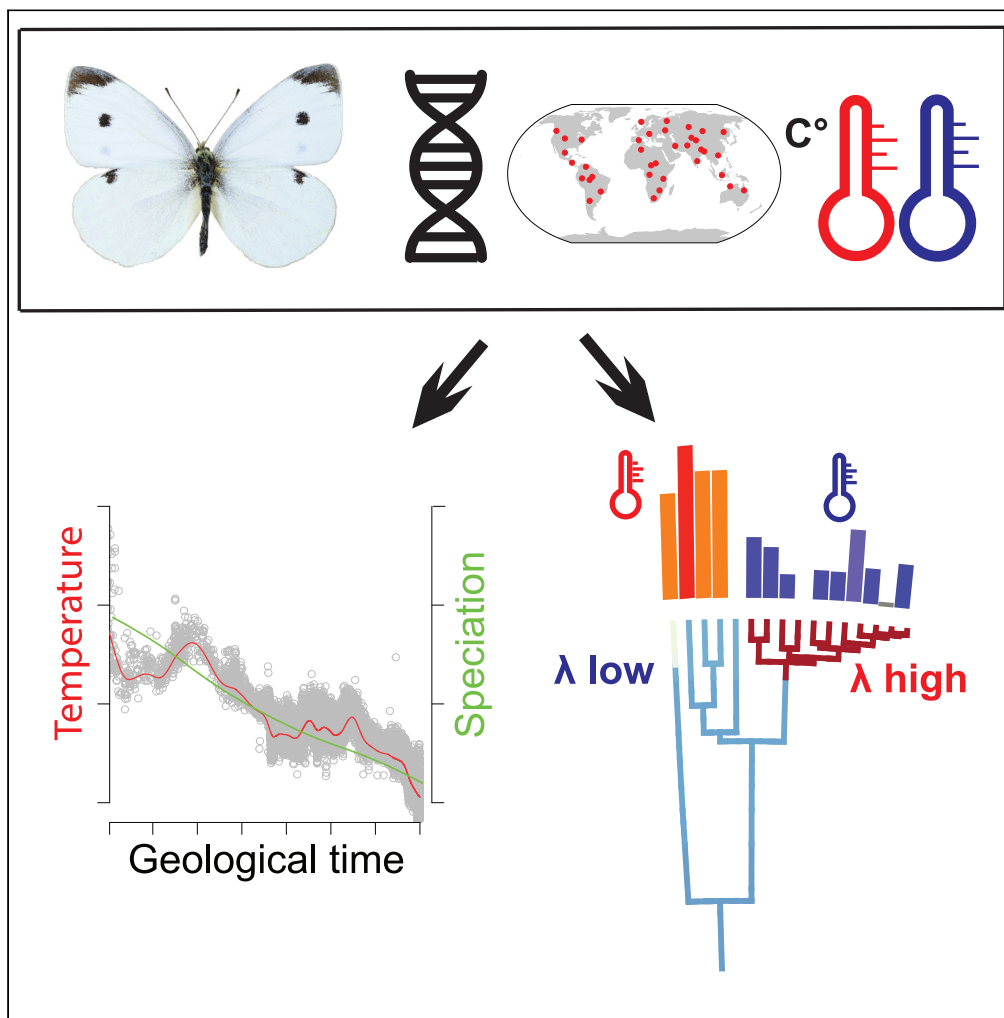


Article

# Comprehensive phylogeny of Pieridae butterflies reveals strong correlation between diversification and temperature



Ana Paula S. Carvalho, Hannah L. Owens, Ryan A. St Laurent, ..., Naomi E. Pierce, Akito Y. Kawahara, David J. Lohman

apsdecarvalho@gmail.com

**Highlights**

The most comprehensive phylogeny of Pieridae to date, with 593 pierid species

Mild warm seasons are associated with higher speciation rates

Habitats with high daily temperatures host species with lower speciation rates

Diversification decreased in tandem with global temperatures through time

Carvalho et al., iScience 27, 109336  
April 19, 2024 © 2024 The Authors.  
<https://doi.org/10.1016/j.isci.2024.109336>



