Selection on size has generated distinctive paired wing flight systems for butterfly flight and migration

Richard Rabideau Childers
Harvard University https://orcid.org/0000-0002-7137-3192
Liming Cai
Department of Integrative Biology, The University of Texas, Austin
Shawan Chowdhury
The University of Queensland
James Crall
Harvard University https://orcid.org/0000-0002-8981-3782
Mark Cornwall
Harvard University
Rachel Hawkins Sipe
Harvard University
Crystal Maier
Harvard University
Sarah Maunsell
Harvard University
Gerard Talavera
CSIC https://orcid.org/0000-0003-1112-1345
Cheng-Chia Tsai
Columbia University https://orcid.org/0000-0001-5662-6908
Katherine Angier
Harvard University
Vijay Barve
University of Florida
Evan Dankowicz
Harvard University
Jomar Hinolan
National Museum of the Philippines https://orcid.org/0000-0002-4446-3803
Micael Itliong
City University of New York
Mark Arcebal Naive
University of Chinese Academy of Sciences
Gunnar Johnson

Harvard University
Francisco Matos
Harvard University
Valentina Morgan
City University of New York
David Plotkin
University of Florida
Mary Salcedo
Cornell University https://orcid.org/0000-0003-3766-8437
Vaughn Shirey
Georgetown University
Kimberly Vermilya
Albany Medical College
Jalen Winstanley
Harvard University
Amy Wu
Harvard University
Robert Guralnick
University of Florida
Akito Kawahara
University of Florida https://orcid.org/0000-0002-3724-4610
David Lohman
City College of New York https://orcid.org/0000-0002-0689-2906
Leslie Ries
Georgetown University
Edward Soucy
Harvard University https://orcid.org/0000-0002-1187-5596
Roger Vila
Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra) https://orcid.org/0000-0002-2447-4388
Nanfang Yu
Columbia University https://orcid.org/0000-0002-9462-4724
Naomi Pierce
Harvard University https://orcid.org/0000-0003-3366-1625
Wei-Ping Chan (w.p.chan23@gmail.com)
Harvard University https://orcid.org/0000-0002-7132-9191

Biological Sciences - Article

Keywords:

Posted Date: May 31st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2968821/v1

Additional Declarations: There is **NO** Competing Interest.

For submission to *Nature*

Selection on size has generated distinctive paired wing flight systems for butterfly flight and migration

Authors: Wei-Ping Chan[1,2*]†, Richard Rabideau Childers[1,2*], Liming Cai[3], Shawan Chowdhury[4-7], James D. Crall1[1,2,8], Mark A. Cornwall[1,2], Rachel Hawkins Sipe[2], Crystal A. Maier[2], Sarah C. Maunsell[1,2], Gerard Talavera[9], Cheng-Chia Tsai[10], Katherine Angier[1,2], Vijay Barve[11], Even Dankowicz[2], Jomar D. Hinolan[12], Micael G. A. Itliong[13], Mark Arcebal K. Naive[14-16], Gunnar Johnson[1,2], Francisco Matos[1,2], Valentina Morgan[17], David Plotkin[11,18], Mary K. Salcedo[19], Vaughn Shirey[20], Kimberly Vermilya[21], Jalen Winstanley[2], Amy Wu[2], Robert Guralnick[11,22], Akito Y. Kawahara[11,18,22], David J. Lohman[13,17,23], Leslie Ries[20], Edward R. Soucy[24], Roger Vila[25], Nanfang Yu[10] and Naomi E. Pierce[1,2]†

Affiliations:

[1] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA
[2] Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138, USA
[3] Department of Integrative Biology, The University of Texas, Austin, Texas, 78712, USA
[4] School of Biological Sciences, The University of Queensland, Australia, 4072
[5] Institute of Biodiversity, Friedrich Schiller University Jena, Dornburger Straße 159, 07743 Jena, Germany
[6] Helmholtz Centre for Environmental Research - UFZ, Department of Ecosystem Services, Permoserstr. 15, 04318 Leipzig, Germany
[7] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstr. 4, 04103 Leipzig, Germany
[8] Department of Entomology, University of Wisconsin-Madison, WI, 53706, USA
[9] Institut Botànic de Barcelona (IBB, CSIC), Barcelona, Catalonia, Spain
[10] Department of Applied Physics & Applied Math, Columbia University, New York, NY10027
[11] Florida Museum of Natural History, McGuire Center for Lepidoptera and Biodiversity, University of Florida, Gainesville, Florida, 32611
[12] Botany Section, National Museum of Natural History, Manila, Philippines.
[13] Biology Department, City College of New York, City University of New York, USA
[14] Center for Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan 666303, China
[15] University of Chinese Academy of Sciences, Beijing 100049, China
[16] Jose Rizal Memorial State University, Tampilisan Campus, ZNAC, Tampilisan 7116, Zamboanga del Norte, Philippines

[17] PhD Program in Biology, Graduate Center, City University of New York, NY, USA
[18] Entomology and Nematology Department, University of Florida, Gainesville, FL, USA
[19] Biological and Environmental Engineering, Cornell University, Ithaca NY, 14850
[20] Department of Biology, Georgetown University, Washington, DC, USA
[21] Albany Medical College, Albany, New York, USA
[22] Department of Biology, University of Florida, Gainesville, FL, USA
[23] Entomology Section, National Museum of Natural History, Manila, Philippines
[24] Center for Brain Science, Harvard University, 52 Oxford St. Room 331, Cambridge, MA 02138
[25] Institut de Biologia Evolutiva (CSIC-UPF), Barcelona, Catalonia, Spain

*These authors contributed equally to this work

†Correspondence to:

Wei-Ping Chan Email: chanw@g.harvard.edu

Naomi E. Pierce Email: npierce@oeb.harvard.edu

Abstract

Butterfly migration across great distances is testimony to their impressive aeronautical skills[1]. Wing size and shape are two key determinants of flight performance[2], and in butterflies, the specific configuration of forewings to hindwings including the overlapping area between them can have profound effects on overall size and shape. Here we use quantitative morphometrics to analyze the wing size, shape, and flight performance of 190 exemplar species representing a tribal-level phylogeny of Papilionoidea. We identify three morphological 'flight systems' that are characterized by these fore- and hindwing configurations (small [beating], medium, and large [gliding]), and our data suggest that the emergence of these systems coincided with periods of paleoclimatic upheaval. A higher proportion of migratory species have wings characteristic of gliders, and their degree of host specialization appears to be more flexible than that of counterparts with beating wings.

Main text

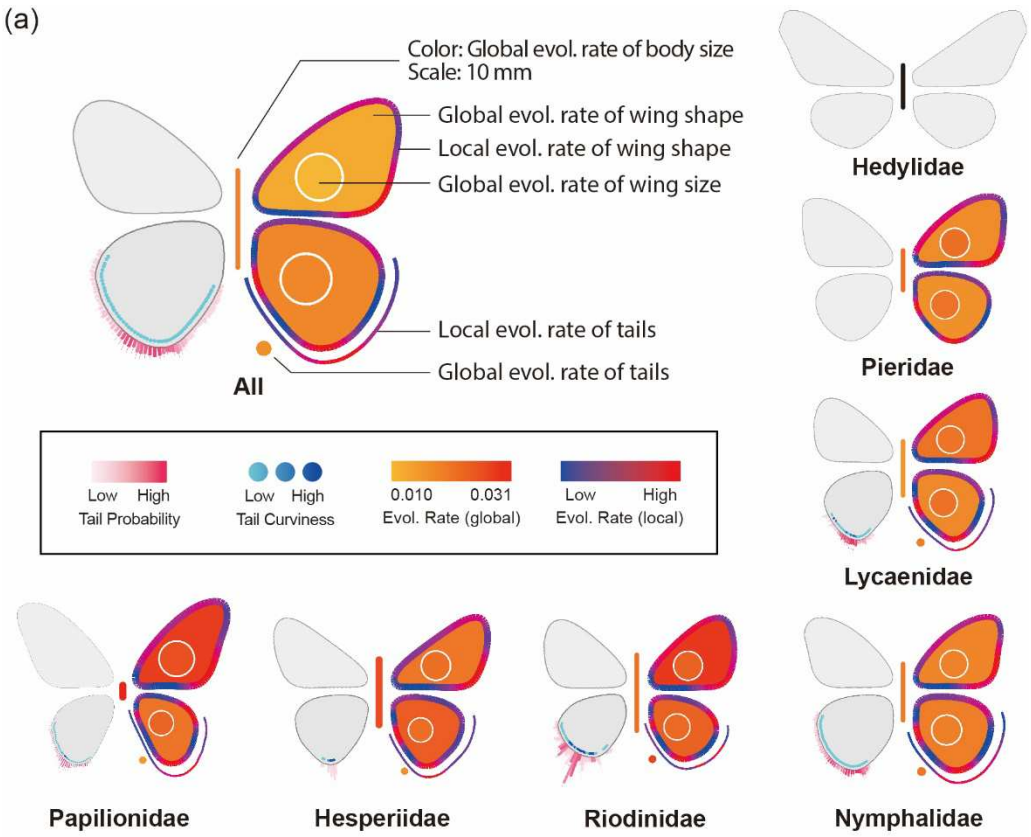
From giant whales in the ocean to tiny insects on land, animals migrate[1,3]. Flight is an energy-effective mode of migration, enabling small-bodied insects, such as butterflies, to complete continental-scale travel. Globally, at least 3% of butterflies show evidence of migratory movements[1], where wing size and shape—two important morphological traits that can be arranged in multiple forewing and hindwing configurations—play crucial roles in determining flight performance[2,4,5]. Flight mechanics of butterflies and other insects with two pairs of thoracic wings can be analyzed using multiple paradigms[6-9]. For example, each lateral forewing-hindwing set can be considered a fused, single wing system[7,10-12]. This simplification facilitates quantifying flight behavior using computational fluid dynamics, setting parameters by carrying out experiments in wind tunnels and testing aeronautic performance using well-defined mathematical models[7,10-19]. Alternatively, the separate actions of forewing and hindwing can be taken into account independently[20-22] and analyzed as a combinatorial, two-wing system that considers traits such as aspect ratio (the ratio of wing length to width), moment of area (a geometrical property of the wing indicating how easily the wing can rotate relative to the body axis), and wing loading (the ratio of body size to wing size), which can be calculated directly from wing morphology[2,23] (see Methods; Extended Data Fig. 1 & 2, Supplementary Discussion).

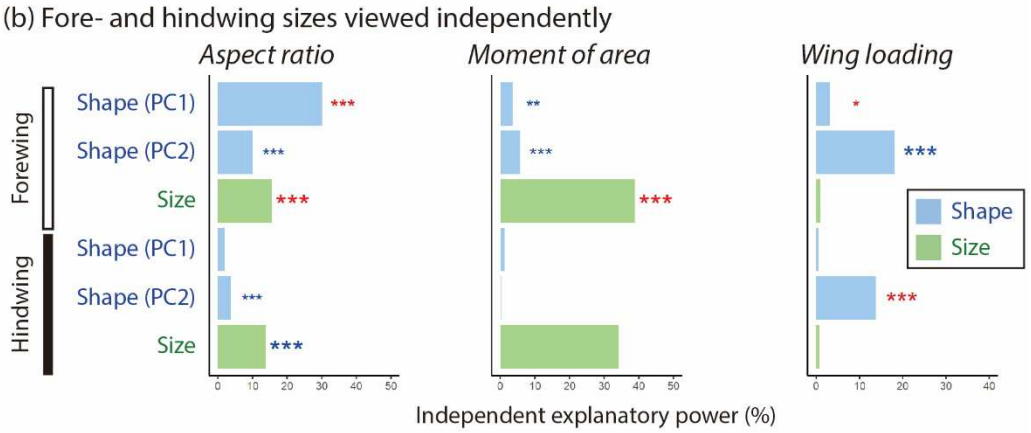
The quantification of shape properties through geometric morphometrics has advanced our understanding of the evolution of wing shape and its implications for flight performance[20-22,24-26]. However, the role of wing size deserves further attention. Apart from aspect ratio, geometric morphometrics have rarely been considered in highly parameterized studies of insect flight[5,7,10-12]. We explore the ways in which absolute and relative wing size can vary in a two-wing system and use this information to identify three characteristic morphological systems employed by most butterflies that we have named according to the dominant morphologies they have: (I) small; (II) medium; and (III) large.

We quantified how wing size and shape are correlated with flight performance (Extended Data Fig. 2) based on fossils and extant representative species included in a tribal level phylogeny[27] representing all families and >90% of the described tribes of butterflies (Methods, Supplementary Data 1 & 2). For each of the 190 species sampled in the phylogeny, we digitized approximately 12 individuals (usually 6 males and 6 females) selected from as many populations across the species' range as possible. The images were then processed by a custom-made, semi-automated pipeline to identify, extract, and characterize wing shape properties[28,29]. We used general Procrustes alignments and elliptical Fourier analyses followed by principal component analyses to summarize the geometric morphometrics of wing shape[21]. Flight characteristics and traits were estimated for each species (Methods and Extended Data Fig. 3). Male and female traits were considered together unless otherwise noted, and evidence for migratory behavior was scored at the species level[1].

Across all butterflies, wing size evolves more slowly than other morphological features including wing shape, body size, and morphology of hindwing tails. Overall, hindwing shape evolves more quickly (σ^2 [evolutionary rate of change] = 0.017, p = 0.20; not significantly different from a Brownian motion model) than hindwing tail morphology (σ^2 = 0.015, p = 0.32), which evolves more quickly than forewing shape (Fig. 1a; σ^2 = 0.012, p = 0.03). Hindwing tail morphologies in males and females both evolve quickly (both σ^2 = 0.017), but the ratio of evolutionary rate (tail/hindwing shape) shows no significant difference over time between sexes (male: 1.01, p = 0.94; female: 1.03, p = 0.80). Evolutionary rates of morphological change also range widely

among different taxonomic families and different body parts. The highest rates of body size change ($\sigma^2 = 0.032$, $p < 0.001$) and forewing shape change ($\sigma^2 = 0.029$, $p < 0.001$) within a clade occur within Papilionidae and Riodinidae, respectively (Fig. 1a). Notably, the leading costal margin and tornal region of the forewing evolve more quickly than other parts of the forewing. Similarly, tails near the anal angle of the hindwing evolve more quickly than other positions on the hindwing (Fig. 1a).


(a)
Color: Global evol. rate of body size
Scale: 10 mm
Global evol. rate of wing shape
Local evol. rate of wing shape
Global evol. rate of wing size
Local evol. rate of tails
Global evol. rate of tails
All
Low High
Tail Probability
Low High
Tail Curviness
0.010 0.031
Evol. Rate (global)
Low High
Evol. Rate (local)
Hedylidae
Pieridae
Lycaenidae
Papilionidae
Hesperiidae
Riodinidae
Nymphalidae


(b) Fore- and hindwing sizes viewed independently
Aspect ratio
Moment of area
Wing loading
Forewing
Hindwing
Shape (PC1)
Shape (PC2)
Size
Shape
Size
Independent explanatory power (%)

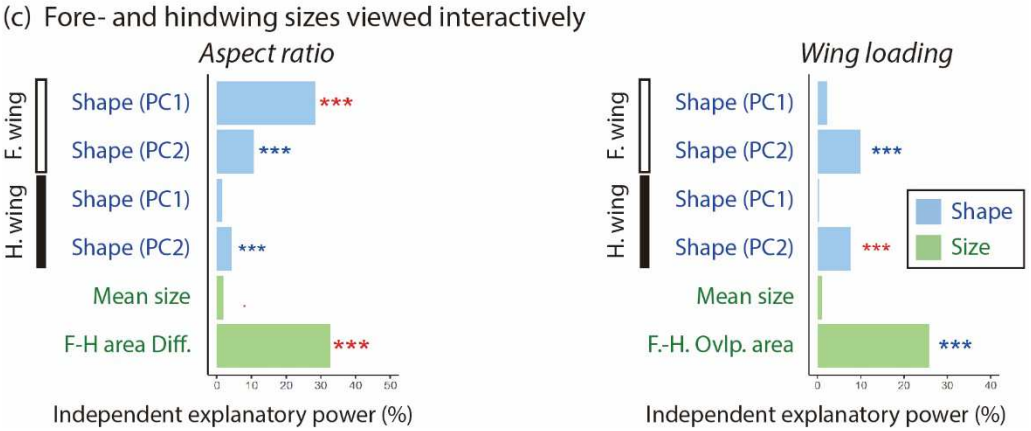

(c) Fore- and hindwing sizes viewed interactively
Aspect ratio
Wing loading
F. wing
H. wing
Shape (PC1)
Shape (PC2)
Mean size
F-H area Diff.
F.-H. Ovlp. area
Shape
Size
Independent explanatory power (%)
Independent explanatory power (%)

Figure 1| Butterfly wing shape and evolutionary rate of morphological change including the relationship between butterfly flight characteristics and morphology in a fused-planform model versus a two-wing model. (a) An example based on all specimens (190 species: 2,388 specimens) showing average wing shape. Left half: average wing shape and tail detail of both sexes (males and females not analyzed separately); right half: global and local evolutionary rates (evol. rate) of shape and size. Global evolutionary rates were derived based on the entire wing shape while local evolutionary rate shows the rate at each point along a wing margin. Wing shape and evolutionary rate of each family except the monotypic Hedylidae (Methods) are shown in the explanatory diagram. (c-d) The independent explanatory power of wing morphology to explain flight characteristics when fore- and hindwing sizes are analyzed independently (c) and integratively (d). Forewing-hindwing size differences are abbreviated as F-H size difference; overlapping area as F-H overlapping area. Asterisks denote significantly positive (red) or negative (blue) beta values under phylogenetic GLMM. Significance levels are labeled: ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

Wing size traits are the most important factors explaining flight performance, as shown by aerodynamic indices with different degrees between fast beating and gliding flight. This may be due to the independent nature of the forewing and the hindwing, which allows for separate adaptations to either wing as well as interactions between them (Extended Data Fig. 1; Supplementary Discussion). By exploring the ways in which overall wing size can vary in a two-wing (fore-and hindwing) system, we observe that wing size plays a significant role that has been frequently overlooked. The key to this is considering the relationship between fore- (F) and hindwing (H) morphology, including the difference between fore- and hindwing areas and the overlapping area between fore- and hindwings, rather than analyzing each of these units separately (Fig. 1b-c and Extended Data Fig. 1; Supplementary Discussion). Forewing shape, which is thought to have a strong effect on butterfly flight performance[29], becomes less important in explaining overall aspect ratio and wing loading (Fig. 1b-c and Extended Data Fig. 2). Moment of area is an integrative parameter accounting for both shape and size parameters (Eq. S3), leading to a tight correlation with size ($r^2 = 0.98$, $p < 0.0001$; Extended Data Fig. 4). The strong correlation between moment of area and wing size as reflected by overall area demonstrates how wing size can act as a primary factor in influencing aerodynamic properties (Extended Data Fig. 2; Supplementary Discussion).

