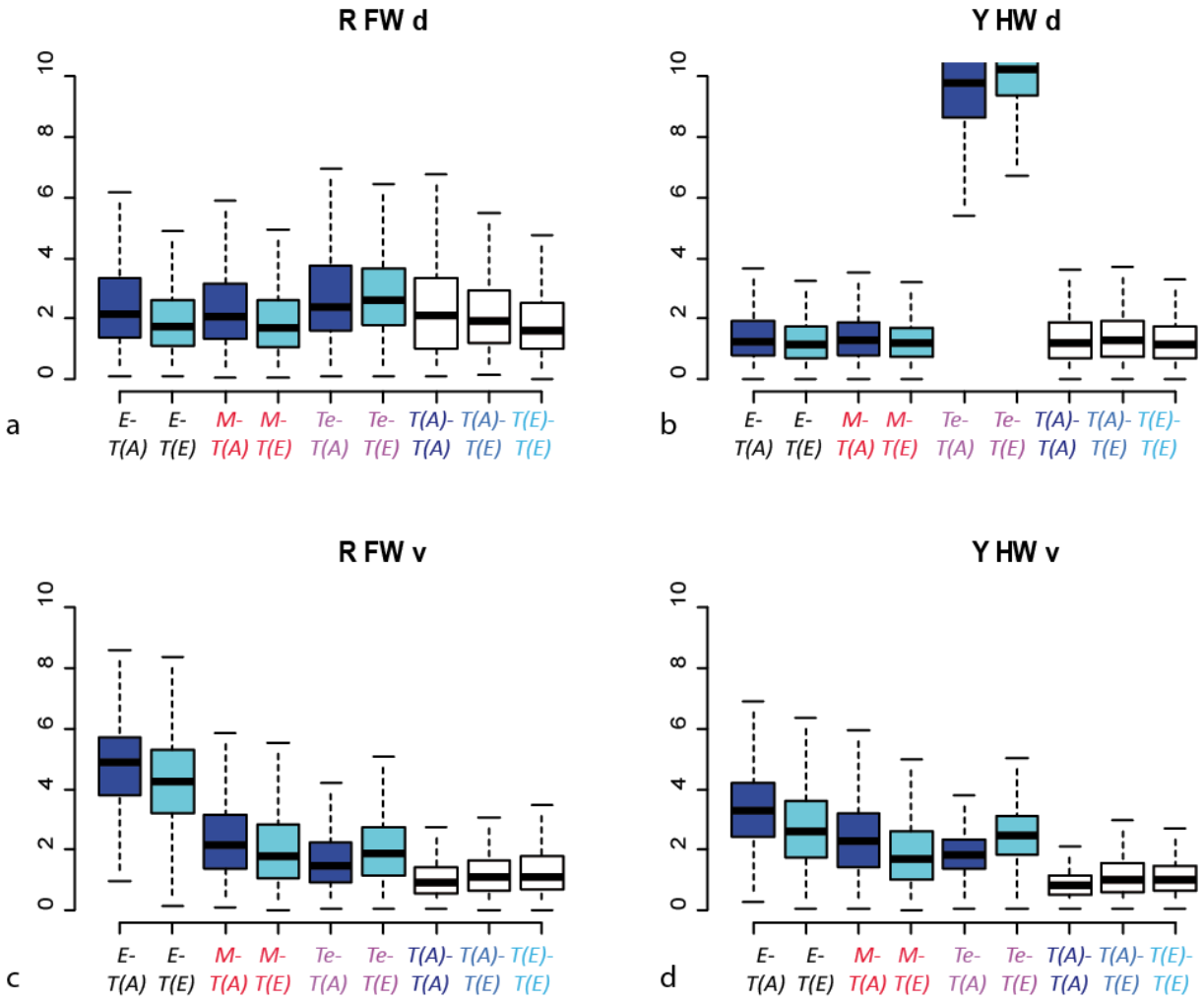


Figure S9: Investigating other vision system (UV-type).
Colour perceptual distances in the model of Vorobyev & Osorio (1998) with a bird UV-type visual system (Blue tit) and a large gap light for the red patches (R) and the yellow patches (Y)