## Article

## Comprehensive phylogeny of Pieridae butterflies reveals strong correlation between diversification and temperature

Ana Paula S. Carvalho,<sup>1,32,\*</sup> Hannah L. Owens,<sup>2,3,4</sup> Ryan A. St Laurent,<sup>1,5</sup> Chandra Earl,<sup>6</sup> Kelly M. Dexter,<sup>1</sup> Rebecca L. Messcher,<sup>1</sup> Keith R. Willmott,<sup>1</sup> Kwaku Aduse-Poku,<sup>7</sup> Steve C. Collins,<sup>8</sup> Nicholas T. Homziak,<sup>1,9</sup> Sugihiko Hoshizaki,<sup>10</sup> Yu-Feng Hsu,<sup>11</sup> Athulya G. Kizhakke,<sup>12</sup> Krushnamegh Kunte,<sup>12</sup> Dino J. Martins,<sup>13,14</sup> Nicolás O. Mega,<sup>15</sup> Sadaharu Morinaka,<sup>16</sup> Djunijanti Peggie,<sup>17</sup> Helena P. Romanowski,<sup>18</sup> Szabolcs Sáfán,<sup>8,19</sup> Roger Vila,<sup>20</sup> Houshuai Wang,<sup>21</sup> Michael F. Braby,<sup>22,23</sup> Marianne Espeland,<sup>24</sup> Jesse W. Breinholt,<sup>1,25</sup> Naomi E. Pierce,<sup>26</sup> Akito Y. Kawahara,<sup>1,9,27,31</sup> and David J. Lohman<sup>28,29,30,31</sup>

## SUMMARY

**Temperature is thought to be a key factor influencing global species richness patterns. We investigate the link between temperature and diversification in the butterfly family Pieridae by combining next generation DNA sequences and published molecular data with fine-grained distribution data. We sampled nearly 600 pierid butterfly species to infer the most comprehensive molecular phylogeny of the family and curated a distribution dataset of more than 800,000 occurrences. We found strong evidence that species in environments with more stable daily temperatures or cooler maximum temperatures in the warm seasons have higher speciation rates. Furthermore, speciation and extinction rates decreased in tandem with global temperatures through geological time, resulting in a constant net diversification.**

<sup>1</sup>McGuire Center for Lepidoptera and Biodiversity, Florida Museum of Natural History, Gainesville, FL, USA

<sup>2</sup>Center for Global Mountain Biodiversity, Globe Institute, University of Copenhagen, Copenhagen, Denmark

<sup>3</sup>Center for Macroecology, Evolution, and Climate, Globe Institute, University of Copenhagen, Copenhagen, Denmark

<sup>4</sup>Florida Museum of Natural History, University of Florida, Gainesville, FL, USA

<sup>5</sup>Department of Entomology, Smithsonian Institution, National Museum of Natural History, Washington, DC, USA

<sup>6</sup>Department of Natural Sciences, Bernice Pauahi Bishop Museum, Honolulu, HI, USA

<sup>7</sup>Biology Department, Howard University, Washington, DC, USA

<sup>8</sup>African Butterfly Research Institute, Karen, Nairobi, Kenya

<sup>9</sup>Department of Entomology and Nematology, University of Florida, Gainesville, FL, USA

<sup>10</sup>Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

<sup>11</sup>Department of Life Science, National Taiwan Normal University, Taipei, Taiwan, R.O.C

<sup>12</sup>National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bengaluru, India

<sup>13</sup>Turkana Basin Institute, Stony Brook University, Stony Brook, NY, USA

<sup>14</sup>Insect Committee of Nature Kenya, The East Africa Natural History Society, Nairobi, Kenya

<sup>15</sup>Programa de Pós-Graduação em Biologia Animal, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

<sup>16</sup>Saitama Study Center, The Open University of Japan, Omiya-ku, Saitama City, Japan

<sup>17</sup>Museum Zoologi Bogor, Research Center for Biosystematics and Evolution, Research Organization for Life Sciences and Environment, National Research and Innovation Agency, Cibinong, Bogor, Indonesia

<sup>18</sup>Laboratório de Ecologia de Insetos, Departamento de Zoologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

<sup>19</sup>Institute of Silviculture and Forest Protection, University of Sopron, Sopron, Hungary

<sup>20</sup>Institut de Biologia Evolutiva (CSIC-Univ. Pompeu Fabra), Barcelona, Spain

<sup>21</sup>Department of Entomology, College of Plant Protection, South China Agricultural University, Guangzhou, China