Across the butterfly phylogeny, the evolutionary rates of change in fore- and hindwing size are significantly slower than a random Brownian motion model ($p < 0.001$ and $p < 0.005$ for fore-

158 and hindwing size, respectively; see Methods). This suggests that wing size or area may be more
159 evolutionarily constrained than other traits, such as wing shape (which has an evolutionary rate
160 of change significantly higher than the Brownian motion model [$p < 0.001$] for both fore- and
161 hindwings). However, changes in fore- and hindwing overlapping area rather than overall wing
162 size (σ^2_{diff} [difference from the BM model] = +0.0042, $p < 0.00001$ versus σ^2_{diff} = +0.0014, $p =$
163 0.009, respectively) were more likely to occur, leading to corresponding changes in wing
164 loading. Thus while overall area remains relatively constant, the way in which this area is formed
165 is malleable/"labile" because its rate of change is more rapid than the random model. In addition,
166 the changes in the size of fore- and hindwings can also be easily achieved by random evolution
167 (having significantly lower evolutionary rate than the random model) comparing to other traits,
168 such as wing shape. This is significant because changing fore- and hindwing sizes alters the
169 aspect ratio and wing loading, and in turn, alters beating and gliding abilities. Selection appears
170 to favor changes in the relative sizes of the fore and hind wings over changes in their shapes. We
171 expect to see changes in these traits expressed primarily as changes in relative areas of fore and
172 hind wings rather than relative shape.
173
174 In addition to the central role of wing size in determining flight performance (Fig. 1c), we find
175 that wing shape correlates with wing size-related morphologies (Fig. 2), as was recently
176 demonstrated in Nymphalidae[30] and bombycoid moths[31]. Fore- and hindwing shape are both
177 correlated with wing size; this relationship is more linear for forewing shape (Fig. 2a, c, & e) and
178 indirect and non-linear for hindwing shape (Fig. 2b, d, & f), implying that forewing shape may
179 be the primary trait changing with size, and that hindwing shape may function more as an
180 auxiliary trait interacting with size.
181

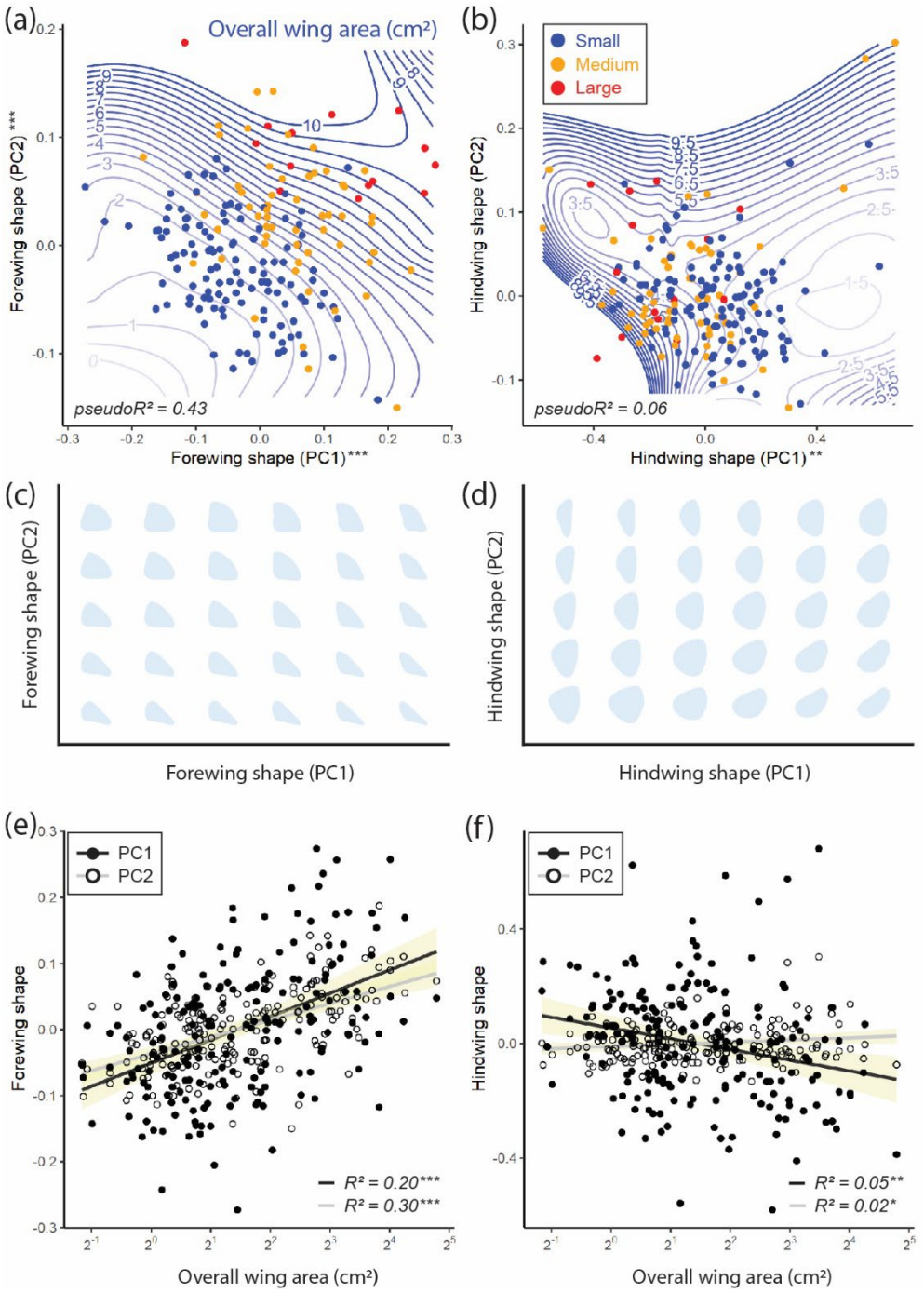
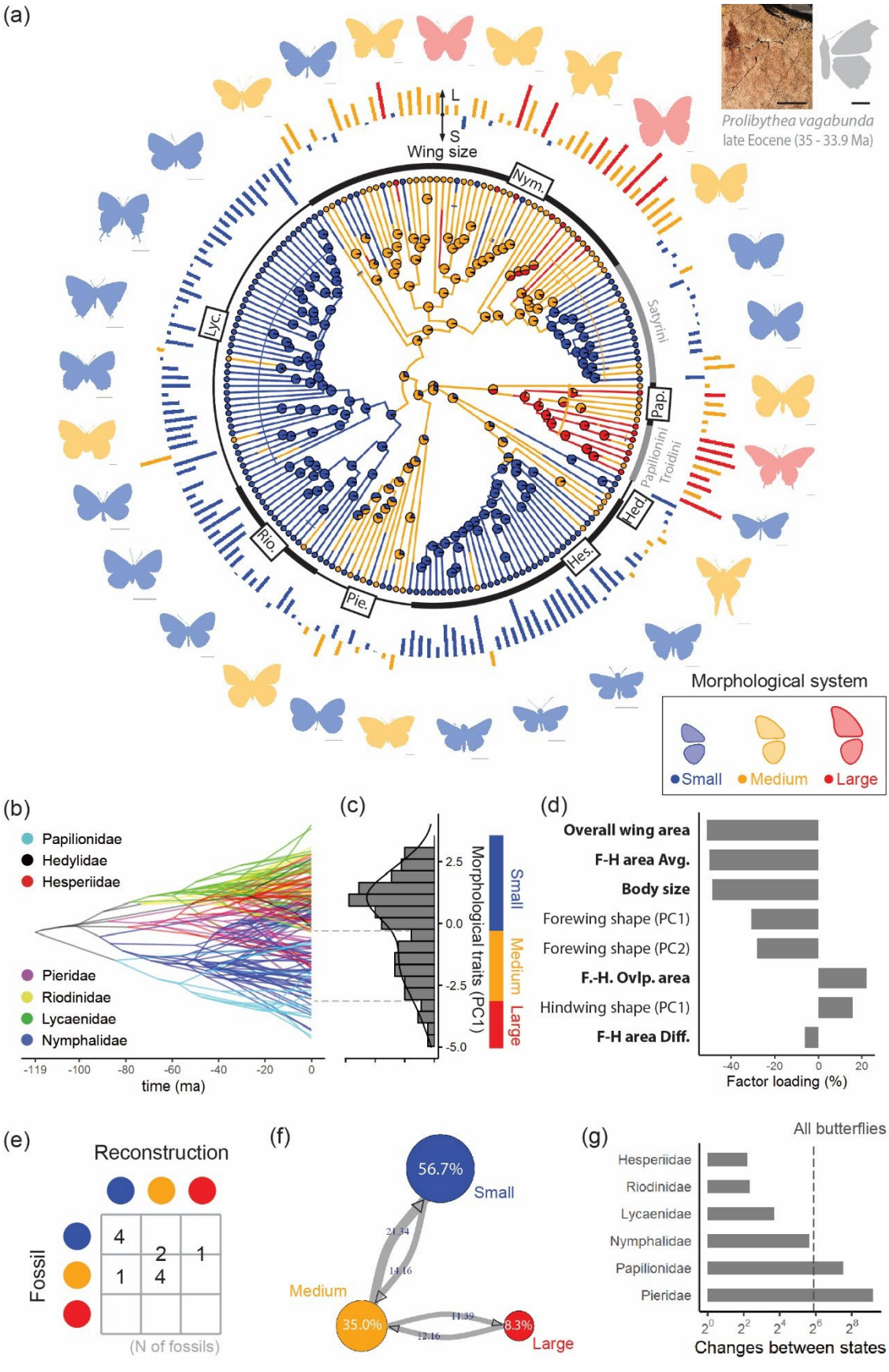


Figure 2| Fore- and hindwing shapes are correlated with wing size. The forewing shape (a) and hindwing shape of butterfly species (b) are plotted in morphological space with contours indicative of overall wing size, and the corresponding shape can be found in (c; forewing) and (d; hindwing). The pseudo R^2 derived from multivariate phylogenetic GLMM is labeled at the lower left of each panel, and the significant levels of $PC1_{shp}$ and $PC2_{shp}$ are labeled with axis titles. Plots are color-coded based on the morphological systems, and the details can be found in Fig. 3. The correlation (univariate GLM) between the first two components ($PC1_{shp}$ & $PC2_{shp}$) of forewing shape (e) and hindwing shape (f) with overall wing size is shown. The shaded area

indicates 95% confidence interval. The statistical results (under GLM) are provided at the lower right of each panel. Significant levels are labeled: ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

To visualize the evolutionary transitions of morphological traits on the butterfly phylogeny, we plotted all morphological traits (shapes and sizes) in morphological space after performing a principal component analysis. With different clustering methods (finite Gaussian mixture modeling, k-means clustering, and density distribution; Methods), roughly three distinct morphological systems can be distinguished along the first principal component (Fig. 3a-c; explaining 42.70% of the variance; Extended Data Fig. 5, Fig. S1 & S2), which is mainly composed of size-related traits (Fig. 3d). We refer to these as small (System S), medium (System M), and large (System L). Given the high variation in size and shape between butterfly species, additional morphological systems could be identified using different approaches, more dimensions, or more species (Methods; Extended Data Fig. 5, Fig. S1 & S2).

Traits in these morphological systems are further translated into aerodynamic indices that we used to derive wing beating and gliding indices (Methods). In the main text, we only show the results based on one of the three clustering methods (others are in the Supplementary Information) because different methods yielded the same conclusions (Extended Data Fig. 5): species falling into System S are generally small and favor wing beating, while species in System L are large and favor gliding (Fig. 4a and Extended Data Fig. 5). Species that are not strong wing beaters or gliders fall into System M and reflect a possible trade-off between beating and gliding abilities on the adaptive landscape, with niche partitioning in flight performance among different families (Fig. 4a; Supplementary Discussion).


(a)
Prolibythea vagabunda
late Eocene (35 - 33.9 Ma)
L
S
Wing size
Nym.
Satyrini
Pap.
Papilionini
Troidini
Hed.
Hes.
Pie.
Rio.
Lyc.
Morphological system
Small
Medium
Large
(b)
Papilionidae
Hedylidae
Hesperiidae
Pieridae
Riodinidae
Lycaenidae
Nymphalidae
-119
-100
-80
-60
-40
-20
0
time (ma)
(c)
2.5
0.0
-2.5
-5.0
Morphological traits (PC1)
Small
Medium
Large
(d)
Overall wing area
F-H area Avg.
Body size
Forewing shape (PC1)
Forewing shape (PC2)
F.-H. Ovlp. area
Hindwing shape (PC1)
F-H area Diff.
-40
-20
0
20
Factor loading (%)
(e)
Reconstruction
Fossil
4
2
1
1
4
(N of fossils)
(f)
56.7%
Small
21.34
14.16
Medium
35.0%
11.39
12.16
8.3%
Large
(g)
All butterflies
Hesperiidae
Riodinidae
Lycaenidae
Nymphalidae
Papilionidae
Pieridae
2^0
2^2
2^4
2^6
2^8
Changes between states

215

Figure 3| Ancestral state reconstruction of butterfly wing morphological systems. (a) The butterfly phylogeny is colored according to the three morphological systems identified in panel **c**. Family names are abbreviated on the inner ring: Papilionidae (Pap); Hedylidae (Hed); Hesperiidae (Hes); Pieridae (Pie); Riodinidae (Rio); Lycaenidae (Lyc); Nymphalidae (Nym). Validated fossil records are labeled as bars segmenting the branch(es) corresponding to their identified taxonomic group(s). The outer ring bar chart shows total wing area (L: large; S: small). Each butterfly silhouette is the color of its morphological system and has a scale bar of 1cm in the lower right corner. An example showing how fossil morphologies were integrated into this study is at the upper right corner (Methods; Extended Data Fig. 7). The average fore- and hindwing shape of each morphological system is shown in the lower left. (b) Traitgram of the evolution of butterfly morphology from the first principal component ($PC1_{morph}$) in morphological space. The outlier in the gliding system is *Historis odius* (Nymphalidae). (c) Histogram of the logical division boundaries (dashed line) for three morphological systems. The probability curve is shown as a solid line. (d) The factor loading of morphologies on $PC1_{morph}$. (e) A table comparing the morphological systems of fossils with ancestral state reconstructions for the nodes where they are placed. A fossil was placed on the boundary of two cells when it has ambiguous morphological system. (f) State changes frequencies among morphological systems. The residence time and state changes are labeled with the vertices and arrows, respectively. (g) Chart showing total changes between states in different taxonomic groups (each family separately, and all butterflies with a dotted line).

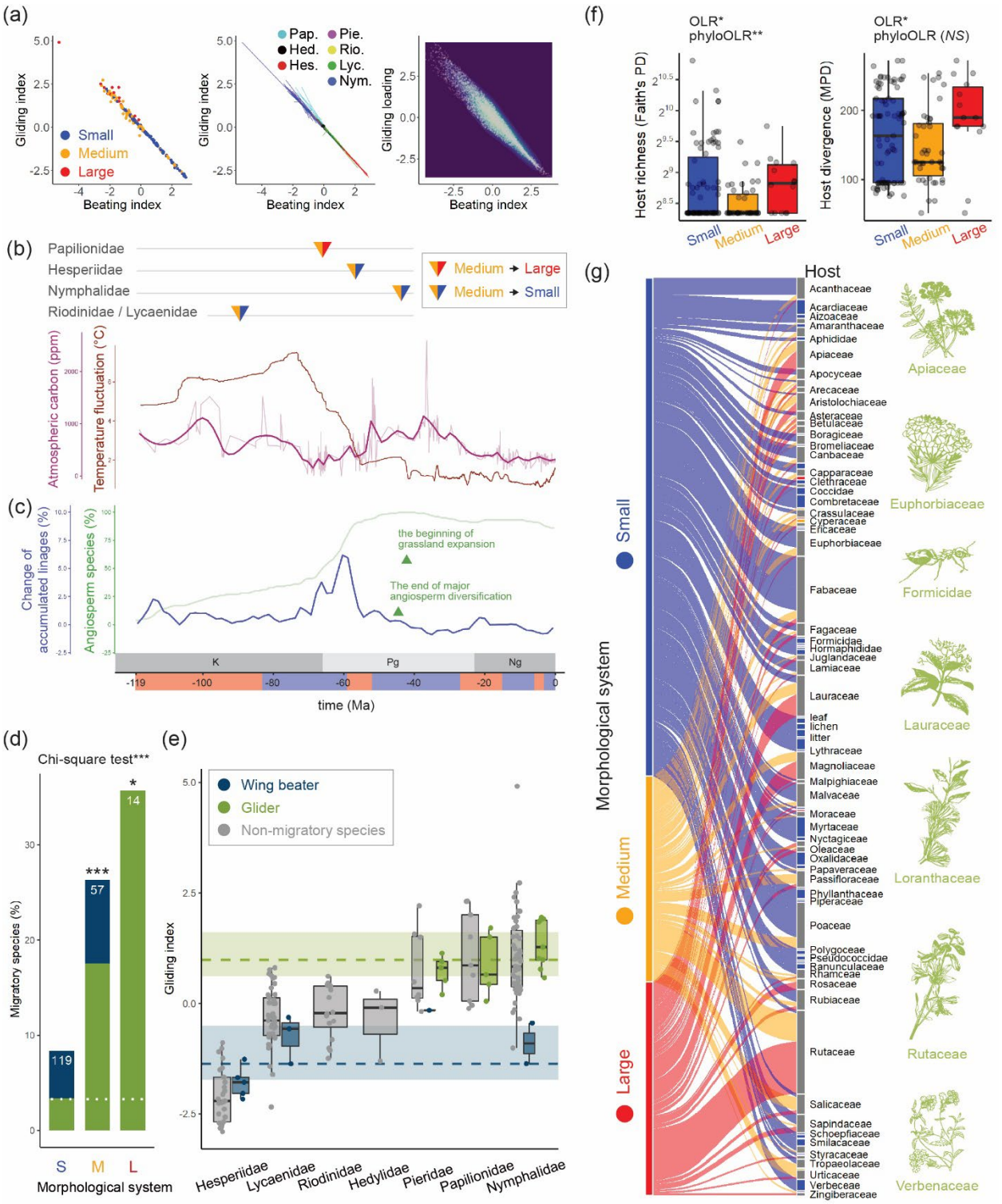


Figure 4| Wing beating and gliding indices in the flight performance landscape and the association with paleoclimate, evolution of angiosperms, butterfly migration, and host richness. (a) Flight performance landscape of all records, colored by morphological system (left panel) and by family in a traitgram (middle panel). A simulation (right panel) based on morphological traits was generated by random selection from the empirical data range. Light

colors indicate greater density of records. Pap, Papilionidae; Hed, Hedylidae; Hes, Hesperiidae; Pie, Pieridae; Rio, Riodinidae; Lyc, Lycaenidae; Nym, Nymphalidae. (b) The inverted triangles share the same x axis with panel (c) and denote the major transitions between morphological systems in each family. The fluctuation of reconstructed temperature and atmospheric carbon in Earth's history[32,33] are in the lower panel. (c) Accumulated angiosperm lineages through time (green line) according to a recent phylogeny[34]. The blue line indicates the change of cumulative angiosperm lineages. The temporal axis indicates the geological time scale (K: Cretaceous, Pg: Paleogene, Ng: Neogene), the bar below indicating expert-identified warming (red) and cooling (blue) phases[35], and major events in Earth's history. (d) Percentage of migratory species within each morphological system. Total number of samples are written within each bar. For each bar, a white dashed line marks the random expected percent and asterisks at the top indicate significant difference compared to random expectation or among all three morphological groups (across bars). (e) The migratory species in each family were categorized as beaters, gliders, or neither, and plotted as boxplots across families, showing the duo-affinity zone—migratory beaters and gliders. (d-e) share the same color legends. (f) Two indices of host plant richness and divergence (Methods; Extended Data Fig. 10) were mapped onto the morphological systems: the sum of phylogenetic branch lengths (Faith's PD) and the mean pairwise distance (MPD). Asterisks between bars show the level of significant difference between that pair of morphological systems. The result of classic ordinal logistic regression (OLR) and phylogenetic OLR are both labeled on the top of each panel. (d, f) Significance levels are labeled: ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; ., $p < 0.1$; NS, $p > 0.1$. (g) Details of host information are shown. Illustrations of exemplar hosts are provided at the right.