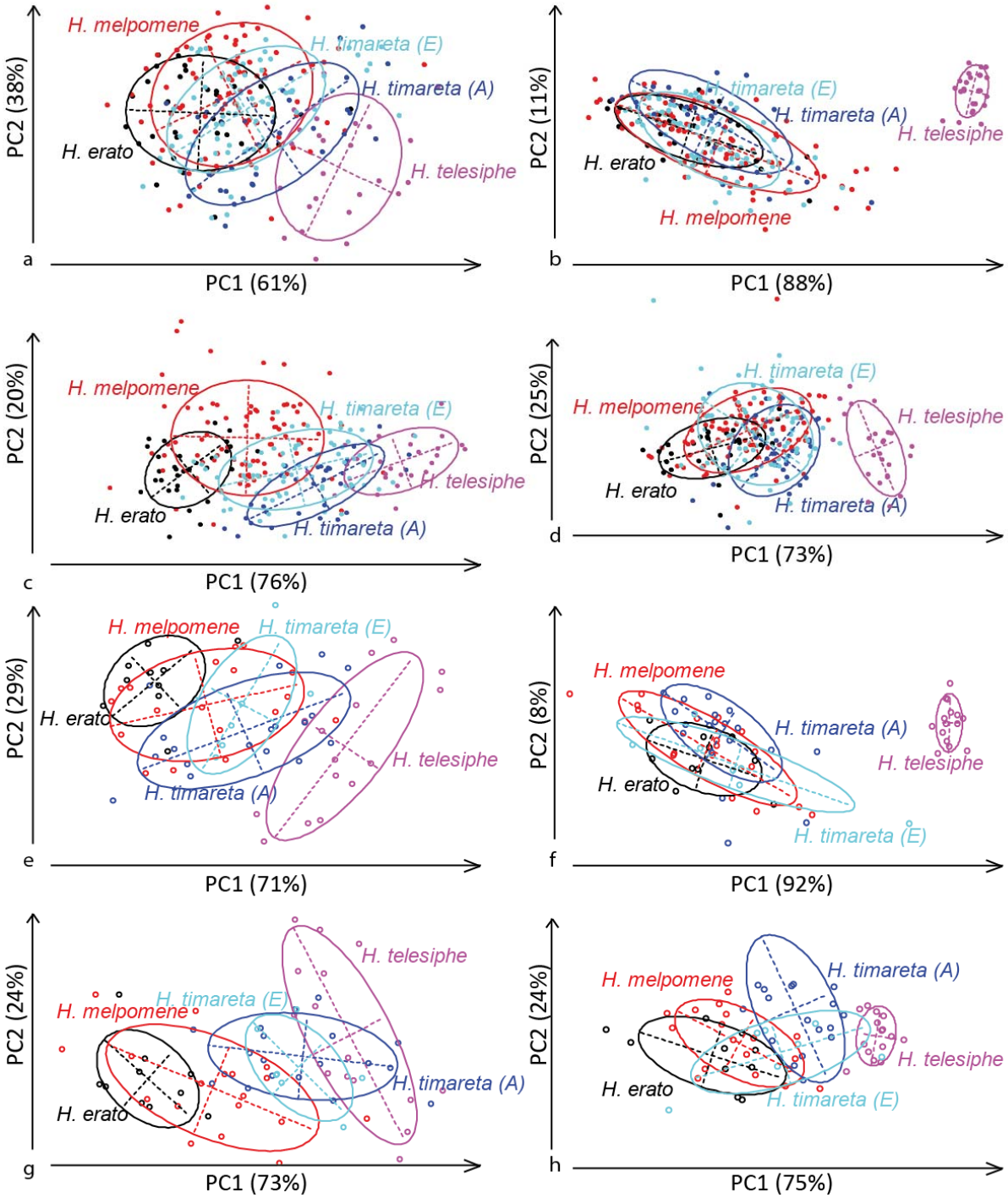


Figure S10: Principal component analyses based on the excitation of the four photoreceptors in the physiological model of Endler & Mielke for each colour patch

Bird V-type visual system and large gap light. (a) red patch on male dorsal FW, (b) yellow patch on male dorsal HW, (c) red patch on male ventral FW, (d) yellow patch on male ventral HW, (e) red patch on female dorsal FW, (f) yellow patch on female dorsal HW, (g) red patch on female ventral FW, (h) yellow patch on female ventral HW.