<sup>22</sup>Division of Ecology and Evolution, Research School of Biology, The Australian National University, Acton, ACT, Australia

<sup>23</sup>Australian National Insect Collection, National Research Collections Australia, Canberra, ACT, Australia

<sup>24</sup>Leibniz Institute for the Analysis of Biodiversity Change, Museum Koenig, Bonn, Germany

<sup>25</sup>Intermountain Healthcare, Intermountain Precision Genomics, St. George, UT, USA

<sup>26</sup>Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, MA, USA

<sup>27</sup>Department of Biology, University of Florida, Gainesville, FL, USA

<sup>28</sup>Department of Biology, City University of New York, New York, NY, USA

<sup>29</sup>PhD Program in Biology, Graduate Center, City University of New York, New York, NY, USA

<sup>30</sup>Entomology Section, National Museum of Natural History, Manila, Philippines

<sup>31</sup>These authors are equal contributors

<sup>32</sup>Lead contact

\*Correspondence: [apsdecarvalho@gmail.com](mailto:apsdecarvalho@gmail.com)

<https://doi.org/10.1016/j.isci.2024.109336>



## INTRODUCTION

Understanding drivers of species diversity across Earth is a central pursuit of ecology and evolutionary biology.<sup>1</sup> While multiple biotic and abiotic factors can influence speciation and extinction,<sup>2–5</sup> temperature is thought to be a key variable impacting spatial diversity patterns.<sup>6–9</sup> Diversity is often greater in warmer environments due to higher evolutionary rates,<sup>2</sup> as these species tend to have higher mutation rates, faster metabolism, and shorter generation times, all of which are directly associated with increased rates of evolution.<sup>2,4,10–12</sup> Due to the lack of pronounced temperature variation in species-rich areas (e.g., some tropical regions), authors also hypothesize that temperature stability could affect diversification.<sup>13</sup> This relationship is based on the premise that climatic stability favors the evolution of specialized adaptations (climatic stability hypothesis<sup>13</sup>). The inherent temperature stability of tropical climates compared to temperate ones is therefore thought to reduce extinction rates.

To examine whether temperature is correlated with diversification, it is necessary to study widely distributed clades of species inhabiting diverse biomes that have sufficiently fine-grained spatial data. Additionally, a well-supported, dated phylogenetic tree is necessary to account for relationships among species when evaluating patterns of diversity. Although many studies have assessed the correlation between ambient temperature and diversity, these studies have primarily focused on vertebrates and plants.<sup>3,14–22</sup> Considerably less attention has been directed to other, substantially more diverse groups like insects, which—as ectotherms—have life histories more likely to be intrinsically connected to temperature.<sup>23</sup> Butterflies are among the best documented insect groups, with nearly 34.5 million locality records in the Global Biodiversity Information Facility (GBIF, as of September 2022).

Studies of butterfly diversification have largely focused on the coevolution with their host plants.<sup>24,25</sup> Prior investigations have found strong evidence for reciprocal adaptations between Pierinae butterflies and their host plants as predicted by an escape-and-radiate coevolutionary scenario.<sup>26,27</sup> However, an extensive study examining evolutionary patterns of host use across all butterflies found weak evidence for coevolution in many clades.<sup>25</sup> Host plants are just one factor that might affect insect diversification rates; less attention has been paid to the role of climate.

