Comparative Phylogenetics of Papilio Butterfly Wing Shape and Size Demonstrates Independent Hindwing and Forewing Evolution

H. L. Owens, D. S. Lewis, Fabien L. Condamine, A Kawahara, R Guralnick

TITLE: Comparative phylogenetics of *Papilio* butterfly wing shape and size demonstrates independent hindwing and forewing evolution

RUNNING TITLE: Comparative phylogenetics of *Papilio* wing morphology

AUTHORS: HL Owens*[,1,2], DS Lewis[3], FL Condamine[4], AY Kawahara[2], RP Guralnick[2]

*Corresponding author

1. Center for Macroecology, Evolution, and Climate, University of Copenhagen, Copenhagen, Denmark
2. Florida Museum of Natural History, University of Florida, Gainesville, FL USA
3. Department of Biology, Burman University, Lacombe, Alberta, Canada
4. CNRS, Institut des Sciences de l'Evolution de Montpellier (Université de Montpellier), Montpellier, France

CORRESPONDING AUTHOR DETAILS

Hannah L. Owens

Center for Macroecology, Evolution, and Climate

University of Copenhagen

Universitetsparken 15, byg 3

2100 Copenhagen Ø, Denmark

Email: hannah.owens@bio.ku.dk

ABSTRACT

The complex forces that shape butterfly wings have long been a subject of experimental and comparative research. Butterflies use their wings for flight, camouflage, mate recognition, warning and mimicry. However, general patterns and correlations among wing shape and size evolution are still poorly understood. We collected geometric morphometric measurements from over 1400 digitized museum specimens of *Papilio* swallowtails and combined them with phylogenetic data to test two hypotheses: 1) forewing shape and size evolve independently of hindwing shape and size, and 2) wing size evolves more quickly than wing shape. We also determined the major axes of wing shape variation and discovered that most shape variability occurs in hindwing tails and adjacent areas. We conclude that forewing shape and size are functionally and biomechanically constrained, while hindwings are more labile, perhaps in response to disruptive selective pressure for Batesian mimicry or against predation. The development of a significant, re-usable, digitized data resource will enable further investigation on tradeoffs between flight performance and ecological selective pressures, along with the degree to which intraspecific, local-scale selection may explain macroevolutionary patterns.

KEYWORDS: geometric morphometrics, museum specimens, Lepidoptera, Batesian mimicry

"…[O]n these expanded membranes Nature writes, as on a tablet, the story of the modifications of species, so truly do all changes of the organization register themselves thereon." – Henry Walter Bates on butterfly wings, 1863

For decades, researchers have examined butterfly wing diversity through lenses of functional adaptation, evolutionary history, and development. For nearly all Lepidoptera species, wings power flight to search for larval host plants, nectar sources, mates, and new territory (Scoble 1992). The physical requirements for powered flight are thought to exert natural selective pressure on lepidopteran wing size and shape; indeed, artificial selection experiments on wing and body size allometries have demonstrated significant fitness advantages for wild-type males compared those selectively bred for alternative allometries (Frankino et al. 2005). However, evidence suggests that forewings and hindwings unequally contribute to flight performance: in a study of 19 species of butterflies and 25 species of moths, all could fly with their hindwings removed, although at the cost of speed and maneuverability (Jantzen and Eisner 2008). Therefore, forewing shape and size may result from stabilizing selection imposed by the biomechanical requirements of flight, whereas hindwing shape and size may respond more readily to neutral or selective processes such as sexual selection (Chazot et al. 2016), and predation pressure (Sourakov 2013, Barber et al. 2015, Willmott et al. 2017, Rubin et al. 2018).

Still, experimental manipulations (e.g. Jantzen and Eisner 2008) cannot characterize processes at evolutionary time scales and across lineages. Comparative analysis of data from natural history collections may ameliorate this shortcoming by bridging the gap between experimental manipulation and observed macroevolutionary patterns. Strauss (1990) quantified variation in wing morphology in select heliconiine and ithomiine butterflies and found hindwings

were much more variable than forewings, providing a tantalizing link between functional studies and the impact of aerodynamic constraints on wing shape evolution. In contrast, a recent study of *Morpho* butterflies found a strong correlation between fore- and hindwing sizes as well as shapes (Chazot et al. 2016). Such datasets can also be used to identify morphological "paths of least resistance," axes along which diversification happens most quickly (Schluter 1996). Comparative studies of *Myotis* bat skulls (Dzeverin 2008) and whole *Pheidole* ants (Pie and Tschá 2013) found size evolved more quickly than shape, but size variation as an evolutionary path of least resistance remains untested in Lepidoptera.

We built on this groundwork to test two predictions via examination of swallowtail butterflies in one clade of the genus *Papilio* (subgenera *Agehana, Alexanoria, Chilasa, Heraclides* and *Pterourus*, hereafter "swallowtails in the clade of interest"): 1) forewing shape and size evolve independently of hindwing shape and size, and 2) wing size evolves more quickly than wing shape (Table 1). Our first prediction is based on the presumption that the forewing is functionally constrained while other selective pressures (e.g. predation, sexual selection) operate on hindwing shape. The second is based on the presumption that overall size change is an evolutionary path of least resistance. To test these hypotheses, we took morphometric measurements from digitized museum specimens and analyzed them in a comparative phylogenetic framework with a well-sampled and resolved species-level phylogeny (Owens et al. 2017).