Distances and cross-validation

Table S7. Mahalanobis distances for each component of the phenotype between populations of *H. timareta* and their co-mimics (mean of the bootstrap ± sd)

		H. erato favorinus		*H. melpomene amaryllis*		*H. telesiphe*	
		H. timareta A	*H. timareta E*	*H. timareta A*	*H. timareta E*	*H. timareta A*	*H. timareta E*
Colour spectra	Red FW d	4.0±1.0	1.2± 0.5	2.3± 0.9	0.4± 0.3	4.3± 1.1	9.1± 1.9
	Red FW v	12.4±2.1	7.5± 1.8	6.7± 1.7	3.1± 1.2	2.9± 0.9	5.9± 1.6
	Yellow HW d	1.4±0.6	0.3± 0.2	1.1± 0.6	0.3± 0.3	53.3± 8.3	66.5± 10.1
	Yellow HW v	4.5±1.2	2.0± 0.8	2.0± 0.8	0.4± 0.3	7.2± 1.2	12.0± 2.2
Colour pattern	Red patch FW d	82.6±11.8	64.4± 11.1	34.3± 4.0	17.4± 2.9	561.7± 62.4	579.4± 62.4
	Yellow bar HW d	25.2±4.1	18.7± 3.7	17.5± 3.2	11.7± 2.8	59.4± 7.6	78.7± 9.2
Wing venation	FW venation	22.6±2.7	21.4± 2.5	15.1± 2.3	10.3± 1.9	37.2± 4.4	42.0± 4.8
	HW venation	25.9±2.9	24.6± 2.8	12.5± 2.0	12.1± 1.9	34.4± 3.7	35.8± 4.1
Wing Outline	FW outline	18.9±4.6	16.5± 3.9	13.4± 3.9	12.1± 3.5	32.9± 7.4	36.8± 8.2
	HW outline	17.3±4.6	12.2± 3.5	6.9± 2.2	6.8± 2.4	42.8± 8.0	38.5± 7.5

		H. timareta A	*H. erato*	*H. telesiphe*	*H. telesiphe*
		H. timareta E	*H. melpomene*	*H. melpomene*	*H. erato*
Colour spectra	Red FW d	1.2±0.5	0.7±0.4	11.9±2.8	15.9± 2.9
	Red FW v	0.9±0.5	3.3±1.3	15.5±3.2	26.1± 4.0
	Yellow HW d	1.3±0.7	0.5±0.4	64.1±10.1	69.9± 10.3
	Yellow HW v	1.3±0.7	1.3±0.7	14.1±2.7	22.0± 3.2
Colour pattern	Red patch FW d	9.6±1.7	43.3±9.9	597.4±61.4	617.3± 62.6
	Yellow bar HW d	5.0±1.1	8.7±2.1	84.7±10.6	88.1± 10.9
Wing venation	FW shape	3.1±0.7	16.6±2.3	52.9±6.5	33.2± 4.4
	HW shape	2.8±0.7	17.2±2.5	50.5±5.8	39.5± 3.9
Wing outline	FW outline	4.7±1.7	10.3±3.2	48.4±10.6	34.1± 7.2
	HW outline	3.7±1.2	14.3±3.7	64.1±11.6	44.7± 9.2

Table S8. Percentage of cross-validation in the linear discriminant analysis for each component of the phenotype between each population of *H. timareta* and their co-mimics

(1 = complete discrimination between the two groups compared, 100% of the individuals can be re-identified); mean of the bootstrap + sd