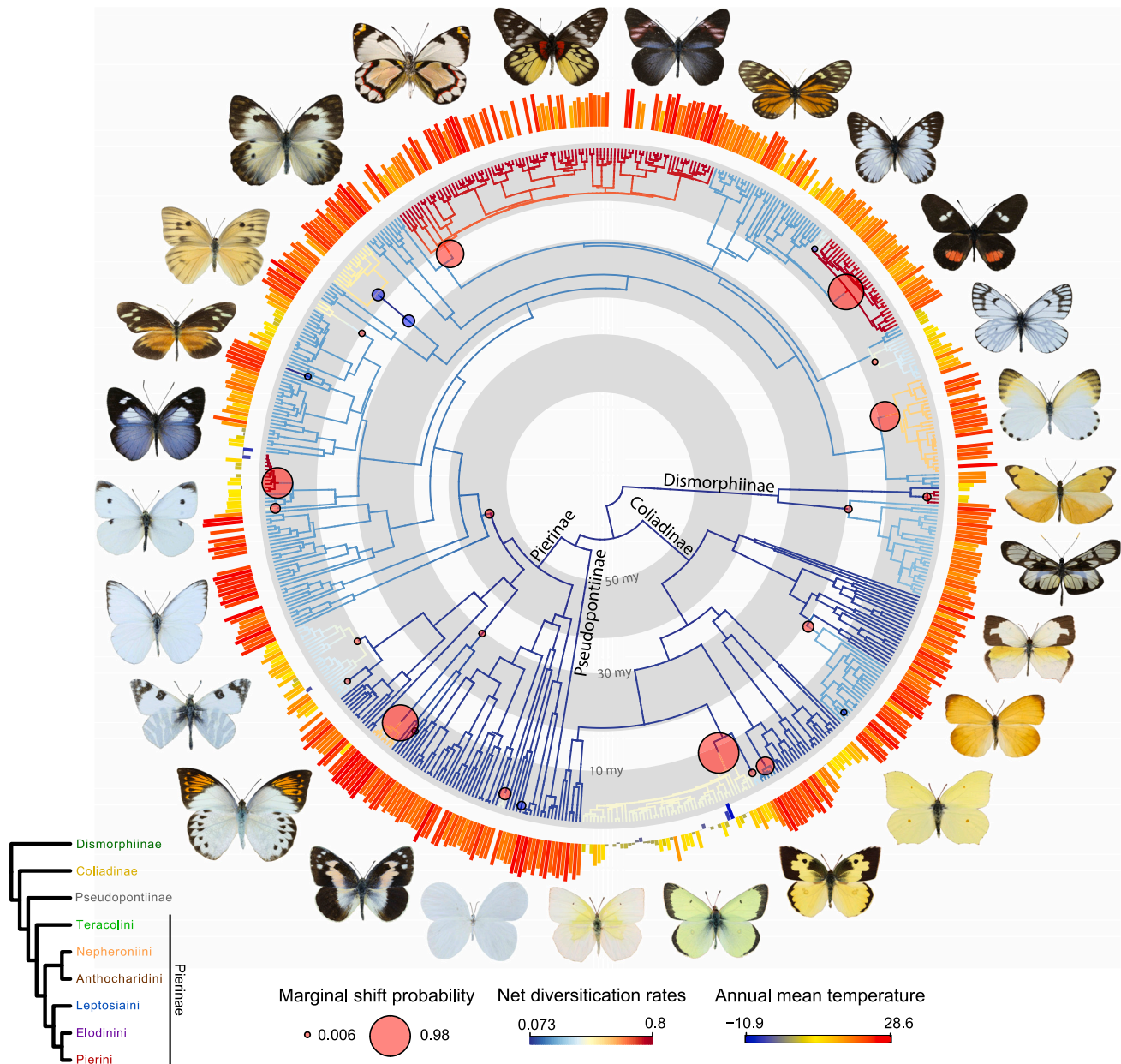
Butterfly species richness is generally much higher in equatorial regions than temperate regions,<sup>28</sup> and is also correlated with ambient temperature.<sup>6,29</sup> As with other ectotherms, temperature strongly influences insect larval development and diapause.<sup>30</sup> Ambient temperature and sunlight have direct effects on behaviors, such as basking, and indirect effects on life history traits, including phenology and voltinism.<sup>31,32</sup> Consequently, extreme or highly variable temperatures can negatively impact fitness.<sup>33</sup> Temperature-mediated diversification is another important context in which to study diversification dynamics in ectothermic organisms such as butterflies.