The ancestral state reconstruction from the continuous morphological value (Fig. 3b; Methods) suggests that the ancestor of butterflies likely had small to medium size. To better understand their evolutionary transitions, we also ran a reconstruction based on the discrete morphological systems (Fig. 3a) and showed that the ancestor of butterflies was in System M (the probability is 74% based on the all-rates-different model (ARD); Methods). It evolved to System S (good wing beaters) at least twice during the early and late stages of butterfly evolution (early stage: ~89.5-93 Ma, late Cretaceous; late stage: ~44-57 Ma, early Paleogene); the lineages Riodinidae and Lycaenidae formed during the earlier stage (Fig. 3a & 4b). The late transition occurred early in the evolution of Hesperiidae (~57 Ma, early Paleogene) and late in Nymphalidae (~44 Ma, mid Eocene). The only transition from System M to System L (good gliders) occurred in Papilionidae (~66-72 Ma, Cretaceous-Paleogene boundary). By tallying state switches between wing morphologies, we see that System S, having better wing beating ability, has the longest evolutionary residence time (Fig. 3f), and switches from System S to System M are most frequent with no direct switches between System S and L (Fig. 3f). Though the evolution of morphological systems is correlated with familial differences and clustering methods, it also

occurs within families (Fig. 3a & g and Extended Data Fig. 6a-b). For example, frequent transitions within the Pieridae can be observed in all clustering methods (Fig. 3g, Fig. S1a & g, and Fig. S2a & g); this is possibly related to the relatively high percentage of migratory species found in this group (Fig. 4e; see discussion below).

We compared ancestral state reconstruction results to those of morphological data quantified from butterfly fossils where the body, both fore- and hindwings, and a likely taxonomic affinity were available (Fig. 3a, Extended Data Fig. 7-9, and Supplementary Data 3). Measurements taken from 10 out of 12 fossil specimens fall within the morphological systems inferred from ancestral state reconstructions based on extant species (Fig. 3a & e). One fossil not matching predicted morphological systems is in the Nymphalidae (it has System M, although System S was suggested by the reconstruction; fossil IDs: F-12; *Barbarothea florissanti*). Three fossils have ambiguous morphological systems: two in Nymphalidae (having the system between S and M, and System M was suggested by the reconstruction; fossil IDs: F-16; *Jupiteria charon*; fossil IDs: F-19; *Prodryas persephone*) and one in Papilionidae (having the system between S and M yet System L was suggested by the reconstruction; fossil ID: F-24; *Doritites bosniaskii*; Extended Data Table 1). Though many factors could have contributed to this discrepancy, including clustering methods, incorrect fossil dating and identification, as discussed above, intra-tribe morphological variation is likely to be greater than captured by our dataset of exemplars with the possibility that key species are extinct, and better predictions necessitate more comprehensive sampling.

We noticed a potential association between paleoclimatic variability and the emergence of butterfly wing morphological systems (Fig. 4b). The early transition of System M seems to have mainly evolved during periods of large temperature fluctuation at annual to millennial timescales[36]: late Cretaceous (~89.5-93 Ma). System L emerges near the K-Pg boundary (~66 Ma), another period of drastic climatic disruption[37]. Given that many studies have shown the changing wing morphologies of Lepidoptera species along environmental gradients[38-40], implying a relatively diverse evolution of butterfly wings under changing environments, we suspect that the evolution of improved wing beating and gliding ability facilitated local species adaptation under fluctuating climatic conditions[41,42]. Although no striking climatic association

312 corresponds to the transition from System M to System S observed in the satyrines in the mid-
313 Eocene (~44 Ma; Satyrinae, Nymphalidae), this transition coincides with the beginning of
314 grassland expansion[43], a period that is also associated with aridification in North America and
315 Africa[43] (Fig. 4c). Most species in Satyrinae have herbaceous host plants and the majority are
316 grass (Poaceae) feeders[44]. Intriguingly, relative changes in paleo-atmospheric carbon, an index of
317 air density or viscosity, did not show potential association with the transitions of morphological
318 flight systems (Fig. 4b).
319
320 At least 568 species of butterflies have been recorded to be migratory[1], although this estimate
321 varies depending on the definition of migration (Methods). Based on an independently compiled
322 list[1], we found that most migratory species fall into System L (having relatively good gliding
323 ability, 35.7%), followed by System M (26.3%), and finally by System S (enhanced wing beating
324 ability, 8.4%) (test of differences among systems: $\chi^2 = 13.81$, $p < 0.001$; Fig. 4d-e). Migratory
325 species can be grouped into good gliders or good wing beaters based on their flight indices, with
326 the groups forming two attractive zones across families (Fig. 4e). The percentage of migratory
327 gliders also increases from System S to System L (Fig. 4d). Our result supports the idea that
328 good gliders are also good migrants[41,45] (Fig. 4d). Two families, Pieridae and Nymphalidae,
329 stand out for containing species in both migratory wing beater and glider groups, which may be
330 related to the frequent switches between morphological states observed in these two groups (Fig.
331 3g, S1g, and S2g). Although the proportion of migratory species within each family varies
332 considerably from one geographic region to another, within the Pieridae, the Coliadinae has the
333 highest proportion of migratory species among all butterfly subfamilies (>25%), with the
334 Pierinae ranking roughly the same as Nymphalinae (>10%), and the Coliadinae are followed by
335 high proportions in three subfamilies of Nymphalidae; the Danainae (>16%); Nymphalinae
336 (>10%); and Heliconiinae (>10%), respectively[1].
337
338 Despite the migrants having striking bimodality in flight performance, we did not detect
339 significant variation in range size or seasonal switching (seasonal variation in habitat
340 suitability[46]) between migratory wing beaters and gliders ($p > 0.5$; Methods). Estimating host
341 plant richness for each tribe based on six different phylogenetic metrics[47] led to varying results
342 depending on the calculation of the sum of branch lengths (Faith's PD), mean pairwise distance

(MPD), mean nearest taxon distance (MNTD), and the distance-based specialization index (DSI) (Fig. 4f-g, Extended Data Fig. 10, and Methods), also we observed a pattern of significantly higher host diversity (Faith's PD) in System S with better wing beating performance (Fig. 4f). However, we also found significantly higher host divergence (MPD) in System L with better gliding performance. These results jointly suggest that the hosts of wing beaters and gliders are diverse when considered at fine and coarse taxonomic levels, respectively, resembling alpha and beta biodiversity, but in a phylogenetic context. In other words, the host specialization of wing beaters (System S) is high at finer taxonomic levels, and while gliders' (System L) hosts are more likely to share common evolutionary histories, their range is nevertheless wider (the mean number of host families per species is 0.90 in System S versus 2.5 in System L). This would have provided gliders greater flexibility in host choice. Differences between species belonging to different morphological systems are not significant in MNTD, which is more sensitive to tip divergence[47], and in distance-based specialization indices (DSI_{MPD} and DSI_{MNTD}; Extended Data Fig. 10). The ecological context (*e.g.*, geographic and climatic factors) behind these different migratory strategies merits further investigation, but our results suggest a more labile relationship between flight and host plant specialization in gliders than in wing beaters.

Our evidence shows high PD in System S [beating] and high MPD in System L [gliding], leading to a question regarding the difference between these two metrics: Faith's PD (total phylogenetic branch length) and mean pairwise distance (MPD). Although both these indices can represent host diversity, PD is more closely reflects overall diversity, whereas MPD also reflects the diversity of clustered linages. High PD and low MPD both indicate high host phylogenetic diversity, but these hosts are more evenly distributed without clustering on the host phylogeny. Low PD and high MPD indicate relatively low host phylogenetic diversity, but these hosts are clustered into a few distant clades. According to our results, System S had the highest PD (but not the lowest MPD) and System L has the highest MPD (but not the lowest PD). Therefore, butterflies in System S feed on diverse hosts with low clustering on the plant phylogeny. Butterflies in System L feed on a few distinct host clades/lineages.

By integrating a set of approaches used by ecologists, evolutionary biologists, and engineers, we demonstrate that wing size—especially the overall wing size, fore- and hindwing size difference,

and fore- and hindwing overlapping area—is a key determinant of butterfly flight performance. Modeling butterfly wings as a two-wing system, where butterflies can change their effective overall wing sizes and flight characteristics by altering the size difference and overlapping area between two wings, demonstrates that most butterflies optimize either wing beating or gliding flight. At least three morphological systems can be identified from the data of tribal-level exemplars sampled for this study, and transitions between systems appear to evolve during periods of high paleoclimatic variability. A majority of migratory species fall into System L, a group characterized by strong gliding ability, but a wing beating strategy is also found for some migrants. Based on morphological evidence from contemporary species and fossils, this study provides phylogeneically-informed species-trait analyses that illustrate a strong link between the evolution of butterfly morphology and flight performance.

Methods

Specimen selection and imaging

Up to six male (6.47 ± 2.07; mean ± S.D.) and six female (5.62 ± 1.86) specimens were sampled for each tip of a tribal level phylogeny[27] (total 190 out of 195 tips representing > 90% of butterfly tribes). If specimens were insufficient for a particular species (fewer than two specimens), another species in the same genus was analyzed (affecting 70 species). A total of 2,388 specimens were examined. A custom-built, high-throughput rig was designed to image the dorsal and ventral sides of each specimen[28]. Potential sampling bias is likely to have occurred when exemplars were selected to represent tribes during phylogeny construction[27]. Although efforts were made to select the type species for each of the genera sampled, if the type species was difficult to obtain, it was replaced with another close relative. It seems likely that more commonly observed species are more likely to be collected and used in inferring phylogeny. Interestingly, these same species are disproportionately composed of migrants, as reflected by the large number of migrants included in the sampled exemplar taxa[48]. Comparing to the percentage of migrants in the studied butterfly species (568/17,296; 3.28%)[1], a significantly higher percentage of migrants can be found in System M and L (Fig. 4d).

Male-female differences

To match the morphology dataset with the migration dataset (species-level records), we initially grouped data from male and female specimens together to generate a species-level morphology dataset for our analyses. Nevertheless, butterflies and moths are strongly sexually dimorphic in many traits (discussed more specifically for this dataset in a separate manuscript, Rabideau Childers, Chan *et al. in review*), and future research will tease apart the effects that these sex-specific differences might have on the flight performance of males versus females.

Image analysis

Wing morphology data from individual specimen images were automatically extracted and run through a pipeline with built-in optional supervision and manual correction steps that performed background removal, fore- and hindwing segmentation, and full wing shape reconstruction[28]. We used the averaged geometric morphometrics of each species (when applicable) in analyses. To reconcile methodologies of two different approaches—one treating wing size as centroid size and independent from shape[20,26], and the other rarely using centroid size[7,10,11]—we used only centroid size in the estimation of evolutionary rates for wing shape and size. This allowed us to distinguish their effects independently[20,26]. Centroid size was derived after the general Procrustes alignments with the R package *Momocs*[49]. Key flight traits including aspect ratio, moment of area, and overall wing area were measured for different wing arrangements using different wing angles. The arrangement with the greatest aspect ratio (and wingspan) was considered the optimal form for flight (i.e. optimal planform) [16], and the corresponding parameters were retained (Extended Data Fig. 3). This method compares butterflies at their best gliding form without further taking into consideration that they can rotate their wings dynamically to switch between optimal gliding and beating forms during their flight. The overlapping area was calculated as the difference between the sum of fore- and hindwing area and the overall wing area (the wing area with the optimal planform). To derive wing loading, body weight was approximated using body size, which was estimated as an ellipsoid based on the length and width of the body[28]. For traits that could be decomposed into multiple dimensions such as wing shapes and tails, the first ($PC1_{shp}$) and second ($PC2_{shp}$) principal components were used as quantifications that jointly explain the majority ($\geq 80\%$) of overall variance. All pipeline coding was done in *Matlab 2018b*.

Connections between morphologies, flight characteristics and aerodynamic indices

Instead of using computational fluid dynamics to simulate butterfly flight, we used aerodynamic indices. Although these are more tractable for large-scale studies, like ours, that include more than 2,000 specimens, it represents a trade-off with respect to other aerodynamic details. Two sets of different but translatable variables are commonly used to describe wing beating mechanisms[7,10]. We use those of Chin and Lentik[7], focusing on Reynolds number (Re), Rossby number (Ro), and advance ratio (J). For characterizing gliding flight, two indices were selected: maximum lift-to-drag ratio (LD_{max}) and glide ratio (Gr; numerically equal to lift-to-drag ratio under constant flight speed). We further condensed these aerodynamic metrics into indices for beating and gliding ability on the flight performance landscape to elucidate the flight performance of butterflies (Supplementary Methods; Extended Data Fig. 2). The three flight characteristics—aspect ratio, wing loading, and moment of area—were embedded into the formulae of our aerodynamic indices (Methods and Supplementary Discussion), and a hierarchical structure based on the mathematical and statistical correlation of these parameters was constructed (Extended Data Fig. 2).

Summary of variables (Extended Data Fig. 2)

Morphology:

Shape: geometric morphometrics of fore- and hindwings, and tail morphology

Size: overall wing area, forewing area, hindwing area, difference between fore- and hindwing areas (F-H area difference or F-H area Diff.), overlapping area between fore- and hindwings (F-H overlapping area or F-H Ovlp. area), forewing centroid size, and hindwing centroid size.

Flight characteristics (for entire butterfly):

aspect ratio, wing loading (mathematically linked to relative wing size; Supplementary Discussion), and moment of area

Aerodynamic indices:

Reynolds number (Re), Rossby number (Ro), advance ratio (J), maximum lift-to-drag ratio (LD_{max}), and glide ratio (Gr)

Key flight traits:

aspect ratio, wing loading, and overall wing area

Macroevolutionary analysis

We calculated the global (considering the wing as a whole) and local (considering different locations on a wing silhouette) evolutionary rates of change (σ^2) for each trait (size and shape). Global evolutionary rates were computed with the first two principal components (which explained over 80% of variation for both fore- and hindwings) in the R package *Geomorph*[50]. Local evolutionary rates were calculated with the geometric morphometrics. By applying the concept of stationary bootstrap[51], a random length of wing-edge segments was drawn based on a geometric distribution at a random position on a wing shape template (128 anchor points in a circular format). For a wing piece of interest within a group (*e.g.*, the forewing of Lycaenidae), a total of 1,000 segments were drawn, and the wing shape of all species in this group were extracted according to each wing-edge segment for their evolutionary rate estimation. The averaged local evolutionary rate was derived for each position. All macroevolutionary analyses described above were done by applying the functions in the R package *geomorph*[50]. These macro-evolutionary analyses were applied to all butterfly families except Hedylidae, which contains the single genus *Macrosoma*[52] comprised of 36 species.

The explanatory power of different morphologies to describe flight characteristics was derived from hierarchical partitioning, which can disentangle the explanatory power among highly collinear variables[53]. A phylogenetic general linear mixed model (phyloGLMM) was also used to indicate the significance of each variable once phylogeny was taken into account; Gaussian distribution and maximal likelihood (ML) were applied (bayes = FALSE, REML = FALSE). Hierarchical partitioning was done in R package *relaimpo*[54], and phyloGLMM was done in package *phyr*[55].

Determination of morphological flight system

Roughly three morphological systems were identified based on the different clustering methods: finite Gaussian mixture modeling[56], k-means clustering[57], and density distribution. For finite Gaussian mixture modeling, three to five states were identified in multivariate morphological space (5-state model [BIC = -1262.1]; 3-state model [BIC = -1332.7]; 3-state model [BIC = -1334.9]), which is composed of overall wing area, the average of fore- and hindwings (F-H area

Avg.), body size, forewing shape ($PC1_{shp}$ and $PC2_{shp}$), hindwing shape ($PC1_{shp}$ and $PC2_{shp}$), the difference between fore- and hindwing areas (F-H area Diff.), overlapping area between fore- and hindwings (F-H Ovlp. area) (Fig. 3d). Two to three states were identified when considering only the first principal component ($PC1_{morph}$) and in the bivariate setting with both $PC1_{morph}$ and $PC2_{morph}$. Given that the models based on only $PC1_{morph}$ show the best model fits among all settings (BIC = -774.6, -779.3 and -785.2 for the top three models), we chose the 3-state model based only on $PC1_{morph}$ (Fig. S1). For k-means clustering, two states were identified in multivariate morphological space based on the elbow and silhouette method[57] (Fig. S2). By integrating the results generated from the two methods above and the density distribution along $PC1_{morph}$ (Fig. 3c), we decided to show the three states (morphological system Small, Medium, and Large) in the main text and used them throughout this paper to discuss their evolution in relation to flight performance (others can be found in Supplementary Information). The model-based clustering and density estimation with Gaussian mixture modelling were done in the R package *mclust*[56]. The elbow and silhouette methods were done in the R package *factoextra*[57].