METHODS

Morphometrics

Standardized dorsal and ventral images of *Papilio* butterfly specimens with scale and color bars were obtained from four natural history museums (Supplemental Fig. 1): the American Museum of Natural History (AMNH), Field Museum (FMNH), Florida Museum of Natural History, McGuire Center for Lepidoptera and Biodiversity (MGCL), and the Smithsonian Institution National Museum of Natural History (NMNH; Supplemental Table 1). Images were taken with a NIKON D300S with an AF Micro-Nikkor 60mm f/2.8D lens (AMNH), Canon EOS 70D with a Canon EF 50mm f/1.4 USM lens (FMNH), Canon EOS 70D with a Canon EF-S 60mm f/2.8 Macro USM lens (MGCL), or a Canon EOS 6D with a Canon EF 28-80mm f/3.5-5.6 lens (NMNH) mounted on either a copy stand or tripod and operated manually (AMNH) or tethered to a MacBook Air with the Canon EOS Utility (FMNH, MGCL, NMNH; Supplemental Fig. 2). Photographs were centered with white-space around images, where labels and color bar are located, to limit lens distortion. Specimen label data, collection date, location and sex of specimen (where available) were transcribed by volunteers via the Notes from Nature platform (Hill et al. 2012). All standardized images used in this study have been made publicly available (AMNH: https://datadryad.org/stash/share/Lq-XOj-IAFl-PyvNF3Es77l3n4WgsEKE6XrKTjPH7z8; FMNH: https://pj.fieldmuseum.org/event/626b0f98-98e7-49c8-903e-d67017fe2356; MGCL: LINK PENDING; NMNH: LINK PENDING).

Landmarks for morphometric measurement were based on previous morphological work on *Heraclides* swallowtails (Lewis et al. 2015; Fig. 1). One forewing landmark (F1 in Fig. 1), was removed from final analysis due to particularly high rate of measurement error; this was largely due to curling of the anterior wing margin in many specimens. To allow full view of

otherwise overlapping wing elements, we used 10 forewing landmarks from dorsal images and 12 hindwing landmarks from ventral images (Fig. 1). Landmark and 1 cm scale bar coordinates were collected in ImageJ 1.49 (https://imagej.nih.gov/ij/) using the PointPicker plugin (http://bigwww.epfl.ch/thevenaz/pointpicker/) and imported into Microsoft Excel for collation and post-processing. We collected 1,449 dorsal and 1,404 ventral landmark measurement sets representing 60 and 59 species, respectively (Supplemental Table 1). Measurements were calibrated with scale bar coordinates, and final coordinate text files were prepared using TextWrangler 5.5.2 (http://www.barebones.com/products/textwrangler/). These data are available in Appendix 1 (https://datadryad.org/stash/share/Lq-XOj-IAFl-PyvNF3Es77l3n4WgsEKE6XrKTjPH7z8). Forewing and hindwing landmark data with scale information were superimposed using Procrustes alignment; species represented by fewer than 10 specimens and statistical outlier landmark sets were removed, and the final datasets were re-aligned. We tested for allometric effects in fore- and hindwing shape variation using the homogeneity of slopes test (grouping specimens by species) and a subsequent Procrustes ANOVA to determine the contribution of size in determining shape variation. Mean species shape was calculated, the resulting dataset was re-aligned, gross outliers were removed, and allometric effects (grouping species into two subclades: *Agehana* + *Pterourus*, *Alexanoria*, *Chilasa*, and *Heraclides*, Fig. 1) were examined. Finally, mean intraspecific Procrustes distance from mean shape, morphospace volume occupied by each species (the product of the range of each principal component excluding values more than 1.5 times greater or less than the interquartile range), mean intraspecific centroid size, and intraspecific centroid variance were calculated from the specimen-level dataset. These four shape summary statistics were each rescaled to values between 0 and 1 using min-max normalization to render them comparable for

downstream analyses. All shape analyses were performed using the R package *geomorph* 3.1.2 (Adams et al. 2013), except for intraspecific Procrustes distance from mean shape, which was calculated using *Evomorph* 0.9 (Cabrera and Giri 2016; details can be found in Appendix 2).

Testing independence of forewing and hindwing shape evolution

Using a recently-published time-calibrated phylogeny for swallowtails in the clade of interest (Fig. 1; Appendix 1; Owens et al. 2017), we performed a suite of comparative phylogenetic analyses of forewing and hindwing morphology to test the independence of fore- and hindwing shape and size evolution. These analyses were all done in R 3.5.1 (R Core Team 2012; details of the analyses can be found in Appendix 2). We estimated phylogenetic signal of fore- and hindwing shapes using K_{mult}, (Blomberg et al. 2003), a generalization of Blomberg's *K* implemented in the R package *geomorph* 3.1.2 (Adams et al. 2013). Just as with the traditional Blomberg's *K* statistic, the closer K_{mult} is to 1, the more variation in species' characters can be explained by their phylogenetic relationships under a Brownian motion (BM) model of evolution, and values of $K < 1$ indicate more variation than expected under BM (Blomberg et al. 2003). We then used *geomorph* to calculate the modular rate ratio (function: *compare.multi.evol.rates*) for the fore- and hindwing datasets (Denton and Adams 2015), and test whether fore- and hindwing shape evolution was correlated (also known as phylogenetic integration, function: *phylo.integration*; Adams and Felice 2014). Both tests were conducted under an assumption of BM, as this is the only model currently available for such multidimensional datasets implemented in *geomorph* (Adams and Collyer 2018).