		H. erato		*H. melpomene*		*H. telesiphe*	
		H. timareta A	*H. timareta E*	*H. timareta A*	*H. timareta E*	*H. timareta A*	*H. timareta E*
Colour spectra	Red FW d	0.82 ± 0.04	0.69 ± 0.05	0.73 ± 0.05	0.55 ± 0.08	0.83 ± 0.03	0.92 ± 0.03
	Red FW v	0.99 ± 0.01	0.92 ± 0.03	0.90 ± 0.03	0.74 ± 0.05	0.78 ± 0.04	0.87 ± 0.03
	Yellow HW d	0.69 ± 0.04	0.54 ± 0.07	0.66 ± 0.05	0.50 ± 0.08	1.00 ± 0.00	0.99 ± 0.01
	Yellow HW v	0.85 ± 0.03	0.69 ± 0.05	0.75 ± 0.05	0.52 ± 0.09	0.98 ± 0.01	0.98 ± 0.01
Colour pattern	Red patch FW d	1.00 ± 0.00	0.99 ± 0.01	0.97 ± 0.02	0.92 ± 0.04	1.00 ± 0.00	1.00 ± 0.00
	Red patch FW v	1.00 ± 0.01	0.99 ± 0.01	0.97 ± 0.02	0.90 ± 0.04	1.00 ± 0.00	1.00 ± 0.00
	Yellow bar HW d	0.94 ± 0.03	0.94 ± 0.03	0.89 ± 0.04	0.85 ± 0.05	0.98 ± 0.01	1.00 ± 0.01
	Yellow bar HW v	0.99 ± 0.01	0.98 ± 0.02	0.96 ± 0.02	0.95 ± 0.03	0.99 ± 0.01	1.00 ± 0.01
Wing venation	FW shape	0.99 ± 0.01	0.69 ± 0.06	0.91 ± 0.04	0.85 ± 0.05	0.93 ± 0.03	0.99 ± 0.01
	HW shape	1.00 ± 0.01	0.67 ± 0.06	0.89 ± 0.04	0.90 ± 0.04	0.90 ± 0.04	1.00 ± 0.01
Wing outline	FW outline	0.93 ± 0.03	0.92 ± 0.04	0.84 ± 0.05	0.85 ± 0.05	0.99 ± 0.01	0.99 ± 0.01
	HW outline	0.89 ± 0.04	0.86 ± 0.05	0.74 ± 0.05	0.76 ± 0.06	0.98 ± 0.01	0.98 ± 0.02

		H. timareta A	*H. erato*	*H. telesiphe*	*H. telesiphe*
		H. timareta E	*H. melpomene*	*H. melpomene*	*H. erato*
Colour spectra	Red FW d	0.70 ± 0.05	0.61 ± 0.06	0.92 ± 0.03	0.98 ± 0.01
	Red FW v	0.64 ± 0.05	0.80 ± 0.05	0.97 ± 0.02	1.00 ± 0.00
	Yellow HW d	0.66 ± 0.05	0.58 ± 0.06	1.00 ± 0.00	1.00 ± 0.00
	Yellow HW v	0.65 ± 0.05	0.63 ± 0.06	0.98 ± 0.01	1.00 ± 0.00
Colour pattern	Red patch FW d	0.84 ± 0.05	0.98 ± 0.02	1.00 ± 0.00	1.00 ± 0.00
	Red patch FW v	0.85 ± 0.05	0.97 ± 0.02	1.00 ± 0.00	1.00 ± 0.00
	Yellow bar HW d	0.75 ± 0.05	0.81 ± 0.05	1.00 ± 0.00	0.99 ± 0.01
	Yellow bar HW v	0.79 ± 0.05	0.91 ± 0.04	1.00 ± 0.00	1.00 ± 0.00
Wing venation	FW shape	0.69 ± 0.06	0.85 ± 0.05	1.00 ± 0.01	0.93 ± 0.03
	HW shape	0.67 ± 0.06	0.90 ± 0.04	1.00 ± 0.01	0.90 ± 0.04
Wing outline	FW outline	0.68 ± 0.06	0.84 ± 0.05	0.99 ± 0.01	0.98 ± 0.02
	HW outline	0.66 ± 0.06	0.89 ± 0.04	1.00 ± 0.01	0.98 ± 0.02