Ancestral state reconstruction

Before we ran the ancestral states reconstruction, multiple settings (equal rate [ER], symmetric rates [SYS], and all rates differ [ARD]) with and without a hidden-state detection were tested in the R package *corHMM*[58] (Extended Data Table 2). We concluded that a hidden state is unlikely, and the best setting (ARD) was selected to generate the final figure. Ancestral states were reconstructed using maximum likelihood, and changes between character states were done by stochastic mapping based on the function *fastAnc* in the R package *phytools*[59]. The principal component eigenvectors change when additional data are integrated into a single model. To preserve the original results based on the phylogeny, we ran another independent PCA for the integrated dataset of both specimens and fossils. Thus, the principal components (PC) from the analyses with fossil data differ from those in the main analyses. To derive the morphological system of each fossil, we plotted all specimens and fossils together in morphological space. Along with the first principal component ($PC1_{morph}$), three different morphological systems were identified and assigned to each fossil (Extended Data Fig. 9, Fig. S3 & S4).

Paleoclimate and angiosperm phylogeny analysis

Temperature was estimated from stable oxygen isotopes in fossils of planktonic Foraminifera[60] based on a method developed by Condamine *et al.* [32]. Temperature fluctuation was further calculated as local standard variation. Carbon concentrations were obtained from Fletcher *et al.* [33], and phylogeny of angiosperms was adapted from Condamine *et al.* [34] with further calculations on accumulated lineages and the change of accumulated lineages.

Migratory species analysis

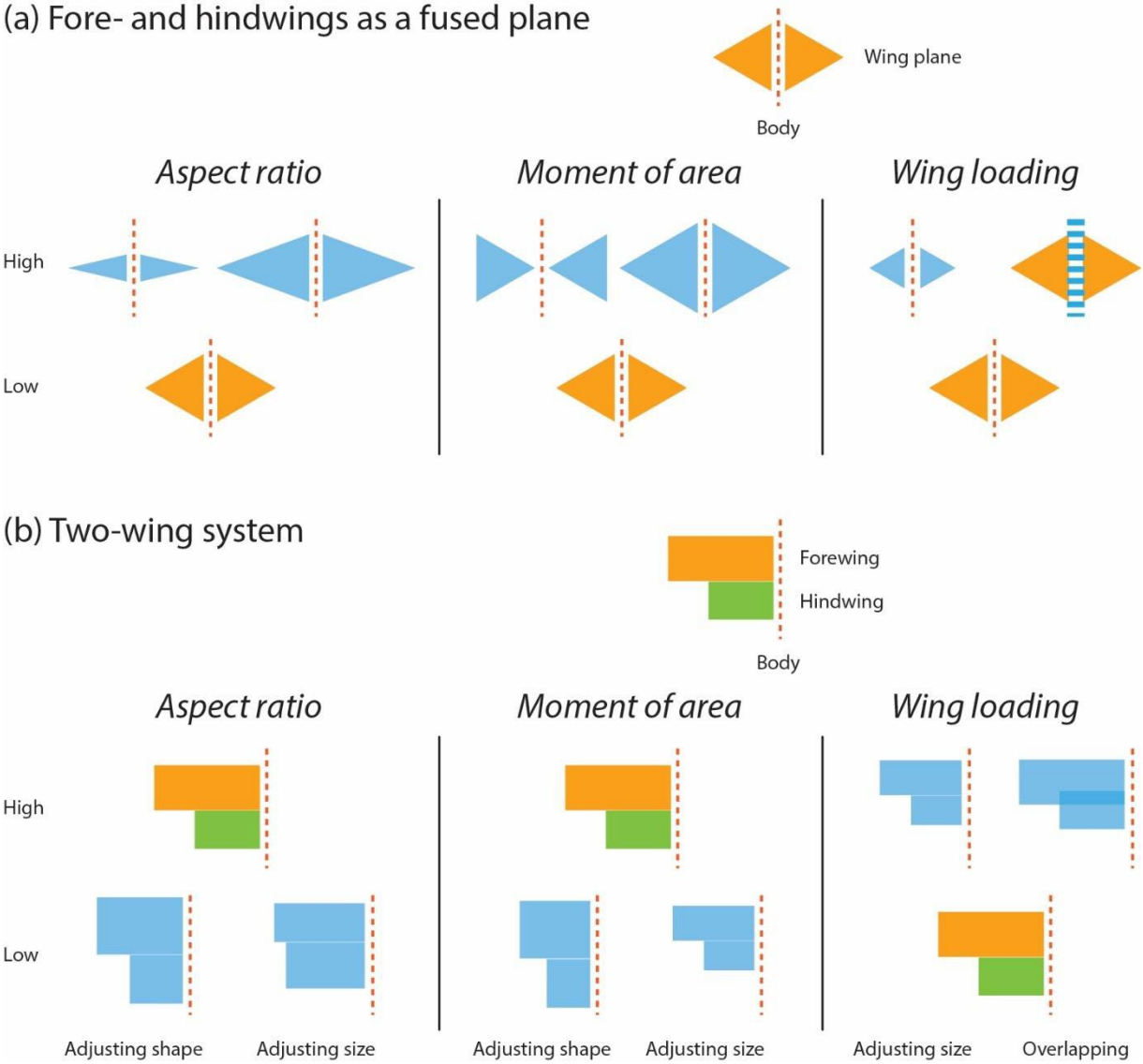
We compared our analysis of wing trait data gathered from exemplar species included in the tribal level phylogeny of Espeland *et al.* [27] against a recently published list of migratory butterfly species by Chowdhury *et al.* [1], including estimates of range size and seasonal switching (seasonal changes in habitat suitability[46]). Migratory behaviors were quantified based on experimental, physiological, or population evidence[1] including: Mark–release–recapture; trap catches; tethered flight measurements; amplified fragment length polymorphism; radar observations; flight in beneficial compass direction; stable isotopes; maintaining constant compass course by using wind drift compensation; larvae incapable of development in unfavorably seasonal climates, extended pre–oviposition period; larval host plant absent; and breeding population not established. Note that we adapted this independently prepared and published dataset directly because the discussion of the definition of migration is beyond the scope of this study. Migratory behaviors are not necessarily shared by closely related species, including those within a tribe[1], so only exact matches were included (Supplementary Data 1). No interpolation or extrapolation was applied. Differences in the percentage of migratory species in each morphological system were compared with random percentages (the number of migratory species compared to the number of overall species surveyed[1]) using a Chi-square test which cannot take into account of phylogeny. ANOVA followed by *post-hoc* Student's t-test with and without the control of phylogenetic signals were used to quantify the differences, such as range size and seasonal switching, between morphological systems. The phylogenetic version of ANOVA and *post-hoc* t-test were done by the function *phylANOVA* in the R package *Phytools*[59].

Host plant analysis

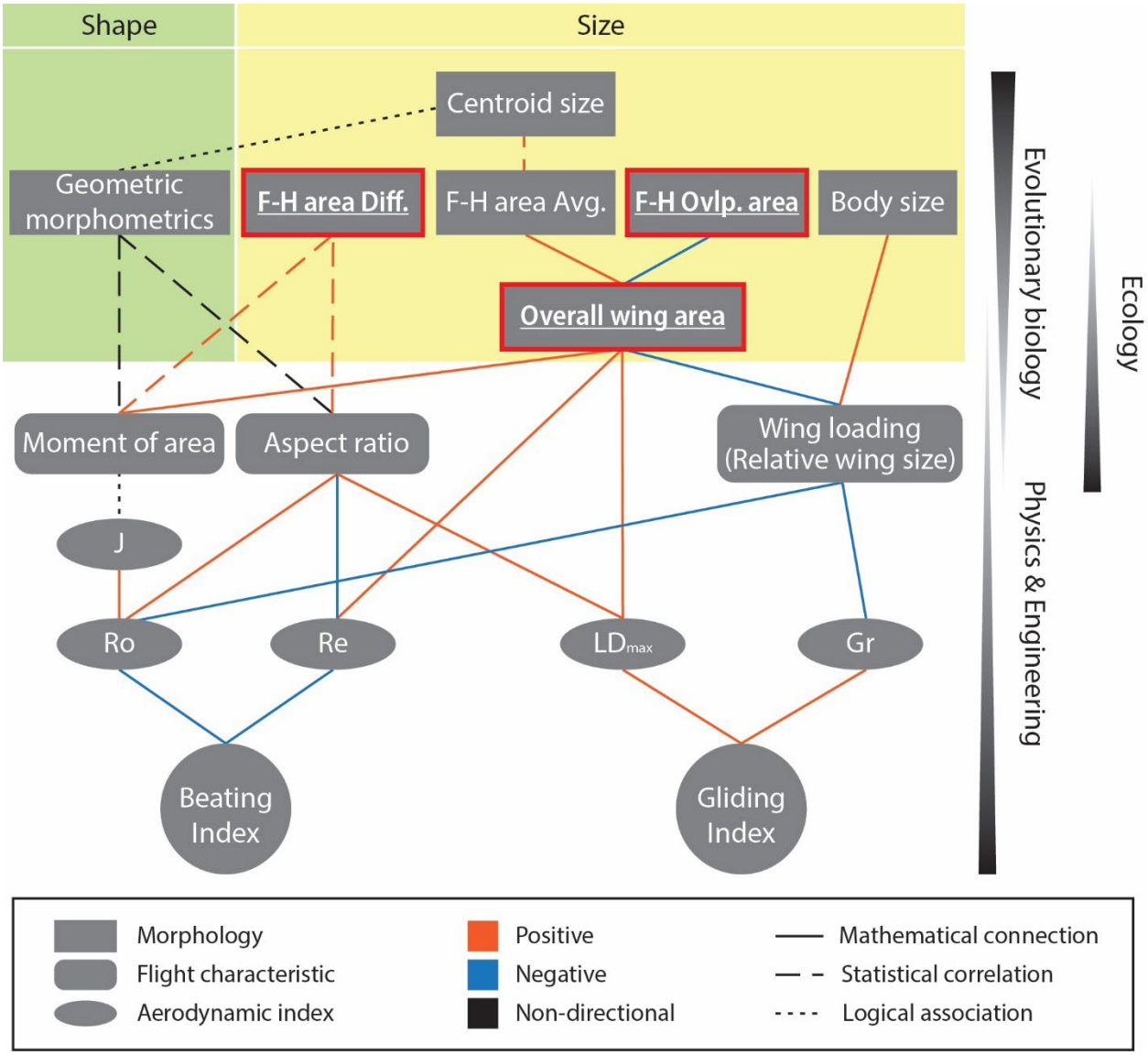
For each butterfly species that feeds on seed plants, we quantified host plant species-richness and divergence based on a host plant dataset[61] with six phylogenetic metrics implemented in the R

558 package *picante*[47]. The taxon sampling of Espeland *et al.* [27] was designed to sample one species
559 per tribe across Papilionoidea. To better reflect dietary breadth within each tribe, we
560 consolidated host plant information from one species per genus for each tribe included in
561 Espeland *et al.* [27]. We derived host plant breadth from the raw number of host plant families and
562 Faith's PD, the sum of phylogenetic branch lengths[62]. To quantify host plant divergence, we
563 calculated mean pairwise distance (MPD), mean nearest taxon distance (MNTD), and the
564 distance-based specialization index (DSI). MPD is the average distance separating all pairs of
565 species on the phylogenetic tree[63]. MNTD is the mean distance between each species and its
566 closest relative[63]. DSI is a Z-score of MPD or MNTD that measures specialization as a deviation
567 from a random expectation[64]. A calibrated phylogeny of seed plants[65] was used as reference. We
568 pruned the tree to include two species per plant family with a divergence time representing the
569 crown group age for each family. This exercise enabled the calculation of pairwise distance-
570 based metrics such as MPD, MNTD, and DSI for monophagous butterflies that feed on only one
571 plant family. Finally, the functions *pd*, *ses.mpd*, and *ses.mntd* were used to calculate these
572 phylogenetic metrics using 1,000 replicates in the R package *picante*[47]. Ordinal logistic
573 regression (OLR) was used to identify correlations between morphological systems and host
574 diversity/divergence/specialization. The phylogenetic version of OLR was done by the R
575 package *Phytools*[59].
576
577 **Figures (embedded in the text)**

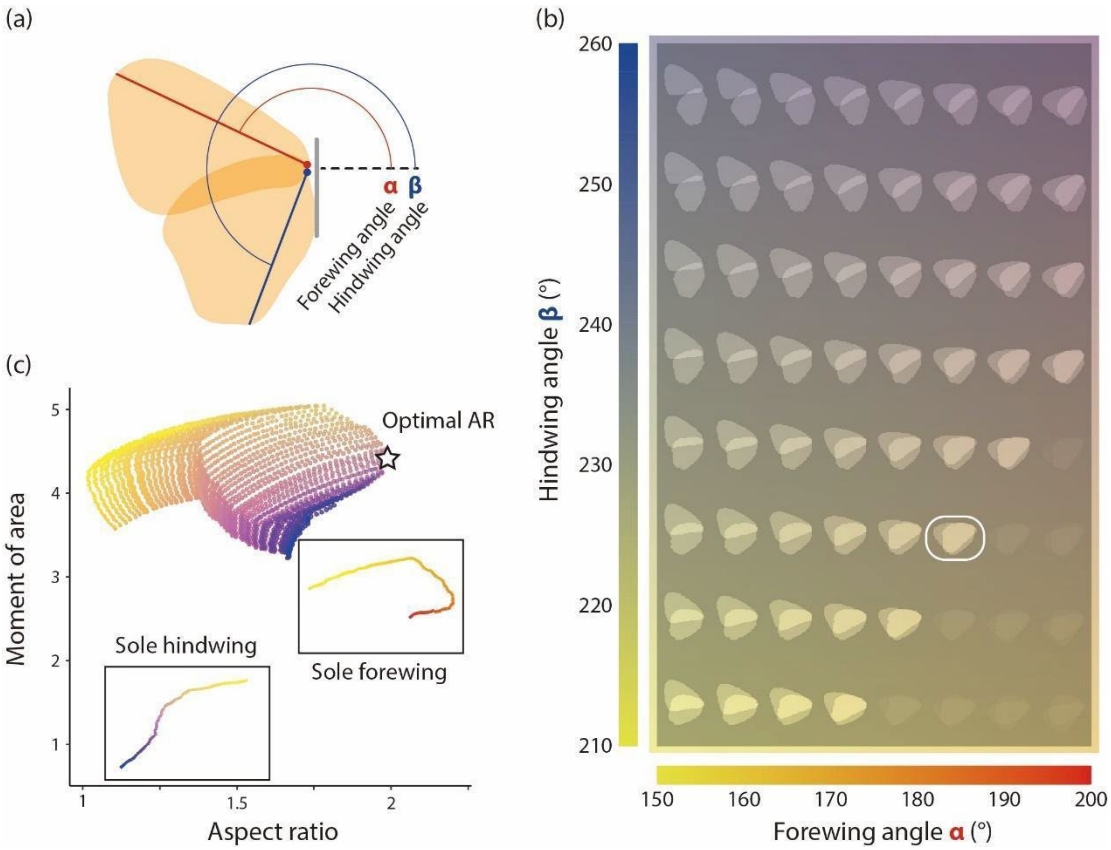
Extended data figures



Extended Data Figure 1| The relationship between butterfly flight characteristics and morphology in a fused-planform system versus a two-wing system. (a) In highly parameterized studies (*e.g.*, applied physics), butterfly wings are frequently modeled as a single, fused planform[11,12,14,15]. A vertical comparison showing a low and high value of three different flight characteristics. (b) A conceptual diagram showing the changing morphology and flight characteristics when butterfly wings are modeled as a two-wing system. A vertical comparison showing low and high values of three flight characteristics. (a-b) The blue area represents how a given flight characteristic can be shifted through changes to shape or size. Thus, flight characteristics can shift via evolutionary changes to wing size with or without changes in shape (Supplementary Discussion).

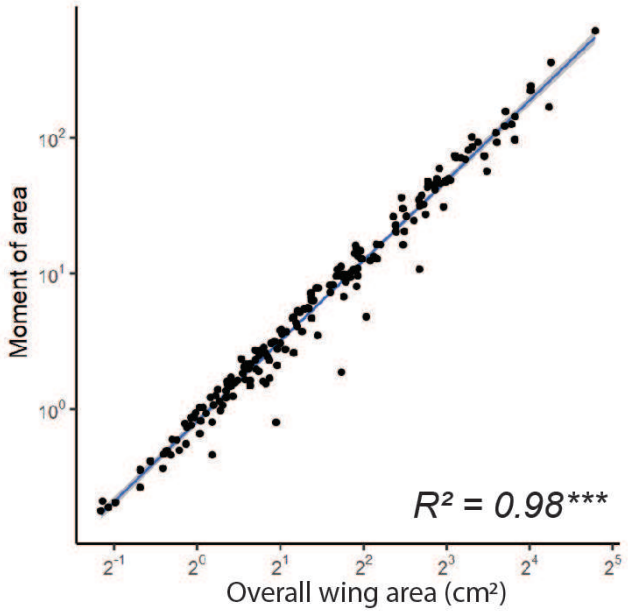


Extended Data Figure 2| The map of variable connections among different fields. Those morphologies identified and/or emphasized in this study are underscored and labeled in red. The frequency of variable applications is indicated at the right based on different fields. Abbreviations are shown for convenience: the difference between fore- and hindwing areas (F-H area Diff.), overlapping area between fore- and hindwings (F-H Ovlp. area); Reynolds number (Re), Rossby number (Ro), advance ratio (J), maximum lift-to-drag ratio (LD_{max}), and glide ratio (Gr).