Testing independence of fore- and hindwing size evolution

We also tested the independence of forewing and hindwing size evolution. First, we estimated the phylogenetic signal (univariate Blomberg's K), then estimated the evolutionary rates of fore- and hindwing shape based on the best-fit evolutionary model for each dataset, and finally, tested for significant fore- and hindwing centroid size correlation. These analyses were done using the R packages *phytools* 0.6-60 (Revell 2012), *geiger* 2.0.6 (Harmon et al. 2008), and *phylolm* 2.6 (Tung Ho and Ané 2014); additional details can be found in Appendix 2. Correlations were assessed using a linear-time algorithm developed by Tung Ho and Ané (2014). For the correlation between fore- and hindwing centroid size, we fit a series of phylogenetic linear regressions with different evolutionary models—BM, Ornstein-Uhlenbeck (OU) with root ancestral state estimated from the data, and early burst (EB). We chose the best-fit model for the relationship between fore- and hindwing centroid size based on its minimum corrected Akaike Information Criterion (AICc) score as calculated using the R package *MuMIn* 1.42.1 (Barton 2018) and examined the ratio of hindwing to forewing evolutionary rates (σ^2) under these best-fit models using *geomorph*. To get a clear picture of the degree to which fore- and hindwing centroid sizes were correlated, we calculated R^2_{pred} for hindwing size as explained by forewing size; this is an extension of the traditional R^2 that weighs the residuals by variance *and* covariance (Ives 2018). These calculations were done using the R package *rr2* 1.0.1 (Ives and Li 2018; Appendix 2).

Differences in speed of shape and size evolution

We made a final comparison of rates of fore- and hindwing shape and size evolution (σ^2) as obtained from the analyses described above to determine the relative evolutionary lability of these four characteristics. Appendix 2 provides an R markdown script of all analyses performed.

RESULTS

Morphometrics

Forewing measurements for 1,449 specimens representing 60 species and hindwing measurements for 1,404 specimens representing 59 species were analyzed after specimen-level datasets were cleaned. When the datasets were reduced to species for which there were available measurements for at least 10 individuals, 49 species remained in the forewing dataset and 47 in the hindwing dataset. Principal component (PC) 1 of the forewing describes elongation of the wing from the apex to the base, explaining 45% of species' mean shape variation; PC 2 describes wing margin angularity, explaining 20% of species' mean shape variation (Fig. 2). PC 1 of the hindwing primarily describes tail length, explaining 67% of variation, while PC 2 describes scalloping of the outer margin, explaining 8% of variation (Fig. 2).

Shape does not covary with size consistently for all species in the specimen-level dataset and for all subgenera in the mean species shape dataset (homogeneity of slopes test, $P < 0.05$), which suggests interspecific allometric effects likely do not play a complicating role in the patterns we examined in this study. However, correlations between log size and shape are significant ($P < 0.05$), albeit small, for all datasets (forewing specimen-level: $R^2 = 0.02$; hindwing specimen-level: $R^2 = 0.15$; forewing mean shape: $R^2 = 0.15$; hindwing mean shape: $R^2 = 0.21$). Summary statistics quantifying intraspecific shape variation (mean Procrustes distance of individuals from mean shape, morphospace volume occupied by each species, mean centroid size, and centroid size variance) can be found in Supplemental Table 2.

Independence of Fore- and Hindwing shape

Comparative phylogenetic analyses of mean species wing shape indicate a stronger phylogenetic signal in forewing than hindwing shape datasets (Table 2, Fig. 2). The evolutionary rate of the hindwing shape (under an assumption of BM) was also 9 times faster (a statistically significant difference, $P < 0.05$) than that of the forewing (forewing $\sigma^2 = 6.6 \times 10^{-5}$; hindwing $\sigma^2 = 6.01 \times 10^{-4}$). However, despite this disparity in evolutionary rates, fore- and hindwing shapes at the species-level are strongly integrated (r_{PLS}: 0. 75; $P < 0.05$).

Independence of Fore- and Hindwing size

Comparative phylogenetic analyses of wing centroid size yielded similar patterns. Phylogenetic signal for both fore- and hindwing centroid size is low to moderate (Table 2, Supplemental Fig. 3), but the statistical significance of these estimates is marginal for the forewing and not significant for the hindwing ($P = 0.04$, $P = 0.18$, respectively). The OU model fits the forewing dataset best, while the white noise model fits the hindwing dataset best. However, for forewing size, the white noise model was the next-best fit, while for hindwing size, the OU model was the next-best fit (Supplemental Table 4). Therefore, we calculated evolutionary rate based on the OU model for both hindwing and forewing, as this was the best-fit model for both datasets from which evolutionary rate can be calculated (Harmon et al. 2008). Hindwing centroid size evolution is 90 times the forewing rate; this difference is statistically significant (forewing $\sigma^2 = 8.4 \times 10^{-8}$, hindwing $\sigma^2 = 1.1 \times 10^{-5}$, $P < 0.05$). Fore- and hindwing centroid size are mildly but positively and significantly ($P < 0.05$) correlated ($R^2 = 0.22$; $\beta = 3.58$).

Differences in speed of shape and size evolution

Based on our σ^2 estimates, hindwing shape is evolving fastest ($\sigma^2 = 6.01 \times 10^{-4}$), followed by forewing shape ($\sigma^2 = 6.6 \times 10^{-5}$), hindwing size ($\sigma^2 = 1.1 \times 10^{-5}$), and forewing size ($\sigma^2 = 8.4 \times 10^{-8}$), in that order (Table 2). Appendix 3 is an R Markdown report of the corresponding results of analytical code supplied in Appendix 2.