Hybrid detection and population structure

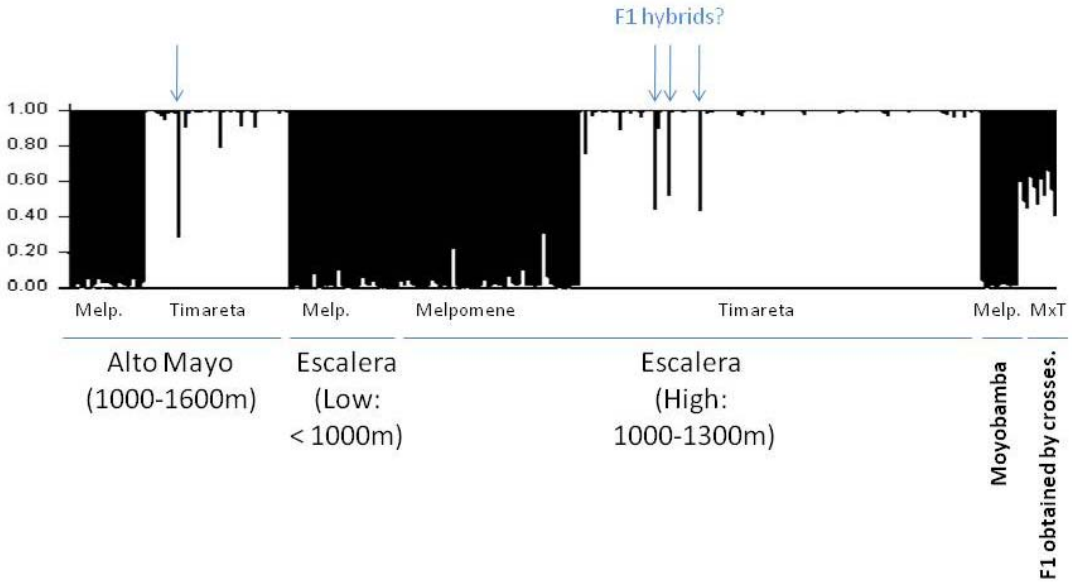


Figure S11. Multilocus Bayesian clustering and assignment analysis with STRUCTURE 2.3.1. Each individual is represented by a column and the colour represents the relative genome contribution of each cluster. The blue arrows point to the specimens identified as F1 hybrids with other assignment analyses (NewHybrids).

References

Belkhir, K., Borsa, P., Chikhi, L., Raufatse, N. & Bonhomme, F. (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

Chai, P. & Srygley, R.B. (1990) Predation and the Flight, Morphology, and Temperature of Neotropical Rain-Forest Butterflies. *American Naturalist,* **135,** 748-765.

Cheng H-D, Sun Y (2000) A hierarchical approach to color image segmentation using homogeneity. *Image Process IEEE Trans On* 9: 2071–2082.

Endler, J.A. & Mielke, P.W. (2005) Comparing entire colour patterns as birds see them. *Biological Journal of the Linnean Society*, **86,** 405-431.

Flanagan, N.S., Blum, M.J., Davidson, A., Alamo, M., Albarran, R., Faulhaber, K., Peterson, E. & McMillan, W.O. (2002) Characterization of microsatellite loci in neotropical *Heliconius* butterflies. *Molecular Ecology Notes*, **2,** 398-401.

Goudet, J. (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www2.unil.ch/popgen/softwares/fstat.htm. Updated from Goudet (1995)

Hart, N.S. (2002) Vision in the peafowl (Aves : *Pavo cristatus*). *Journal of Experimental Biology*, **205,** 3925-3935.

Hart, N.S., Partridge, J.C., Cuthill, I.C. & Bennett, A.T.D. (2000) Visual pigments, oil droplets, ocular media and cone photoreceptor distribution in two species of passerine bird: the blue tit (*Parus caeruleus* L.) and the blackbird (*Turdus merula* L.). *Journal of Comparative Physiology, A*, **186,** 375-387.