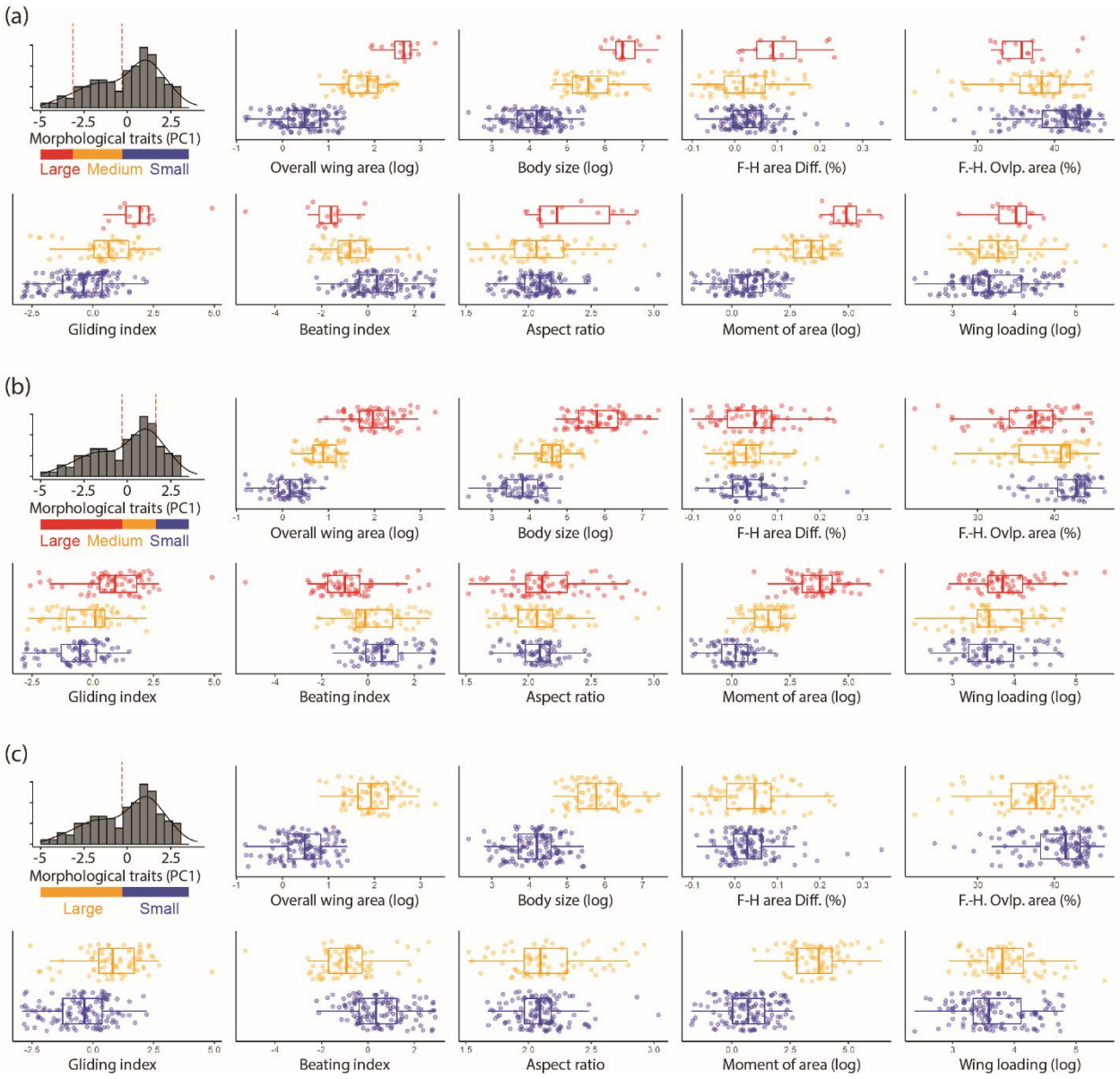


600

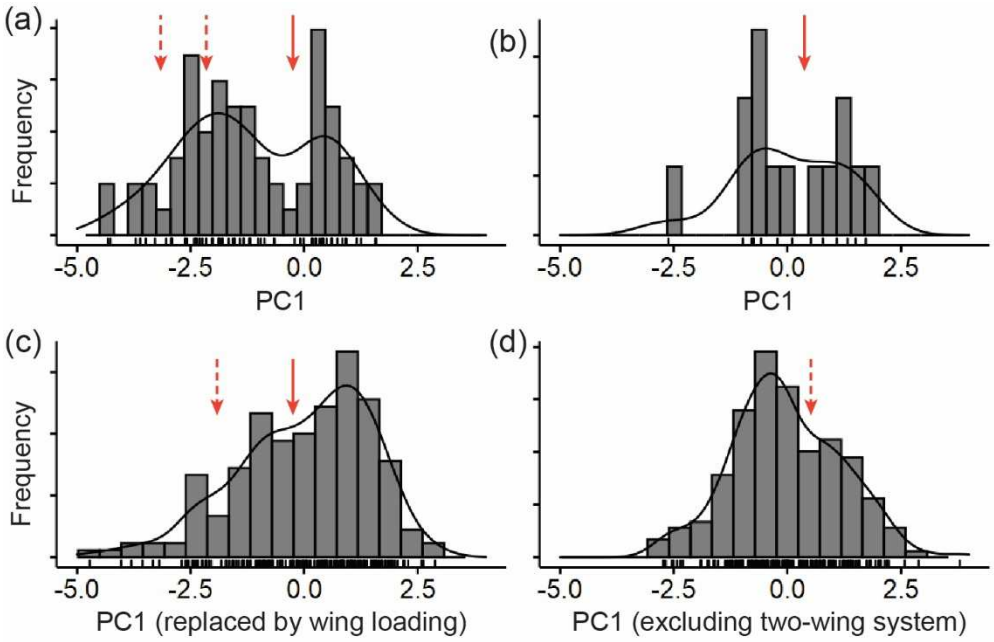
601 **Extended Data Figure 3| Diagram of the simulation of fore- and hindwing arrangements.**
602 (a) The definition of fore- and hindwing angles. (b) Visualization of different arrangements of an
603 exemplar species. (c) Aspect ratio (AR)-moment of area (MA) plot. The color corresponds to
604 that in **b**, and the star indicates the arrangement with the optimal aspect ratio (wingspan), which
605 is circled in **b**.
606



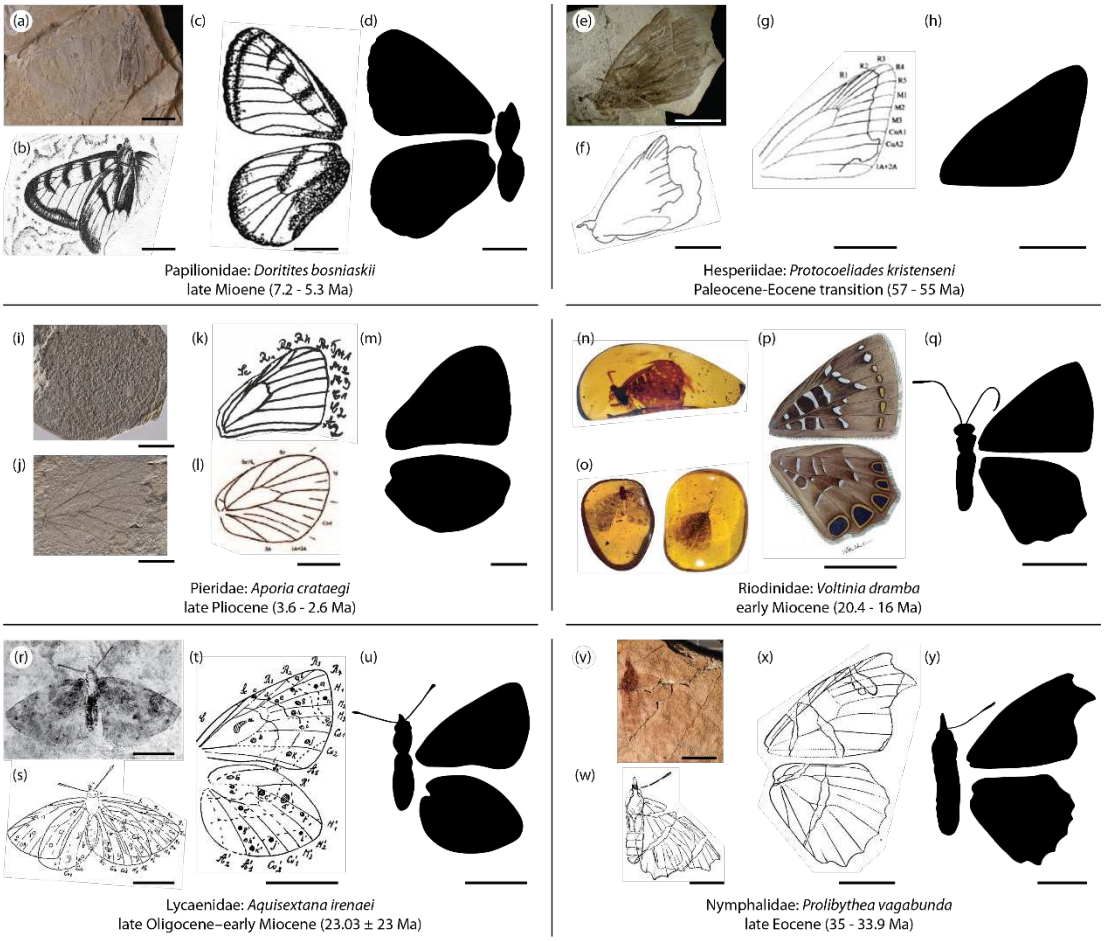
Extended Data Figure 4| The correlation between overall wing area and moment of area. A regression line is shown with 95% confidence interval (shaded area). See Extended Data Fig. 2 for the definition.



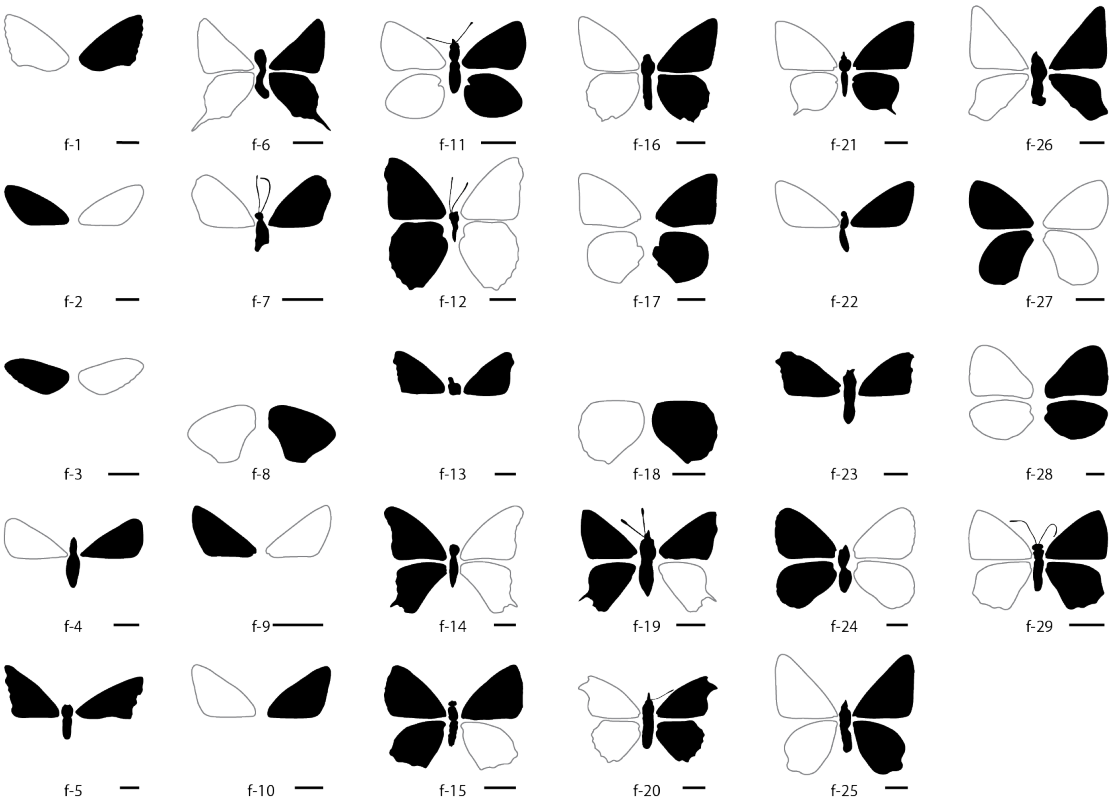
Extended Data Figure 5| The size-related morphologies, flight characteristics, and flight indices based on three different clustering approaches. (a) the clustering guided by the statistical results and density distribution of $PC1_{morph}$ (Fig. 3). (b) the clustering based on finite Gaussian mixture modeling (Fig. S1). (c) the clustering based on k-means clustering (Fig. S2). The patterns are generally similar among all three clustering methods, showing better gliding abilities in the larger system than the smaller one. Abbreviations are shown for convenience: the difference between fore- and hindwing areas (F-H area Diff.), overlapping area between fore- and hindwings (F-H Ovlp. area).



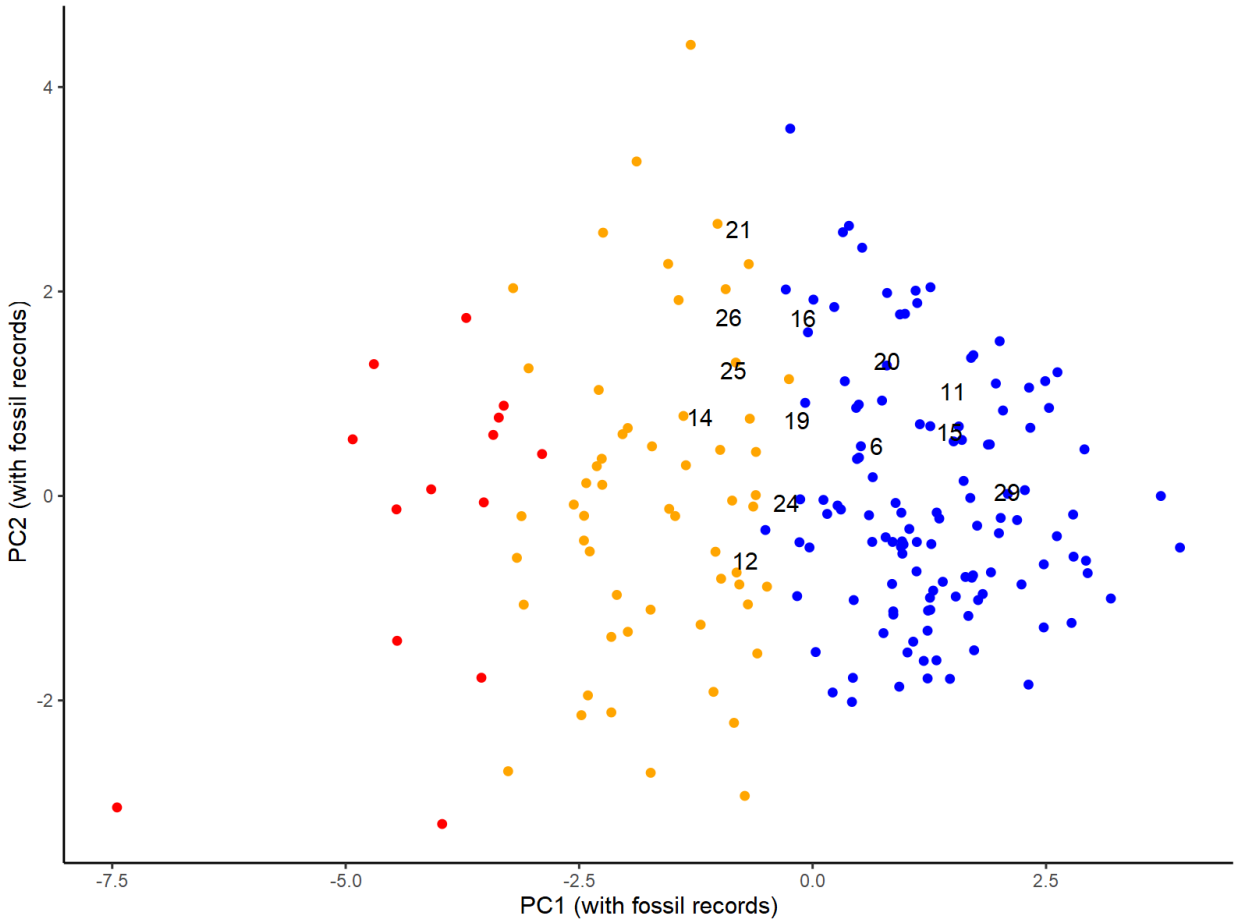
Extended Data Extended Data Figure 6| Replication of Fig. 3c with (a) Nymphalidae only, (b) Pieridae only, (c) body size and wing size replaced by wing loading (relative size), and (d) additional removal of F-H wing difference and F-H wing overlap based on c. (a-b) $PC1_{morph}$ is the same as in Fig. 3, and the samples are labeled above the x axis. Solid arrows indicate visible breaks of morphological systems, and dashed arrows indicate uncertain break of morphological systems. Note that the small sample sizes may lead to gaps in the histogram.