DISCUSSION

Our results demonstrate that fore- and hindwings are subject to different selective pressures and are evolving autonomously from each other, although there is also evidence of balancing constraint. Notably, phylogenetic signal is stronger in the forewing compared to the hindwing. This pattern is consistent with findings from an early comparative study that demonstrated much higher hindwing shape variation compared to forewings in ithomiine and heliconiine butterflies (Strauss, 1990). Our findings also provide evidence that swallowtail forewing shape and size evolve more slowly than respective hindwing measures, perhaps due to stabilizing selection imposed by dependence on forewings for flight (Jantzen and Eisner 2008). If so, hindwings may be responding more readily to stochastic forces than forewings as the result of factors that can be geographically quite localized and variable (e.g. predation pressure; Barber et al. 2015; Rubin et al. 2018). However, further comparative scrutiny of this trend among butterflies with differing flight behaviors would be beneficial. A recent study in experimental wing reduction in *Pierella helvina*, a floor-gliding heteriine butterfly, demonstrated enlarged hindwings greatly improved flight performance (Stylman et al. 2019).

In contrast to previous studies on bats and ants that found size evolved more quickly than shape (Dzeverin 2008; Pie and Tschá 2013), swallowtails in the clade of interest, shape has evolved at least an order of magnitude more quickly than size in both the hindwing and forewing (Table 2). Indeed, for the full clade of interest as well as both subclades, hindwing shape has the fastest rate of evolution, followed by forewing shape, hindwing size, and forewing size, in that order. This suggests that for butterflies, and in contrast to predictions (Table 1), hindwing shape is the path of least resistance to morphological diversification. This is not to say the evolution of shape and size in the clade is necessarily adaptive — indeed, the difference between hindwing and forewing evolutionary rate ratios for centroid size and wing shape, while dramatic, may also bear the signal of stochasticity in hindwing evolution. Our results should also be interpreted carefully due to a methodological limitation—evolutionary rates for the shape datasets could only be inferred under a BM model using existing methods (Adams and Collyer 2018). If the fit of an OU model could be assessed and was found to be a better fit than BM, it is likely that lower evolutionary rates would be inferred for shape than those found here, as the OU model constrains character evolution around a central location.

One of the most labile characteristics of hindwing shape is presence and size of tails, as can be seen examining our hindwing shape deformation grids in relation to PC1 (Fig. 2), which explains 77% of hindwing shape variation. While it is tempting to think of tails as a presence-absence trait, there are a wide range of tail shapes and relative sizes compared to the rest of the wing (e.g. Fig. 1), ranging from highly prominent to entirely absent. Tail form lability, visible both in the wide range of tail shapes and sizes and in models of hindwing evolution that appear primarily stochastic, remains understudied. Longer tails have been argued to increase aerodynamic performance for lepidopterans (Park et al. 2010) as well as improving the odds of

escaping predators (Barber et al. 2015). However, the tradeoff between costs associated with producing tails versus their benefit requires closer examination, as it is likely dependent on complex interactions among flight behavior characteristics, biotic interactions, and microhabitat.

Mimicry may be especially critical as a driver of hindwing shape evolution and tail shape differences. Mimicry selection in *Papilio* butterflies, often for dramatically different models, appears to have had a profound effect across the clade. This is particularly true for hindwings in swallowtails in the clade of interest, and especially in relation to tail shape because most species with strongly reduced tails (e.g. those species with positive values along PC1) are also mimetic (Supplemental Fig. 4). Overall, 39 of the 60 taxa included in our study have at least one mimetic sex (Supplemental Table 5; Supplemental Fig. 4). Examples of putative Batesian mimic taxa include *P. (Pterourus) zagreus*, which mimics *Heliconius* and a number of genera in Danainae (Tyler et al. 1994); *P. (Chilasa) clytia*, which mimics *Euploea* (Kunte 2009), and *Heraclides* of the *anchisiades* group, which mimics *Parides* (Srygley and Chai 1990; Tyler et al. 1994). Additional mimetic taxa that were not well-sampled enough to incorporate into our final analyses or were removed because they were strong outliers, such as *P. (Chilasa) laglaizei*, a mimic of the uraniid moth *Alcidis agarthyrsus* (Collins and Morris 1985), and *P. (Pterourus) euterpinus*, a mimic of heliconiine butterflies (Tyler et al. 1994).

Our results suggest that for swallowtails, selective pressure for mimicry is a much stronger driver of morphology than shared phylogenetic history. This highlights the importance of predation interactions for the evolution of these lineages. Kunte (2009) demonstrated *Papilio* butterflies do not appear to have co-evolved with their models, but instead may have adapted to existing models after colonizing new areas. This result is consistent with the findings of studies in other organisms, such as ant-mimic jumping spiders (Ceccarelli and Crozier 2007), coral–

294 snake–mimic colubrid snakes (Rabosky et al. 2016), and army–ant–mimic rove beetles
295 (Maruyama and Parker 2017). None of these groups appear to have co-evolved with their
296 models, but instead may have taken advantage of already-established model patterns. This
297 hypothesis deserves further study by reconstructing the evolution of shape in mimetic *Papilio*
298 lineages (and outward into the family Papilionidae) and comparing the result to the inferred
299 biogeographic history of these lineages and their putative models.