Ibáñez L, Insight Software Consortium (2003) The ITK software guide. [Clifton Park, N.Y.]: Kitware.
Kurugollu F, Sankur B, Harmanci AE (2001) Color image segmentation using histogram multithresholding and fusion. *Image Vis Comput* 19: 915–928.
Langham, G.M. (2006) Rufous-tailed jacamars and aposematic butterflies: do older birds attack novel prey? *Behavioral Ecology,* **17,** 285-290.
Le Poul, Y., Whibley, A., Chouteau, M., Prunier, F., Llaurens, V. & Joron, M. (2014) Evolution of dominance mechanisms at a butterfly mimicry supergene. . *Nature Communications,* **5**.
Martin K, Hoffman B (2003) Mastering CMake: a cross-platform build system. [Clifton Park, New York]: Kitware Inc.
Mattes D, Haynor DR, Vesselle H, Lewellen TK, Eubank W (2001) Nonrigid multimodality image registration. *Med Imaging* 4322: 1609–1620.
Mattes D, Haynor DR, Vesselle H, Lewellen TK, Eubank W (2003) PET-CT image registration in the chest using free-form deformations. *Med Imaging IEEE Trans On* 22: 120–128.
Mavarez, J. & Gonzalez, J. (2006) A set of microsatellites markers for *Heliconius melpomene* and closely related species. *Molecular Ecology Notes,* **6,** 20-23.
Mérot, C., Mavarez, J., Evin, A., Dasmahapatra, K.K., Mallet, J., Lamas, G. & Joron, M. (2013) Genetic differentiation without mimicry shift in a pair of hybridizing *Heliconius* species (Lepidoptera: Nymphalidae). *Biological Journal of the Linnean Society,* **109,** 830-847.
Meyer F, Beucher S (1990) Morphological segmentation. *J Vis Commun Image Represent* 1: 21–46.
Meyer F (1992) Color image segmentation. Image Processing and its Applications, 1992., International Conference on. pp. 303–306.
Nikolaev DP, Nikolayev PP (2004) Linear color segmentation and its implementation. *Comput Vis Image Underst* 94: 115–139. doi:10.1016/j.cviu.2003.10.012.
Odeen, A. & Hastad, O. (2003) Complex distribution of avian color vision systems revealed by sequencing the SWS1 opsin from total DNA. *Molecular Biology and Evolution,* **20,** 855-861.
Rohlf FJ, Slice D (1990) Extensions of the Procrustes Method for the Optimal Superimposition of Landmarks. *Syst Zool* 39: 40–59. doi:10.2307/2992207.
Rousset, F. (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources,* **8,** 103-106.
Shafarenko L, Petrou H, Kittler J (1998) Histogram-based segmentation in a perceptually uniform color space. *Image Process IEEE Trans On* 7: 1354–1358.
Styner M, Gerig G (1997) Evaluation of 2D/3D bias correction with 1+ 1ES-optimization. *Rapp Rech* 179.
Styner M, Brechbuhler C, Szckely G, Gerig G (2000) Parametric estimate of intensity inhomogeneities applied to MRI. *Med Imaging IEEE Trans On* 19: 153–165.
Théry, M., Pincebourde, S. & Feer, F. (2008) Dusk light environment optimizes visual perception of conspecifics in a crepuscular horned beetle. *Behavioral Ecology,* **19,** 627-634.
Thévenaz P, Unser M (2000) Optimization of mutual information for multiresolution image registration. *Image Process IEEE Trans On 9*: 2083–2099.
Yoo TS, Ackerman MJ, Lorensen WE, Schroeder W, Chalana V, et al. (2002) Engineering and algorithm design for an image processing api: a technical report on itk-the insight toolkit. *Stud Health Technol Inform*: 586–592.
Vorobyev, M. & Osorio, D. (1998) Receptor noise as a determinant of colour thresholds. *Proceedings of the Royal Society B-Biological Sciences,* **265,** 351-358.
Weir, B.S. & Cockerham, C.C. (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution,* **38,** 1358-1370.