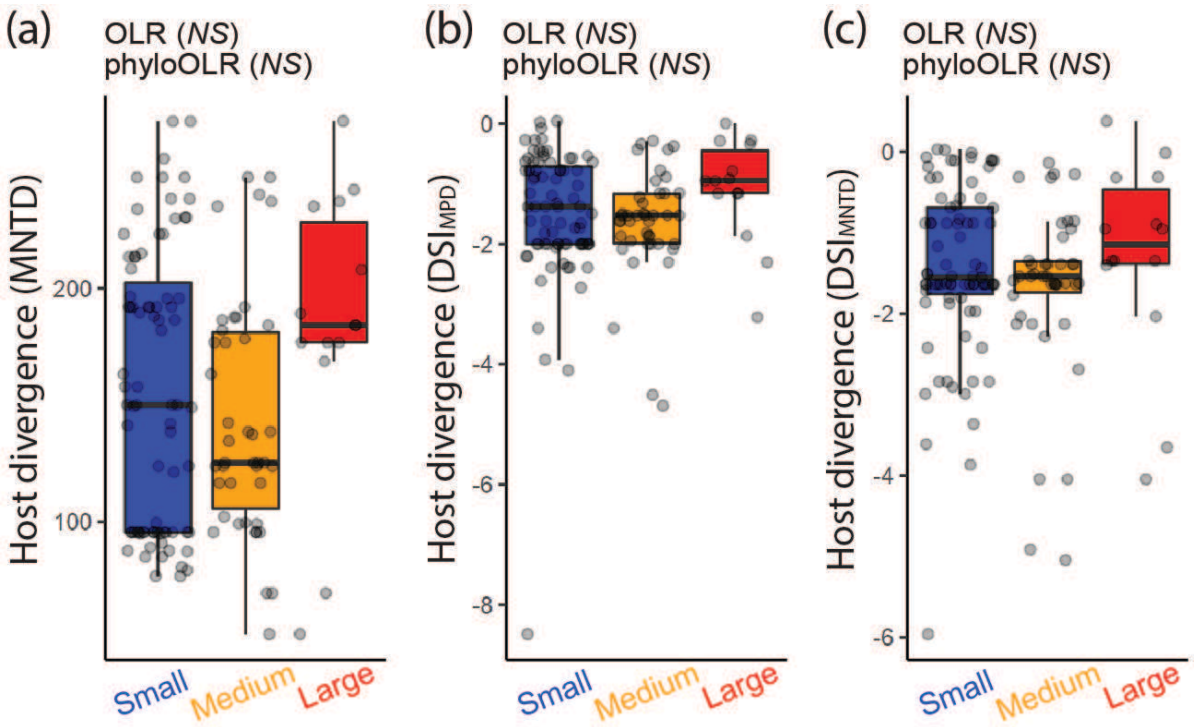
Extended Data Figure 7| Exemplar fossil images, expert illustrations, and extracted wing masks. One exemplar fossil was chosen per family (apart from Hedylidae where no fossils are known). The corresponding illustrations by paleobiologists are shown. The black mask was extracted by superimposing illustrations on the fossil images (Methods). All available fossil masks are shown in Extended Data Fig. 8 and documented in Extended Data Table 1.



Extended Data Figure 8| Morphological mask extracted from fossil images and illustrations. Available morphological masks extracted from a total of 29 fossils. Solid masks are those extracted from images, and hollow shapes are mirrored from the extracted mask for better visualization. See Methods for details and Extended Data Table 1 for corresponding fossil information: Supplementary Data 3 includes the vectorized images.



Extended Data Figure 9| An independent morphological PC plot used to identify the morphological system of fossils. All existing species are plotted and colored by morphological system (System S: blue, System M: orange, System L: red), and the fossil record IDs are plotted (Extended Data Table 1). Along $PC1_{morph}$, the morphological systems of each fossil can be clearly identified.



Extended Data Figure 10| Host plant breadth of butterfly species in different morphological systems. Two of the four indices of host plant breadth (Methods; Fig. 4f) were applied and mapped onto the morphological systems: (a) mean nearest taxon distance (MNTD), (b) the distance-based specialization index for the MPD, and (c) the distance based speciation index for the MNTD. Asterisks between bars show the significance of the difference between that pair of morphological systems. Significance levels are labeled: **, $p < 0.01$; *, $p < 0.05$; ., $p < 0.1$; NS, $p > 0.1$. The result of classic ordinal logistic regression (OLR) and phylogenetic OLR are both labeled on the top of each panel.

Extended data Table 1| Available fossil records with measurable morphology

ID[‡]	Family	Scientific name	Date	Date (Ma)[$] (Illustration)
F-1	*incertae sedis*	*Apanthesis leuce*	Late Eocene: Late Priabonian	33.9 – 35 ([66])
F-2	*incertae sedis*	*Archaeolycorea ferreirai*	Late Oligocene–Early Miocene boundary: Chattian–Aquitanian	23.03 ± 23[#] ([67])
F-3	*incertae sedis*	*Coliates proserpina*	Late Oligocene–Early Miocene boundary: Chattian–Aquitanian	23.03 ± 23[#] ([68])
F-4	*incertae sedis*	*Lithopsyche antiqua*	Late Eocene: Late Priabonian	33.9 – 35 ([66])
F-5	*incertae sedis*	*Nymphalites scudderi*	Late Eocene: Late Priabonian	33.9 – 35 (NA)
F-6	*incertae sedis*	*Riodinella nympha*	Middle Eocene: Early Lutetian	44 - 47.8 ([69])
F-7	*incertae sedis*	*Stolopsyche libytheoides*	Late Eocene: Late Priabonian	3.94 – 35 ([66])
F-8	*incertae sedis*	*Vanessa vetula*	Late Oligocene: Chattian	23 - 28.1 ([68])
F-9	Hesperiidae	*Pamphilites abdita*	Paleocene-Eocene transition	55 – 57 ([68])
F-10	Hesperiidae	*Protocoeliades kristenseni*	Late Paleocene–Early Eocene: Thanetian– Ypresian	56 ± 55[#] ([70])
F-11*	Lycaenidae	*Aquisextana irenaei*	Late Oligocene–early Miocene	23.03 ± 23[#] ([71])
F-12*	Nymphalidae	*Barbarothea florissanti*	Late Eocene: Late Priabonian	33.9 – 35 ([72])
F-13	Nymphalidae	*Chlorippe wilmattae*	Late Eocene: Late Priabonian	33.9 – 35 (NA)
F-14*	Nymphalidae	*Cyllo sepulta*	Late Oligocene–Early Miocene boundary: Chattian–Aquitanian	23.03 ± 23[#] ([68])
F-15*	Nymphalidae	*Dynamine alexae*	Early Miocene: Burdigalian	16 - 20.4 ([73])
F-16*	Nymphalidae	*Jupiteria charon*	Late Eocene: late Priabonian	28.1 - 33.9 ([66])
F-17	Nymphalidae	*Lithopsyche styx*	Late Eocene: Late Priabonian	33.9 – 35 ([66])
F-18	Nymphalidae	*Oligodonta florissantensis*	Late Eocene: Late Priabonian	33.9 – 35 ([74])
F-19*	Nymphalidae	*Prodryas persephone*	Late Eocene: Late Priabonian	33.9 – 35 ([75])
F-20*	Nymphalidae	*Prolibythea vagabunda*	Late Eocene: late Priabonian	33.9 – 35 ([66])
F-21*	Nymphalidae	*Pseudoneorina coulleti*	Early Oligocene: Stampien	28.1 - 33.9 ([76])
F-22	Nymphalidae	*Satyrites reynesii*	Late Oligocene–Early Miocene boundary: Chattian–Aquitanian	23.03 ± 23[#] ([68])
F-23	Nymphalidae	*Vanessa amerindica*	Late Eocene: Late Priabonian	33.9 – 35 ([67,77])

F-24*	Papilionidae	*Doritites bosniaskii*	Late Miocene: Messinian	5.3 - 7.2 ([78,79])
F-25*	Papilionidae	*Praepapilio colorado*	Middle Eocene: Early Lutetian	41.2 - 47.8 ([69])
F-26*	Papilionidae	*Praepapilio gracilis*	Middle Eocene: Early Lutetian	41.2 - 47.8 ([69])
F-27	Papilionidae	*Thaites ruminiana*	Late Oligocene–Early Miocene boundary: Chattian–Aquitanian	23.03 ± 23[#] ([67])
F-28	Pieridae	*Aporia crataegi*	Late Pliocene	2.58 - 3.6 ([80])
F-29*	Riodinidae	*Voltinia dramba*	Early Miocene	16 - 20.4 ([81])

*Those fossils are labeled in Fig. 3a & e. (include fossils with body, both fore- and hindwings and with known taxonomic information)

[$] The dates of fossils are based on De Jong[67]

[#]The value was estimated according to the GSA geologic time Scale

[‡]Only some fossils have available catalog numbers: F-1, PALE-4; F-5, AMNH_FI-18827; F-7, PALE-6; F-13, PALE-8602; F-16, USNM-38110; F-17, PALE-2; F-19, PALE-1; F-20, PALE-5; F-23, UF-21999. (PALE is the catalog number used in Museum of Comparative Zoology; AMNH, American Museum of Natural History; USNM, American Museum of Natural History; UF, The Florida Museum of Natural History)

Extended data Table 2| Model rankings for the ancestral state reconstruction (Fig. 3a & f-g). "Models separated by a '/' indicate a hidden rates model and the split distinguishes between the two state-dependent process (e.g. ER/ARD represents a hidden rate model where R1 is an equal rates model and R2 is an all rates differ model). AICc is the sample size corrected Akaike information criterion. AICcWt is the relative likelihood of each model and is used in model averaging. Mean rate is the average transition rate for a particular model. ASR is the most likely ancestral state reconstruction [of the butterfly ancestor] for a particular model and its marginal probability. k rates is the number of independent rate parameters being estimated for a given model. The row in bold is the best fitting model according to AICc" [58]

	AICc	AICcWt	Mean Rate	ASR	K rates
ER	270.19	<0.01	<0.01	(Sys. S, R1) 78%	1
SYM	249.87	0.05	0.01	(Sys. L, R1) 50%	3
ARD	**244.01**	**0.9**	**0.01**	**(Sys. M, R1) 74%**	**6**
ER/ER	269.37	<0.01	<0.01	(Sys. S, R1) 78%	4
SYM/SYM	251.39	0.02	0.12	(Sys. M, R2) 46%	8
ER/ARD	250.55	0.03	16.87	(Sys. M, R2) 74%	9
ARD/ARD	258.78	<0.01	0.01	(Sys. M, R2) 31%	14

Supplementary Information

Supplementary Methods

Imaging and image analysis

Specimen data labels were removed before placing specimens on the imaging platform. Specimens were imaged in a customized rig consisting of a high-resolution SLR camera (Nikon D800) modified to image up to thirty pinned butterfly specimens simultaneously[28].

After being run through the imaging processing pipeline, all results were checked for accuracy and quality (*e.g.*, wing completeness). Tails were defined as regions exceeding the hindwing region reconstructed by the top five harmonics in elliptical Fourier analysis. The tail length, curviness, and probability of occurrence were measured from multiple images of a species[28]. All pipeline coding was done in *Matlab 2018b* and run on the FASRC Cannon cluster supported by the FAS Division of Science Research Computing Group at Harvard University.

Evaluation of wing beating performance

Reynolds number (Re) is a unitless index that helps predict the flow under different situations. The lower the Re, the more stable flight should be[7]. According to Eq. S1 and Eq. S4 (see Supplementary Equations), we can have the following equation:

$Re \propto \sqrt{\frac{S}{AR}}$ (Eq. 1),

where S denotes the area of a wing, and AR denotes aspect ratio.

Rossby number (Ro) is a unitless index describing flow dynamics. From Eq. S9 and S10, we derive the following equation describing the relationship between Rossby number and flight characteristics:

$Ro \propto \frac{AR}{WL}$ (Eq. 2),

where AR denotes aspect ratio and WL denotes wing loading.

From Eq. 1 & 2, we can generate two indices of Re and Ro based on butterfly morphology. Given the size and aspect ratio of a butterfly, improved lift force and maneuverability (generating a range of lift force by adjusting overall wing area and aspect ratio) in wing beating

flight would correspond to low Rossby number (Ro) and Reynolds number (Re), respectively[82]. Thus, we derived a wing beating index from the sum of normalized indices of -Re and -Ro.

The evaluation of gliding performance

From Eq. S11 and Eq. S12, we can describe maximal lift-to-drag ratio (LD_{max}), which is the maximal amount of lift generated by a wing compared to its drag, as follows:

$LD_{max} \propto \sqrt{AR \times S}$ (Eq. 3)

According to Eq. S2 and S13, glide ratio (Gr), which is the travel distance divided by the altitude lost, is inversely correlated to WL.

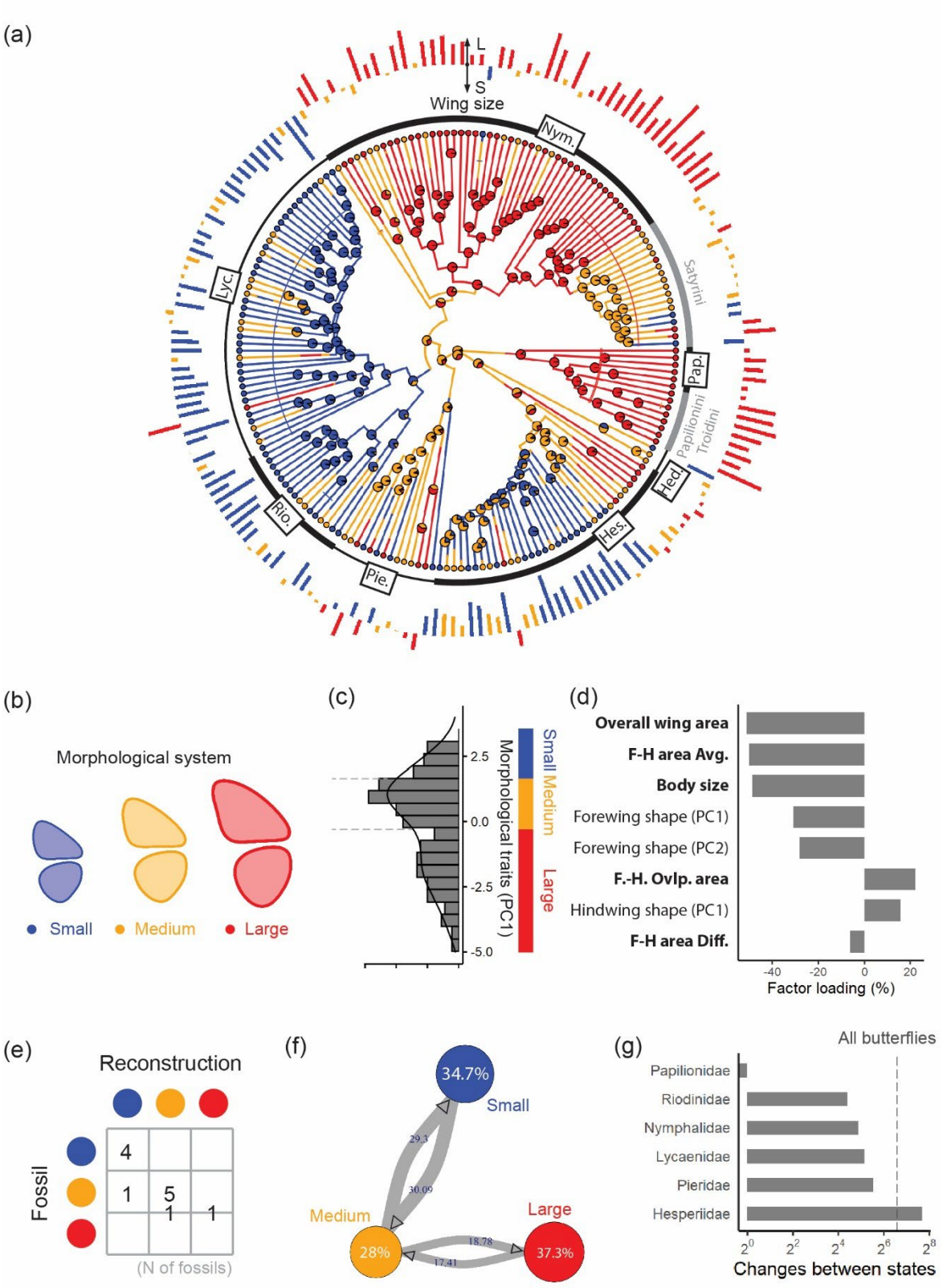
$Gr \propto \frac{1}{WL}$ (Eq. 4)

From Eq. 3 & 4, we can generate two indices of LD_{max} and Gr based on butterfly morphology. For better gliding ability, high maximal lift-to-drag ratio (LD_{max}) and high glide ratio (Gr) are expected[83], so the gliding index was derived from the sum of normalized indices of LD_{max} and Gr.

Summary of archived scripts

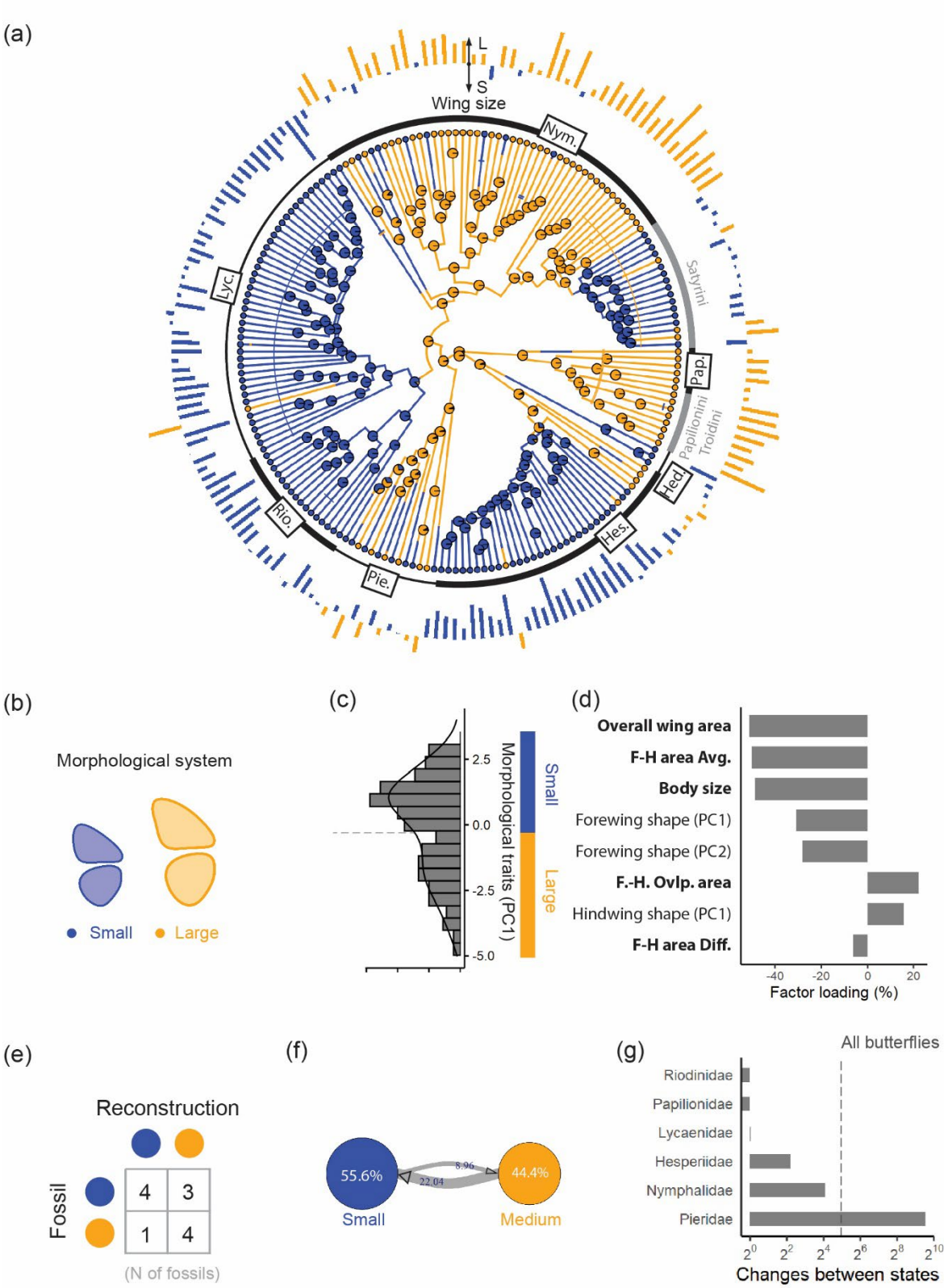
- Image processing[28]
- The simulation of fore- and hindwing arrangements: https://github.com/weipingchan/F-H_wing_arrangement_simulation
- Fossil image processing: https://www.protocols.io/private/A300FAE189F811EDB5D90A58A9FEAC02
- All remaining analyses, statistics, and visualization in this study: https://github.com/weipingchan/scripts_for_Chan_et_al_2023