300 Owing to a preponderance of male and unsexed specimens in our study (and natural
301 history collections in general), we were unable to fully explore differential patterns of wing
302 shape evolution in the context of sexual dimorphism. Previous studies of evolutionary processes
303 related to sex-limited mimicry have not focused on wing shape; instead, studies have diagnosed
304 mimics based on coloration (Kunte, 2009) or examined evolution of key mimicry-relevant genes
305 (Timmermans et al. 2017). Thus, a critical, unanswered question is how different drivers of
306 dimorphism operate across the clade, and we hope that this can be addressed in the future by
307 leveraging our growing database of landmarked *Papilio* wings. Despite current sample
308 limitations, we were able to examine two species (*P. androgeus* and *P. scamander*) for which we
309 had data to make statistically relevant comparisons between male and female specimens. For
310 those two species, we found no evidence of sexual dimorphism in shape or size in forewings or
311 hindwings (script Appendix 3; R Markdown report Appendix 4).

312 In conclusion, our study demonstrates how digitized museum specimens can bridge the
313 gap between taxonomically- and temporally-limited experimental manipulations (e.g. Frankino
314 et al. 2005; Jantzen and Eisner 2008) and broad-scale macroevolutionary hypotheses. We
315 recovered evidence that *Papilio* forewings and hindwings are evolving independently, which is
316 consistent with experimental observations that fore- and hindwings have different effects on

butterfly flight dynamics. Furthermore, hindwing shape may be an evolutionary path of least resistance for morphological diversification in butterflies and may reflect strong disruptive selection for mimicry and/or for predation defense. Still, this study is a first glance at the evolutionary relationships between hindwing and forewing shape and size in insects, and future studies are needed to investigate these patterns more thoroughly within swallowtails and more generally across the insect tree of life. Such work will require detailed information on species' phylogenetic relationships and wing morphologies, as well as factors including mimetic systems, within-species geographic variation, and flight behavior. Fortunately, with the systematic digitization of museum specimens and the increasing capacity of researchers to manage large, complex datasets, the answers to these questions are closer to our reach than ever before.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: https://datadryad.org/stash/share/Lq-XOj-IAFl-PyvNF3Es77l3n4WgsEKE6XrKTjPH7z8

FUNDING

This work was supported by National Science Foundation grant no. DEB 1523732 to H.L.O., DEB 1557007 to A.Y.K. and R.P.G., and an “Investissements d’Avenir” grant managed by Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01) to F.L.C.

ACKNOWLEDGEMENTS

We are very grateful to the curators and collections managers that granted us access to their collections in order to image specimens, particularly David Grimaldi, Courtney Richenbacher,

and Suzanne Rab Green (AMNH); Crystal Maier (FMNH, now MCZ); Andrew Warren and Laurel Kaminsky (MGCL); and Robert K. Robbins and Brian Harris (NMNH). Laura Brenskelle, Josh Dieringer, Toshita Barve, and Vijay Barve (University of Florida) were instrumental in data collection. Chris Hamilton (University of Idaho) engineered the portable light box used on museum visits.

REFERENCES

Adams, D.C., Collyer, M.L. 2018. Multivariate Phylogenetic Comparative Methods: Evaluations, Comparisons, and Recommendations. *Syst. Biol.* 67:14–31.

Adams, D.C., Felice, R.N. 2014. Assessing trait covariation and morphological integration on phylogenies using evolutionary covariance matrices. *PloS One*. 9:565-572.

Adams, D. C., Otárola-Castillo, E., Paradis, E. 2013. geomorph: An R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* 4:393–399.

Allio, R., Scornavacca, C., Nabholz, B., Clamens, A.L., Sperling, F.A.H., Condamine, F.L. 2020. Whole genome shotgun phylogenomics resolves the pattern and timing of swallowtail butterfly evolution. *Syst. Biol.* 69:38–60.

Bai, Y., Bin Ma, L., Xu, S.-Q., Wang, G.-H. 2015. A geometric morphometric study of the wing shapes of *Pieris rapae* (Lepidoptera: Pieridae) from the Qinling Mountains and adjacent regions: an environmental and distance-based consideration. *Fla. Entomol.* 98:162–169.

Barber, J.R., Leavell, B.C., Keener, A.L., Breinholt, J.W., Chadwell, B.A., McClure, C.J.W., Hill, G.M., Kawahara, A.Y. 2015. Moth tails divert bat attack: evolution of acoustic deflection. *P. Natl. Acad. Sci. USA.* 112:2812–2816.

Bartoń, K. 2019. MuMIn: Multi-Model Inference. R package version 1.43.6. https://CRAN.R-project.org/package=MuMIn

Bates, H. W. 1863. *The Naturalist on the River Amazons*. John Murray, London, UK.

Blomberg, S.P., Garland, T., Ives, A.R. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57:717–745.

Cabrera, J.M., Giri, F. 2016. Evomorph: evolutionary morphometric simulation. R package version 0.9. https://cran.r-project.org/web/packages/Evomorph/index.html.

Ceccarelli, F.S., Crozier, R.H. 2007. Dynamics of the evolution of Batesian mimicry: molecular phylogenetic analysis of ant-mimicking Myrmarachne (Araneae: Salticidae) species and their ant models. *J. Evol. Biol.* 20:286–295.

Cespedes, A., Penz, C.M., DeVries, P.J. 2015. Cruising the rain forest floor: butterfly wing shape evolution and gliding in ground effect. *J. Anim. Ecol*. 84:808–816.

Chazot, N., Panara, S., Zilbermann, N., Blandin, P., Le Poul, Y., Cornette, R., Elias, M., Debat, V. 2016. Morpho morphometrics: shared ancestry and selection drive the evolution of wing size and shape in *Morpho* butterflies. *Evolution* 70:181–194.

Collins, N. M., and M. G. Morris. 1985. *Threatened Swallowtail Butterflies of the World*. IUCN, Gland and Cambridge. 401 pp.

Frankino, W.A., Zwaan, B.J., Stern, D.L., Brakefield, P.M. 2005. Natural selection and developmental constraints in the evolution of allometries. *Science* 307:718-720.

Dzeverin, I. 2008. The stasis and possible patterns of selection in evolution of a group of related species from the bat genus *Myotis* (Chiroptera, Vespertilionidae). *J. Mamm. Evol.* 15:123–142.

384 Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E., Challenger, W. 2008. GEIGER: investigating
385 evolutionary radiations. *Bioinformatics* 24:129–131.
386 Hill, A., Guralnick, R.P., Smith, A., Sallans, A., Gillespie, R., Denslow, M., Gross, J., Murrell,
387 Z., Conyers, T., Oboyski, P., Ball, J., Thomer, A., Prys-Jones, R., de la Torre, J., Kociolek,
388 P., Fortson, L. 2012. The notes from nature tool for unlocking biodiversity records from
389 museum records through citizen science. *ZooKeys* 209:219-233.
390 Ives, A.R. 2018. R^2s for correlated data: phylogenetic models, LMMs, and GLMMs. *Syst. Biol.*
391 68:234–251.
392 Ives, A.R., Li, D. 2018. rr2: an R package to calculate R^2s for regression models. *J. Open Source*
393 *Softw*. 3:1028.
394 Jantzen, B., Eisner, T. 2008. Hindwings are unnecessary for flight but essential for execution of
395 normal evasive flight in Lepidoptera. *P. Natl. Acad. Sci. USA* 105:16636–16640.
396 Kunte, K. 2009. The diversity and evolution of Batesian mimicry in *Papilio* swallowtail
397 butterflies. *Evolution* 63:2707–2716.
398 Lewis, D.S., Sperling, F.A.H., Nakahara, S., Cotton, A.M., Kawahara, A.Y, Condamine, F.L.
399 2015. Role of Caribbean Islands in the diversification and biogeography of Neotropical
400 *Heraclides* swallowtails. *Cladistics* 31:291-314.
401 Maruyama, M., Parker, J. 2017. Deep-time convergence in rove beetle symbionts of army ants.
402 *Curr. Biol.* 27:920–926.
403 Owens, H.L., Lewis, D.S., Dupuis, J.R., Clamens, A.L., Sperling, F.A.H., Kawahara, A.Y.,
404 Guralnick, R.P., Condamine, F.L. 2017. The latitudinal diversity gradient in New World
405 swallowtail butterflies is caused by contrasting patterns of out- of- and into- the- tropics
406 dispersal. *Glob. Ecol. Biogeogr.* 26:1447–1458.

Park, H., Bae, K., Lee, B., Jeon, W.P., Choi, H. 2010. Aerodynamic performance of a gliding swallowtail butterfly wing model. *Exp. Mech.* 50:1313–1321.

Pie, M.R., Tschá, M.K. 2013. Size and shape in the evolution of ant worker morphology. *PeerJ* 1:e205.

Punnett, R.C. 1916. Mimicry in butterflies. *Nature* 97:237–238.

R Core Team. 2012. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rabosky, A.R.D., Cox, C.L., Rabosky, D.L., Title, P.O., Holmes, I.A., Feldman, A., McGuire, J.A. 2016. Coral snakes predict the evolution of mimicry across New World snakes. *Nat. Commun.* 7:11484.

Revell, L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3:217–223.

Rubin, J.J., Hamilton, C.A., McClure, C.J., Chadwell, B.A., Kawahara, A.Y., Barber, J.R. 2018. The evolution of anti-bat sensory illusions in moths. *Sci. Adv.* 4:eaar7428.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* 50:1766–1774.

Schluter, D. 2000. *The Ecology of Adaptive Radiation*. Oxford University Press, New York.

Scoble, M.J. 1992. *The Lepidoptera: form, function and diversity*. Oxford University Press, Oxford, UK.

Sourakov, A. 2013. Two heads are better than one: false head allows *Calycopis cecrops* (Lycaenidae) to escape predation by a Jumping Spider, *Phidippus pulcherrimus* (Salticidae). *J. Nat. Hist.* 47:1047–1054.

Srygley, R.B., Chai, P. 1990. Flight morphology of neotropical butterflies: palatability and distribution of mass to the thorax and abdomen. *Oecologia* 84:491–499.

Strauss, R.E. 1990. Patterns of quantitative variation in lepidopteran wing morphology: the convergent groups Heliconiinae and Ithomiinae. *Evolution* 44:86–103.

Stylman, M., Penz, C.M., and P. DeVries. 2020. Large hind wings enhance gliding performance in ground effect in a neotropical butterfly (Lepidoptera: Nymphalidae). *Ann. Entomol. Soc. Am.* 113:15–22.

Timmermans, M.J.T.N., Thompson, M.J., Collins, S., Vogler, A.P. 2017. Independent evolution of sexual dimorphism and female-limited mimicry in swallowtail butterflies (*Papilio dardanus* and *Paplio phorcas*). *Mol. Ecol*. 26:1273-1284.