735 **Supplementary Figures**



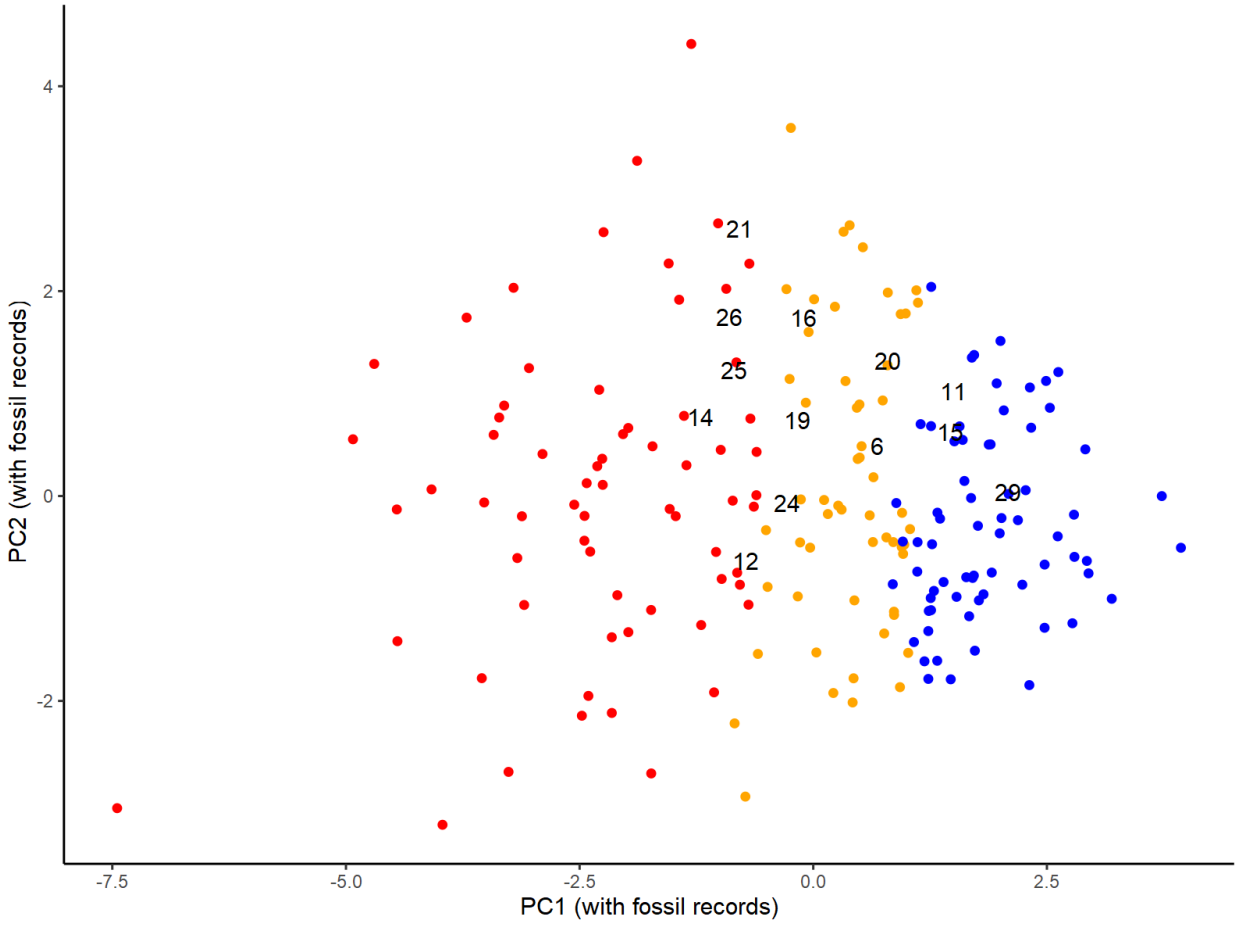
736

Supplementary Figure 1| Ancestral state reconstruction of butterfly wing morphological systems based on finite Gaussian mixture modelling. (a) The butterfly phylogeny is colored according to the three morphological systems identified in panel **c**. Family names are abbreviated on inner ring: Papilionidae (Pap); Hedylidae (Hed); Hesperiidae (Hes); Pieridae (Pie); Riodinidae (Rio); Lycaenidae (Lyc); Nymphalidae (Nym). Validated fossil records are labeled as bars segmenting the branch(es) corresponding to their identified taxonomic group(s). The outer ring bar chart shows total wing area (L: large; S: small). (b) Averaged fore- and hindwing shape of each morphological system is shown. (c) Histogram of the logical division boundaries (dashed line) for three morphological systems based on finite Gaussian mixture modeling. The probability curve is shown as a solid line. (d) The factor loading of morphologies on $PC1_{morph}$, which is identical to Fig. 3d. (e) A table comparing the morphological systems of fossils with ancestral state reconstructions for the nodes where they are placed. A fossil was placed on the boundary of two cells when it has ambiguous morphological system. (f) State changes frequencies among morphological systems. The residence time and state changes (Fig. S4) are labeled with the vertices and arrows, respectively. (g) Chart showing total changes between states in different taxonomic groups (each family separately, and all butterflies with a dotted line).



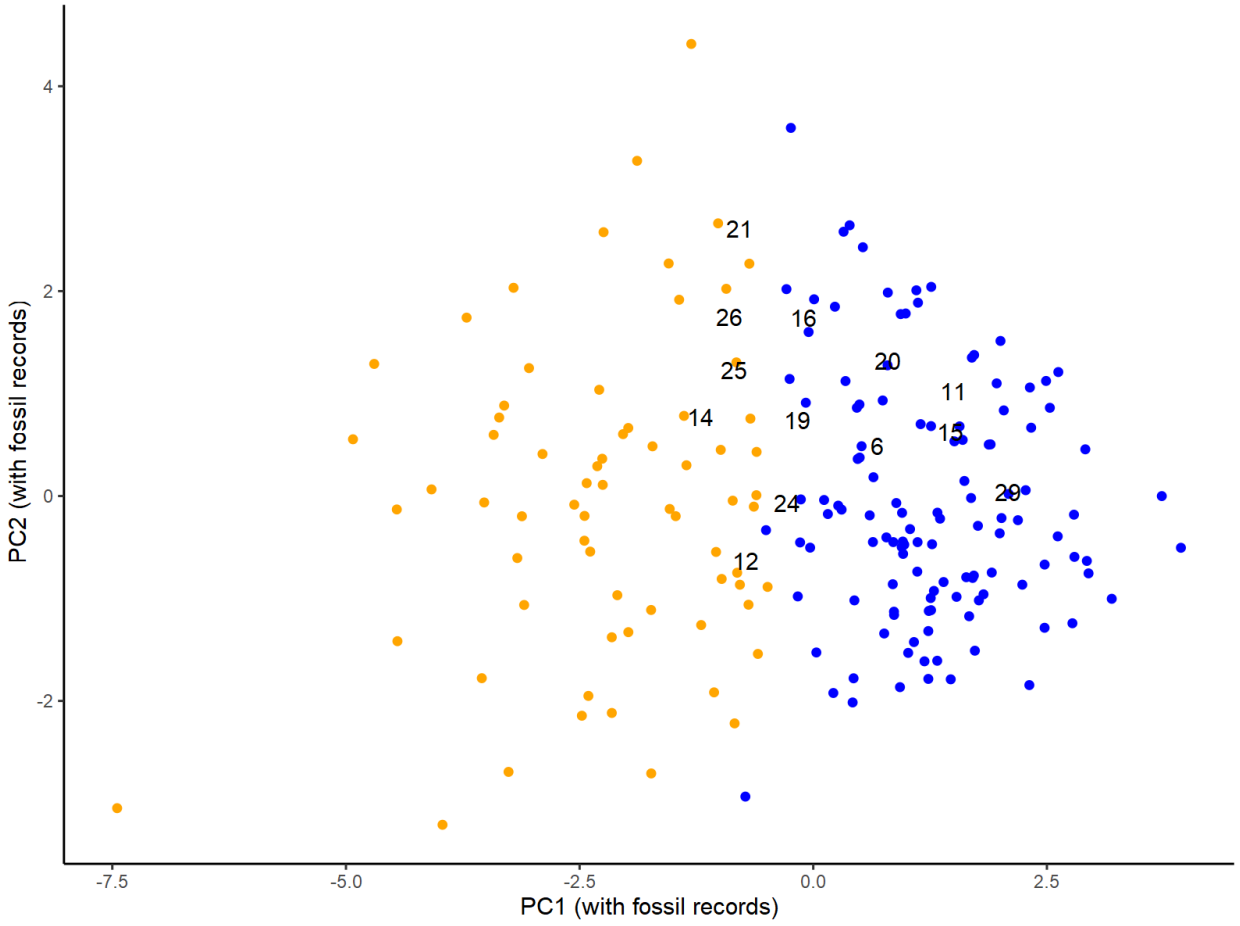
Supplementary Figure 2| Ancestral state reconstruction of butterfly wing morphological systems based on k-means clustering. (a) The butterfly phylogeny is colored according to the three morphological systems identified in panel **c**. Family names are abbreviated on inner ring:

Papilionidae (Pap); Hedylidae (Hed); Hesperiidae (Hes); Pieridae (Pie); Riodinidae (Rio); Lycaenidae (Lyc); Nymphalidae (Nym). Validated fossil records are labeled as bars segmenting the branch(es) corresponding to their identified taxonomic group(s). The outer ring bar chart shows total wing area (L: large; S: small). (b) Averaged fore- and hindwing shape of each morphological system is shown. (c) Histogram of the logical division boundaries (dashed line) for three morphological systems based on k-means clustering. The probability curve is shown as a solid line. (d) The factor loading of morphologies on $PC1_{morph}$, which is identical to Fig. 3d. (e) A table comparing the morphological systems of fossils with ancestral state reconstructions for the nodes where they are placed. A fossil was placed on the boundary of two cells when it had an ambiguous morphological system. (f) State changes frequencies among morphological systems. The residence time and state changes (Fig. S4) are labeled with the vertices and arrows, respectively. (g) Chart showing total changes between states in different taxonomic groups (each family separately, and all butterflies with a dotted line).



773

774 **Supplementary Figure 3|An independent morphological PC plot used to identify the**
775 **morphological system of fossils.** All existing species are plotted and colored by morphological
776 system (System S: blue, System M: orange, System L: red), and the fossil record IDs are plotted
777 (Extended Data Table 1). Along $PC1_{morph}$, the morphological systems of each fossil can be
778 clearly identified.
779



Supplementary Figure 4|An independent morphological PC plot used to identify the morphological system of fossils. All existing species are plotted and colored by morphological system (System S: blue, System M: orange, System L: red), and the fossil record IDs are plotted (Extended Data Table 1). Along $PC1_{morph}$, the morphological systems of each fossil can be clearly identified.

Supplementary Discussion

Relative wing size and wing loading

Relative size and allometry are critical concepts in biological studies applying comparative methods[84-86]. To derive relative size, wing size is usually statistically or mathematically standardized by body size[86-88], often by simply dividing wing size by body size (relative wing size) for each individual[10,86]. Relative wing size is mathematically the reciprocal of wing loading (body size divided by wing size^2). In many biological studies related to flight[2,23,30], relative wing size is therefore expressed as a function of wing loading– a trait that is commonly used in studies of aerodynamics. However, relative wing size (wing loading) has a strong effect on only two aerodynamic indices, whereas absolute wing affects all aerodynamic indices (Extended Data Fig.

797 2). Given its common role in aerodynamic studies[10,29,30], wing loading is treated as a flight
798 characteristic in this study (Extended Data Fig. 1 & 2), rather than the reciprocal of relative wing
799 size, and wing size when mentioned refers to wing size not being controlled by body size, so as
800 not to be mathematically redundant with wing loading. For clarity, we illustrate the effect of
801 using either wing loading (*i.e.*, relative wing size) or absolute wing size in generating three
802 morphological systems (Extended Data Fig. 6c). These morphological systems are not
803 identifiable if we exclude the fore-hindwing size differences and fore-hindwing overlapping
804 areas from the list of traits used in running a PCA analysis (Extended Data Fig. 6d). This result
805 emphasizes the importance of these traits for flight evolution in the two-wing systems of
806 butterflies.
807
808 **Detailed description regarding the link between wing size and flight characteristics**
809 **(Extended Data Fig. 1)**
810 Flight characteristics can evolve through alterations to wing size with or without changes in
811 shape (Extended Data Fig. 1). For example, aspect ratio can increase with increasing forewing-
812 hindwing size differences; moment of area increases with increasing wing size; and wing loading
813 can increase with smaller overall wing size and a higher percentage of overlapping area. Thus,
814 flight characteristics can evolve through alterations to wing size with or without changes in
815 shape.
816
817 **Moment of area in relation to flight**
818 Among the aerodynamic indices, it is difficult to define the role of moment of area, which is
819 generally thought to be a descriptor of overall shape rather than size[2,23]. However, in this large-
820 scale study across a wide range of wing sizes, moment of area is more representative of size than
821 shape, and variation in moment of area in the dataset can be largely explained by variation in size
822 rather than shape (Extended Data Fig. 2 & 4). To reconcile this finding with past interpretations
823 of moment of area, we statistically controlled for size by running a general linear model and
824 taking the residuals to create a new moment of area variable, MA_{ds}. By trying to embed MA_{ds}
825 into the aerodynamic functions, we found that MA_{ds} is not the primary characteristic determining
826 overall energy used for flight (but rather wing loading; Eq. S9). Alternatively, we interpret MA_{ds}
827 as relating to the balance between wing beating frequency and amplitude (Eq. S8). High MA_{ds}

would be associated with low-frequency, high-amplitude beating(*e.g.*, most Papilionidae); while low MA_{ds} would correspond to high-frequency, low-amplitude beating(*e.g.*, Hesperiidae). Earlier descriptions of the moment of area in ecology may have merged these in interpreting flight patterns.

An important trade-off between wing beating and gliding performance

The flight performance landscape (Fig. 4a left and middle panels; mapping wing beating and gliding indices together) visualizes how butterfly evolution is not significantly different from the null model (two-dimensional Kolmogorov-Smirnov test: $p = 0.23$), which is comprised of morphological traits randomly generated within the empirical data range (Fig. 4a right panel). A significant trade-off ($r = -0.995$, $p < 0.0001$ [empirical data]; $r = -0.969$, $p < 0.0001$ [null model]) between beating and gliding flight appears in both the flight performance landscape and its corresponding null model (Fig. 4a), indicating that this trade-off is of a fundamental nature to butterfly aerodynamics. Though the wing beating and gliding indices are highly negatively correlated, there is still a range in gliding ability accessible for a given beating ability in the flight performance landscape, and *vice versa*. Different families also optimize different regions of the adaptive landscape (Fig. 4a middle panel), indicating clear niche partitioning in flight performance.

Supplementary Equations

Aspect ratio (AR)

$AR = R/c = R^2/S = S/c^2$ (Eq. S1)

Where R represents the wing length; c, the chord length of a wing; S, wing area.

Wing loading (WL)

$WL = mg/S$ (Eq. S2)

Where m denotes, mass; g, gravity constant.

Moment of area (MA)

$MA = dA$ (Eq. S3) [23]

Where r denotes the perpendicular distance from the rotational axis of butterfly wings; dA is the infinitesimal area element.

Reynolds number (Re)

$Re = \sqrt{\frac{U_f \times c}{V}}$ (Eq. S4) [7]

Where U_f is the forward velocity; c, the chord length of a wing; V, the viscosity of the medium.

Rossby number (Ro)

$Ro = \frac{R}{c}\sqrt{J^2 + 1}$ (Eq. S5) [7]

Where J denotes advance ratio.

Advance ratio (J)

$J = \frac{U_f}{U_w}$ (Eq. S6) [7]

Where U_w denotes the wingtip velocity in the stroke plane

Since wing tip velocity can be determined by the amplitude of wingbeat and the frequency of wing beats, eq. S4 can be further rewritten as follows.

$J \propto \frac{U_f}{2\,f \times A}$ (Eq. S7) [7]

Where f denotes the frequency of wingbeats; A, the amplitude of each wingbeat.

Energy (E)

The higher the wingbeat frequency and amplitude, the more energy (E) is expended by the wing beating flight system. For a given amount of energy, the frequency and amplitude of wingbeats show the trade-off between higher wing beating frequency and large amplitude. Moment of area is also thought to influence the frequency-amplitude relationship (Supplementary Discussion).

$E \propto f \times A$ (Eq. S8)

For a given size of wing, higher wing loading (WL) indicates more energy being applied on a unit area of a wing.

$E \propto WL$ (Eq. S9)

From Eq. S3, S5, S7, and S8, we can derive the following equation describing the relationship between Rossby number and flight characteristics.

$Ro \propto AR \frac{1}{E}$ (Eq. S10)

Maximal lift to drag ratio (LD_{max})

$LD_{max} = \frac{1}{2}\sqrt{\frac{AR \times \pi \times \varepsilon}{C_{D,0}}}$ (Eq. S11) [83]

Where ε denotes span efficiency factor; $C_{D,0}$, zero lift-drag coefficient.

$C_{D,0} \propto C_D = 1.456 \times 10^5 \left(\frac{\eta \times P}{\sigma \times S \times {U_f}^3}\right)$ (Eq. S12) [83]

Where η denotes propulsive efficiency; P, power; σ, atmospheric density ratio.

Glide ratio (Gr; numerically equal to lift to drag ratio when the flying speed is constant)

At any given flying speed, glide ratio is positively correlated to wing area and negatively correlated to weight[89], and can be described as follows:

$Gr \propto \frac{S}{mg}$ (Eq. S13)

Supplementary Data

Supplementary data 1

The species list with migration notes integrated from Espeland *et al.* [27] and Chowdhury *et al.* [1]

Supplementary data 2

The specimen list we imaged in this study. Note that the barcodes beginning with "MCZ" are from Museum of Comparative Zoology, Harvard University, and those beginning with "MGCL" are from McGuire Center for Lepidoptera and Biodiversity, Florida Museum of Natural History.

Supplementary data 3

The fossil masks we integrated from available illustrations and images. They are provided as vectorized images.

Data availability

https://doi.org/10.5061/dryad.n8pk0p30g

Code availability

https://github.com/weipingchan/scripts_for_Chan_et_al_2023

Acknowledgements

We thank Olivier Bethoux (National Museum of Natural History), Andreas Kroh (Natural History Museum Vienna), and Alexander Gehler (Geoscience Center of the University of Göttingen) for helping us to image fossil butterflies. We thank Clayton Ziemke, Christian Alessandro Perez, Avalon Owens, Zhengyang Wang, Cong Liu, Evan Hoki, Beaziel Ombajen, Zoe Flores, Jocelyn Wang, Han-Ting Yang, Jingqian Wang, Jiale Chen, Annina Kennedy-Yoon and Atreyi Mukherji for helping with data inspection and manual corrections. Paul Moorcroft, David Haig, Scott Edwards, Andrew Davies, and Andrew Biewener provided useful comments and suggestions while we were developing this study. W-PC, MAC and KA were supported by graduate fellowships from the Department of Organismic and Evolutionary Biology at Harvard University; RARC and JDC were supported by Graduate Research Fellowships (GRFP) from the National Science Foundation (NSF), C-CT was supported by a graduate fellowship from the Department of Applied Physics and Mathematics at Columbia University. This research was funded by NSF DEB-0447242 and PHY-1411123 awarded to NEP and NY; and NSF DEB-1541500 to AYK and RG, 1541557 to DJL, and 1541560 to NEP.

References

1 Chowdhury, S., Fuller, R. A., Dingle, H., Chapman, J. W. & Zalucki, M. P. Migration in butterflies: a global overview. *Biol. Rev. Camb. Philos. Soc.* **96**, 1462-1483 (2021).

2 Weis-Fogh, T. Quick estimates of flight fitness in hovering animals, including novel mechanisms for lift production. *J. Exp. Biol.* **59**, 169-230 (1973).

3 Pomilla, C. & Rosenbaum, H. C. Against the current: an inter-oceanic whale migration event. *Biol. Lett.* **1**, 476-479 (2005).

4 Harvey, C., Baliga, V., Wong, J., Altshuler, D. & Inman, D. Birds can transition between stable and unstable states via wing morphing. *Nature* **603**, 648-653 (2022).

5 Bomphrey, R. J., Nakata, T., Henningsson, P. & Lin, H. T. Flight of the dragonflies and damselflies. *Philos Trans R Soc Lond B Biol Sci* **371** (2016).

6 Alexander, D. E. *Nature's Flyers: Birds, Insects, and The Biomechanics of Flight*. (JHU Press, 2002).

7 Chin, D. D. & Lentink, D. Flapping wing aerodynamics: from insects to vertebrates. *J Exp. Biol.* **219**, 920-932 (2016).

8 Gao, Z., Shi, Q., Fukuda, T., Li, C. & Huang, Q. An overview of biomimetic robots with animal behaviors. *Neurocomputing* **332**, 339-350 (2019).

9 Suárez-Tovar, C. & Sarmiento, C. Beyond the wing planform: morphological differentiation between migratory and nonmigratory dragonfly species. *J. Evol. Biol.* **29**, 690-703 (2016).

10 Shyy, W., Kang, C. K., Chirarattananon, P., Ravi, S. & Liu, H. Aerodynamics, sensing and control of insect-scale flapping-wing flight. *Proc. Math. Phys. Eng. Sci.* **472**, 20150712 (2016).

11 Suzuki, K. & Yoshino, M. A trapezoidal wing equivalent to a Janatella leucodesma's wing in terms of aerodynamic performance in the flapping flight of a butterfly model. *Bioinspiration Biomimetics* **14**, 036003 (2019).

12 Zhang, Y., Wang, X., Wang, S., Huang, W. & Weng, Q. Kinematic and aerodynamic investigation of the butterfly in forward free flight for the butterfly-inspired flapping wing air vehicle. *Appl. Sci.* **11**, 2620 (2021).

13 Bhat, S., Zhao, J., Sheridan, J., Hourigan, K. & Thompson, M. Aspect ratio studies on insect wings. *Phys. Fluids* **31**, 121301 (2019).

14 Fei, Y. H. & Yang, J. T. Enhanced thrust and speed revealed in the forward flight of a butterfly with transient body translation. *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.* **92**, 033004 (2015).

15 Nguyen, A. T., Han, J. S. & Han, J. H. Effect of body aerodynamics on the dynamic flight stability of the hawkmoth Manduca sexta. *Bioinspir. Biomim.* **12**, 016007 (2016).

16 Ortega Ancel, A. *et al.* Aerodynamic evaluation of wing shape and wing orientation in four butterfly species using numerical simulations and a low-speed wind tunnel, and its implications for the design of flying micro-robots. *Interface focus* **7**, 20160087 (2017).

17 Yeaton, I. J., Socha, J. J. & Ross, S. D. Global dynamics of non-equilibrium gliding in animals. *Bioinspir. Biomim.* **12**, 026013 (2017).

18 Nakata, T. *et al.* Aerodynamic imaging by mosquitoes inspires a surface detector for autonomous flying vehicles. *Science* **368**, 634-637 (2020).

19 Biewener, A. A., Bomphrey, R. J., Daley, M. A. & Ijspeert, A. J. Stability and manoeuvrability in animal movement: lessons from biology, modelling and robotics. *Proc Biol Sci* **289**, 20212492 (2022).

20 de Camargo, W. R., de Camargo, N. F., Correa Ddo, C., de Camargo, A. J. & Diniz, I. R. Sexual dimorphism and allometric effects associated with the wing shape of seven moth species of sphingidae (Lepidoptera: Bombycoidea). *J. Insect. Sci.* **15** (2015).

21 Owens, H. L., Lewis, D. S., Condamine, F. L., Kawahara, A. Y. & Guralnick, R. P. Comparative phylogenetics of *papilio* butterfly wing shape and size demonstrates independent hindwing and forewing evolution. *Syst. Biol.* **69**, 813-819 (2020).

22 Rubin, J. J. *et al.* The evolution of anti-bat sensory illusions in moths. *Sci. Adv.* **4**, eaar7428 (2018).

23 Le Roy, C., Debat, V. & Llaurens, V. Adaptive evolution of butterfly wing shape: from morphology to behaviour. *Biol. Rev.* **94**, 1261-1281 (2019).

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

25 Chazot, N., Blandin, P., Debat, V., Elias, M. & Condamine, F. L. Punctuational ecological changes rather than global factors drive species diversification and the evolution of wing phenotypes in *Morpho* butterflies. *J. Evol. Biol.* **34**, 1592-1607 (2021).

26 Klingenberg, C. P. Size, shape, and form: concepts of allometry in geometric morphometrics. *Dev. Genes. Evol.* **226**, 113-137 (2016).

27 Espeland, M. *et al.* A comprehensive and dated phylogenomic analysis of butterflies. *Curr. Biol.* **28**, 770-778 e775 (2018).

28 Chan, W.-P. *et al.* A high-throughput multispectral imaging system for museum specimens. *Commun. Biol.* **5**, 1-13 (2022).

29 Dudley, R. *The Biomechanics of Insect Flight: Form, Function, Evolution*. (Princeton University Press, 2002).

30 Le Roy, C. *et al.* Adaptive evolution of flight in Morpho butterflies. *Science* **374**, 1158-1162 (2021).
31 Aiello, B. R. *et al.* The evolution of two distinct strategies of moth flight. *J. R. Soc. Interface* **18**, 20210632 (2021).
32 Condamine, F. L., Rolland, J. & Morlon, H. Assessing the causes of diversification slowdowns: temperature-dependent and diversity-dependent models receive equivalent support. *Ecol. Lett.* **22**, 1900-1912 (2019).
33 Fletcher, B. J., Brentnall, S. J., Anderson, C. W., Berner, R. A. & Beerling, D. J. Atmospheric carbon dioxide linked with Mesozoic and early Cenozoic climate change. *Nat. Geosci.* **1**, 43-48 (2008).
34 Condamine, F. L., Silvestro, D., Koppelhus, E. B. & Antonelli, A. The rise of angiosperms pushed conifers to decline during global cooling. *Proc. Natl. Acad. Sci. USA* **117**, 28867-28875 (2020).
35 Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**, 686-693 (2001).
36 Steuber, T., Rauch, M., Masse, J. P., Graaf, J. & Malkoc, M. Low-latitude seasonality of Cretaceous temperatures in warm and cold episodes. *Nature* **437**, 1341-1344 (2005).
37 Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. Extraterrestrial cause for the Cretaceous-Tertiary extinction. *Science* **208**, 1095-1108 (1980).
38 Montejo-Kovacevich, G. *et al.* Genomics of altitude-associated wing shape in two tropical butterflies. *Mol Ecol* **30**, 6387-6402 (2021).
39 Sanzana, M.-J., Parra, L. E., Sepúlveda-Zúñiga, E. & Benítez, H. A. Latitudinal gradient effect on the wing geometry of Auca coctei (Guérin)(Lepidoptera, Nymphalidae). *Rev. Bras. Entomol.* **57**, 411-416 (2013).
40 Wells, C., Munn, A. & Woodworth, C. Geomorphic morphometric differences between populations of Speyeria diana (Lepidoptera: Nymphalidae). *Fla. Entomol.* **101**, 195-202 (2018).
41 Chapman, J. W., Reynolds, D. R. & Wilson, K. Long-range seasonal migration in insects: mechanisms, evolutionary drivers and ecological consequences. *Ecol. Lett.* **18**, 287-302 (2015).
42 Dockx, C. Directional and stabilizing selection on wing size and shape in migrant and resident monarch butterflies, *Danaus plexippus* (L.), in Cuba. *Biol. J. Linn. Soc.* **92**, 605-616 (2007).
43 Strömberg, C. A. Evolution of grasses and grassland ecosystems. *Annu. Rev. Earth Planet. Sci.* **39**, 517-544 (2011).
44 Pena, C. & Wahlberg, N. Prehistorical climate change increased diversification of a group of butterflies. *Biol. Lett.* **4**, 274-278 (2008).
45 Altizer, S. & Davis, A. K. Populations of monarch butterflies with different migratory behaviors show divergence in wing morphology. *Int. J. Org. Evol.* **64**, 1018-1028 (2010).
46 Chowdhury, S. *et al.* Seasonal spatial dynamics of butterfly migration. *Ecol. Lett.* **24**, 1814-1823 (2021).
47 Kembel, S. W. *et al.* Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463-1464 (2010).
48 García-Berro, A. *et al.* Migratory behaviour is positively associated with genetic diversity in butterflies. *Mol. Ecol.* (2022).

49 Bonhomme, V., Picq, S., Gaucherel, C. & Claude, J. Momocs: outline analysis using R. *Journal of Stat. Software* **56**, 24 p. (2014).

50 Adams, D. C. & Otárola-Castillo, E. Geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol. Evol.* **4**, 393-399 (2013).

51 Politis, D. N. & Romano, J. P. The stationary bootstrap. *J. Am. Stat. Assoc.* **89**, 1303-1313 (1994).

52 Kawahara, A. Y. *et al.* Phylogenetics of moth-like butterflies (Papilionoidea: Hedylidae) based on a new 13-locus target capture probe set. *Mol. Phylogenet. Evol.* **127**, 600-605 (2018).

53 Chevan, A. & Sutherland, M. Hierarchical partitioning. *Am. Stat.* **45**, 90-96 (1991).

54 Grömping, U. Relative importance for linear regression in R: the package relaimpo. *J. Stat. software* **17**, 1-27 (2007).

55 Li, D., Dinnage, R., Nell, L. A., Helmus, M. R. & Ives, A. R. Phyr: an R package for phylogenetic species-distribution modelling in ecological communities. *Methods Ecol. Evol.* **11**, 1455-1463 (2020).

56 Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E. Mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R journal* **8**, 289 (2016).

57 Kassambara, A. & Mundt, F. Package 'factoextra'. *Extract and visualize the results of multivariate data analyses* **76** (2017).

58 Boyko, J. D. & Beaulieu, J. M. Generalized hidden Markov models for phylogenetic comparative datasets. *Methods Ecol. Evol.* **12**, 468-478 (2021).

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

60 Veizer, J. & Prokoph, A. Temperatures and oxygen isotopic composition of Phanerozoic oceans. *Earth Sci. Rev.* **146**, 92-104 (2015).

61 Storer, C. *et al.* Comprehensive phylogeny of butterflies reveals their ancestral hostplant and biogeographic origin. *Nat. Ecol. Evol.* (2023).

62 Faith, D. P. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* **61**, 1-10 (1992).

63 Webb, C. O., Ackerly, D. D., McPeek, M. A. & Donoghue, M. J. Phylogenies and community ecology. *Annu. Rev. Ecol. Evol. Syst.* **33**, 475-505 (2002).

64 Jorge, L. R., Prado, P. I., Almeida-Neto, M. & Lewinsohn, T. M. An integrated framework to improve the concept of resource specialisation. *Ecol. Lett.* **17**, 1341-1350 (2014).

65 Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.* **105**, 302-314 (2018).

66 Scudder, S. H. *The Fossil Butterflies of Florissant*. (US Government Printing Office, 1889).

67 De Jong, R. Fossil butterflies, calibration points and the molecular clock (Lepidoptera: Papilionoidea). *Zootaxa* **4270**, 1-63 (2017).

68 Scudder, S. H. *Fossil Butterflies*. Vol. 1 (FW Putnam, 1875).

69 Durden, C. J. & Rose, H. *Butterflies From the Middle Eocene: the Earliest Occurrence of Fossil Papilionoidea (Lepidoptera).* (Texas Memorial Museum, The University of Texas at Austin, 1978).

70 De Jong, R. Reconstructing a 55-million-year-old butterfly (Lepidoptera: Hesperiidae). *Eur. J. Entomol.* **113** (2016).

71 Théobald, N. Note complémentaire sur les insectes fossiles Oligocènes des gypses d'Aix-en-Provence. *Extrait du Bulletin de la Société des Sciences de Nancy* **6**, 157-178 (1937).

72 Scudder, S. H. *Some Insects of Special Interest from Florissant, Colorado, and Other Points in The Tertiaries of Colorado and Utah*. (US Government Printing Office, 1892).

73 Penalver, E. & Grimaldi, D. A. New data on Miocene butterflies in Dominican amber (Lepidoptera: Riodinidae and Nymphalidae) with the description of a new nymphalid. *Am. Mus. Novit.* **2006**, 1-17 (2006).

74 Brown, F. M. *Oligodonta Florissantensis, Gen. N., Sp. Nov: Lepidoptera, Pieridae*. (Allyn Museum of Entomology, 1976).

75 LeConte, J. *Elements of Geology: A Text-book for Colleges and for the General Reader*. (Appleton, 1902).

76 Nel, A. & Descimon, H. Une nouvelle espèce de Lépidoptère fossile du Stampien de Cérestre (04)(Lepidoptera-Satyridae). *Gèologie Mèditerranèenne* **11**, 287-293 (1984).

77 Miller, J. Y. & Brown, F. M. A new Oligocene fossil butterfly, Vanessa amerindica (Lepidoptera: Nymphalidae), from the Florissant formation, Colorado. *Bull. Allyn Mus.* (1989).

78 Ponomarenko, A. G. & Schultz, O. *Typen der Geologisch-Paläontologischen Abteilung: Fossile Insekten*. (Naturhistorisches Museum, 1988).

79 Rebel, H. *Fossile Lepidopteren aus der Miocänformation von Gabbro*. (1898).

80 Branscheid, F. Funde von Tagfaltern (Rhopalocera, Lepidopt.) im Pliozän von Willershausen. *Berichte der Naturhistorischen Gesellschaft, Hannover* **113**, 101-106 (1969).

81 Hall, J. P., Robbins, R. K. & Harvey, D. J. Extinction and biogeography in the Caribbean: new evidence from a fossil riodinid butterfly in Dominican amber. *Proc. Biol. Sci.* **271**, 797-801 (2004).

82 Bhat, S. S., Zhao, J., Sheridan, J., Hourigan, K. & Thompson, M. C. Uncoupling the effects of aspect ratio, Reynolds number and Rossby number on a rotating insect-wing planform. *J. Fluid Mech.* **859**, 921-948 (2019).

83 Loftin, L. K. *Quest for Performance: the Evolution of Modern Aircraft*. (Scientific and Technical Information Branch, NASA, 1985).

84 García-barros, E. Body size, egg size, and their interspecific relationships with ecological and life history traits in butterflies (Lepidoptera: Papilionoidea, Hesperioidea). *Biol. J. Linn. Soc.* **70**, 251-284 (2000).

85 Gittleman, J. L. & Luh, H.-K. On comparing comparative methods. *Annu. Rev. Ecol. Evol. Syst.*, 383-404 (1992).

86 Glazier, D. S. Complications with body-size correction in comparative biology: possible solutions and an appeal for new approaches. *J. Exp. Biol.* **225** (2022).

87 Frankino, W. A., Zwaan, B. J., Stern, D. L. & Brakefield, P. M. Internal and external constraints in the evolution of morphological allometries in a butterfly. *Evolution* **61**, 2958-2970 (2007).

88 Rosenberg, R. H. & Enquist, M. Contest behaviour in Weidemeyer's admiral butterfly Limenitis weidemeyerii (Nymphalidae): the effect of size and residency. *Anim. Behav.* **42**, 805-811 (1991).

89 Thorington Jr, R. W. & Heaney, L. R. Body proportions and gliding adaptations of flying squirrels (Petauristinae). *J. Mammal.* **62**, 101-114 (1981).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryData1.csv
- SupplementaryData2.csv