Tyler, H., Brown, K.S.J., Wilson, K. 1994. *Swallowtail butterflies of the Americas. A study in biological dynamics, ecological diversity, biosystematics and conservation*. Scientific Publishers, Gainesville, FL, USA.

Tung Ho, L.S., Ané, C. 2014. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol*. 63:397–408.

Willmott, K.R., Willmott, J.C.R., Elias, M., Jiggins, C.D.. 2017. Maintaining mimicry diversity: Optimal warning colour patterns differ among microhabitats in Amazonian clearwing butterflies. *P. Roy. Soc. B-Biol. Sci*. 284:20170744.

TABLES

Table 1: Hypotheses examined in this study, with predictions regarding phylogenetic signal and evolutionary rate.

Hypothesis	Predictions
Forewing shape is evolving independently of hindwing shape	Forewing K_{mult} > Hindwing K_{mult} Forewing $\sigma^2 \neq$ Hindwing σ^2 (Forewing σ^2 / Hindwing σ^2) > 1 Forewing and hindwing $r_{PLS} < 1$
Forewing size is evolving independently of hindwing size	Forewing K_{mult} > Hindwing K_{mult} Forewing $\sigma^2 \neq$ Hindwing σ^2 (Forewing σ^2 / Hindwing σ^2) > 1 Forewing and hindwing $R^2 < 1$
Size is evolving more quickly than shape	Forewing shape σ^2 < Forewing size σ^2 Hindwing shape σ^2_{shape} < Hindwing size σ^2

Table 2: Comparative phylogenetics statistics for species-level shape and size. Wing shape generally has more significant phylogenetic signal and evolves more quickly than centroid size; forewing size and shape generally shows stronger phylogenetic signal and evolves more slowly than hindwing size and shape. Full clade: *Agehana* + *Alexanoria* + *Chilasa* + *Heraclides* + *Pterourus*; Non-*Heraclides*: *Agehana* + *Alexanoria* + *Chilasa* + *Pterourus*. Statistically significant values ($p < 0.05$) are bolded.

		Full Clade	*Heraclides*	Non-*Heraclides*
Shape	Forewing K_{mult}	**0.4977**	**0.2870**	**0.5495**
	Hindwing K_{mult}	**0.3234**	**0.2448**	**0.5691**
	Forewing $\sigma^{2\dagger}$	6.59×10^{-5}	7.03×10^{-5}	5.88×10^{-5}
	Hindwing $\sigma^{2\dagger}$	6.00×10^{-4}	7.09×10^{-4}	4.59×10^{-4}
	Evolutionary Rate Ratio	**9.1106**	**10.0861**	**7.8094**
	Fore- and hindwing correlation (r_{PLS})	**0.7534**	**0.8182**	**0.7518**
Centroid Size	Forewing K	**0.4511**	0.1787	**0.6636**
	Hindwing K	0.2172	0.0960	0.2635
	Forewing σ^2	$\mathbf{8.50 \times 10^{-8}}$	$\mathbf{5.70 \times 10^{-8}}$	$\mathbf{1.12 \times 10^{-7}}$
	Hindwing σ^2	$\mathbf{1.10 \times 10^{-5}}$	$\mathbf{8.60 \times 10^{-6}}$	$\mathbf{1.31 \times 10^{-5}}$
	Evolutionary Rate Ratio	**130.2796**	**151.2126**	**116.9315**
	Fore- and hindwing correlation (R^2)†	0.2174	0.0147	0.1455

† Does not have a *p* value

FIGURES

Figure 1. Geometric morphometric landmarks and phylogeny used for analysis. Phylogeny shows non-monophyly of New World swallowtails (*Heraclides* + *Pterourus*) and subgenus *Pterourus*. Shape landmarks, indicated by dots, adapted from Lewis et al. (2015); landmark F1 removed from final analysis due to particularly high rate of measurement error. Phylogenetic relationships from Owens et al. (2017) with bars indicating currently recognized subgenera; bar colors correspond with subsequent figure plots. Images depict species corresponding to each labeled clade.

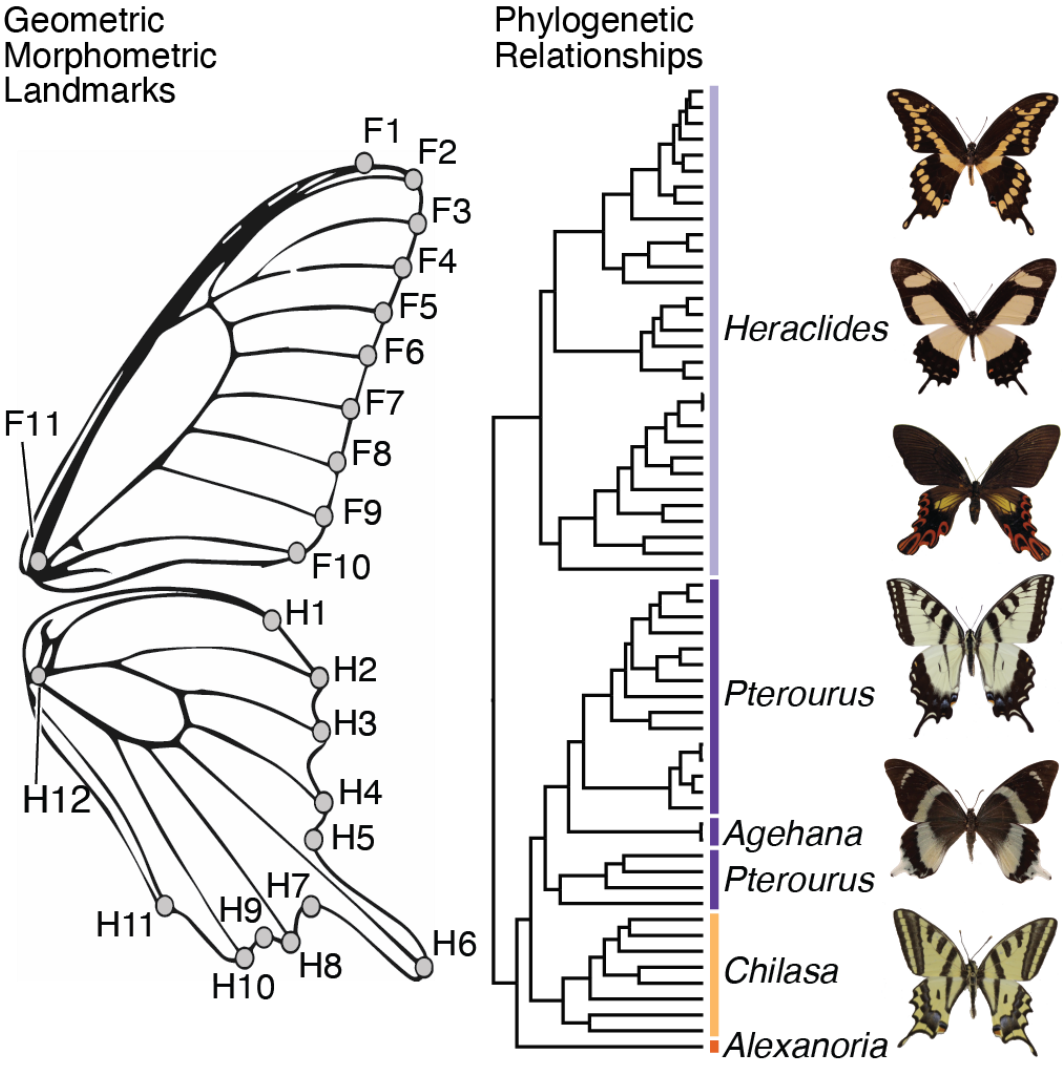

Figure 2. Forewing and hindwing phylomorphospace plots; species' forewing shapes tend to cluster phylogenetically, whereas hindwing shapes do not. Principal components for each dataset were calculated in *geomorph*; the phylogeny is plotted on top of these, along with inferred node states (under a Brownian motion model—no other models are yet available for this method, Revell 2012), which are color-coded by clade. Warp grid deformations show contributions of principal components 1 and 2 to overall shape. Plots include Blomberg's *K* and variance contributions for the first two principal axes of the fore- and hindwing; tips and nodes of phylogeny are color-coded by clade.

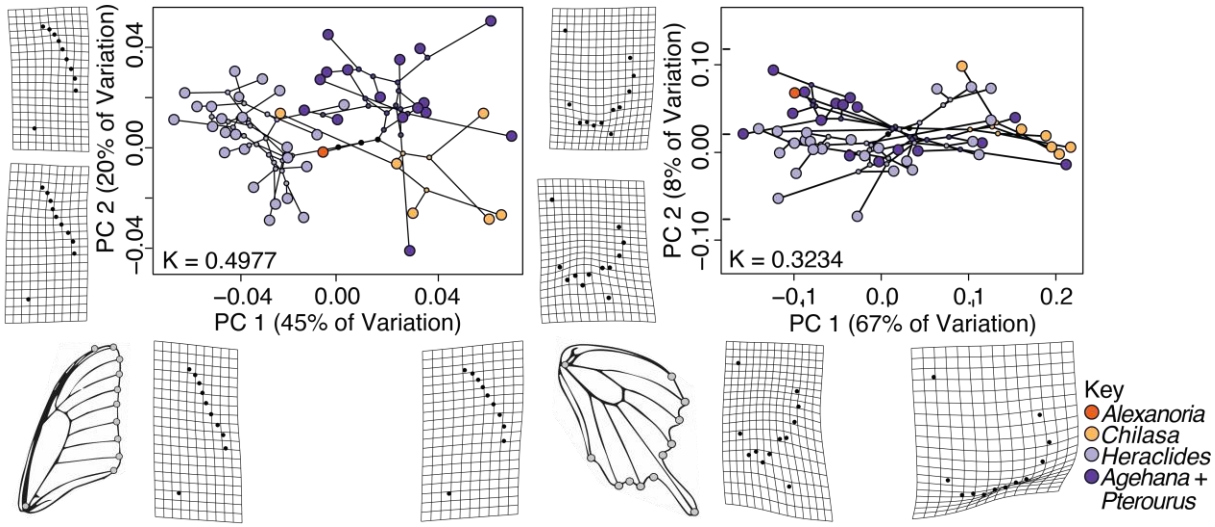

Figure 3. Mean species centroid size in phylomorphospace; *Chilasa* fore- and hindwing sizes are highly correlated and unique compared to the rest of the clade of interest. Scatterplot shows fore- and hindwing centroid sizes. Black lines indicate phylogenetic relationships; small points at nodes indicate inferred ancestral character values (under a Brownian motion model—no other models are yet available for this method, Revell 2012). Tips and nodes are color-coded by clade.