The family Pieridae, sometimes called the white and yellow butterflies, comprises at least 1,159 species in 86 genera distributed across the globe.<sup>34</sup> Most of their diversity is concentrated in the tropics,<sup>35</sup> but Pieridae is known for having species adapted to extreme cold or dry conditions, such as high montane habitats,<sup>36</sup> the Arctic<sup>37</sup> and deserts.<sup>38</sup> The family is sensitive to ambient temperature, and morphological adaptations for regulating temperature have been identified in several lineages.<sup>39–42</sup> Pierid monophyly is well established,<sup>43,44</sup> but relationships among subfamilies and tribes remain contentious (see Braby M.F. and Pierce N.E., Braby M.F. et al., and Warren-Gash H. et al.<sup>45–47</sup> for clade- or region-specific studies) (Figure 1), and current pierid phylogenies are neither sufficiently robust nor comprehensive enough to address macroevolutionary questions. A robust phylogeny of Pieridae will not only resolve its evolutionary relationships but also help answer questions regarding the role of thermal ecology in insect macroevolution. Here, we reconstruct a comprehensive species-level dated tree of Pieridae with 593 species from 84 genera and infer their diversification patterns, wherever possible accounting for incomplete sampling. We combine our phylogenetic reconstruction with climate data extracted using distribution records to examine how speciation and extinction are correlated with paleotemperatures and how present-day (1) warmer climates (annual mean temperature), (2) seasonality and daily temperature variability (temperature annual range and annual mean diurnal range), and (3) extreme temperatures (maximum temperature of the warmest month and minimum temperature in the coldest month) impact pierid diversity. We evaluated three related hypotheses, phrased here as questions: (1) Are temperature stability and warmer climates associated with higher pierid diversity as predicted by the climatic stability hypothesis?; (2) How does climate variability over time and space impact diversification dynamics?; and (3) What aspects of climate are most strongly associated with diversification rate shifts?

## RESULTS

### Pieridae phylogeny

We inferred the most taxonomically comprehensive phylogeny of Pieridae to date (Figures 1 and S1), including genetic data for 593 pierid species and 15 outgroups. Eighty-four of the 86 Pieridae genera recognized by Lamas<sup>34</sup> are included; the two missing genera, *Calopieris* and *Piercolias* are rarely collected, and the validity of the latter has been recently questioned.<sup>48</sup> Subfamilial relationships were strongly supported (UFBoot and SH-aLRT >95%) in general, but support for the position of Pseudopontiinae, which we recovered as the sister group to Pierinae, was low (UFBoot = 75.2, SH-aLRT = 77). This corroborates the results of Kawahara et al.<sup>25</sup> and Espeland et al.<sup>49</sup> based on 391 loci, but not those of Chazot et al.<sup>50</sup> which used nine loci (Figure 2). Recently, Zhang et al.<sup>48</sup> used phylogenomic analyses to validate previous studies indicating the non-monophyly of some taxa. They suggested that *Catasticta* should be a junior synonym of *Archonias* and that *Eurema* is paraphyletic, among other issues. Regarding *Catasticta* and *Archonias*, for example, phylogenetic, morphological, and natural history traits had already suggested that these genera are closely related to each other.<sup>46,51,52</sup> However, sampling in Zhang et al.<sup>48</sup> was limited in some clades, and our study highlights additional issues. Namely, *Euchloe* is paraphyletic, *Rhabdodryas* is nested within *Phoebis*, *Glennia* is nested in *Ganyra*, and *Eurema* is still not monophyletic even after proposed changes by Zhang et al.<sup>48</sup>



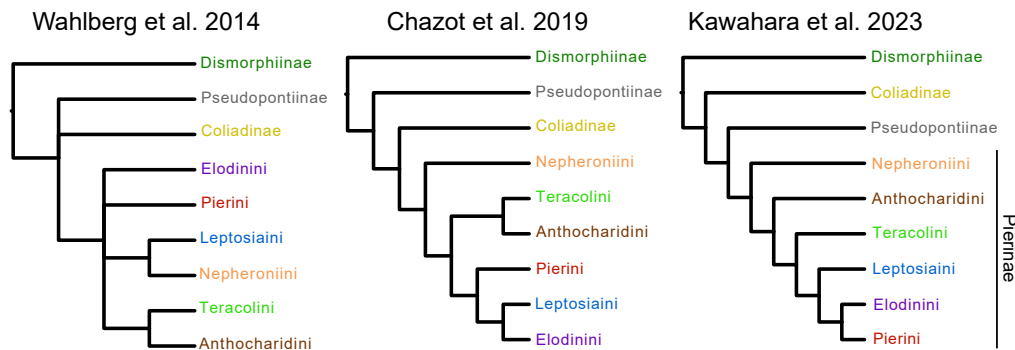
**Figure 1. Evolutionary relationships and diversification patterns of Pieridae**

Time-calibrated tree of 593 species. Branches with significant net diversification rate shifts as estimated by BAMM are indicated with red and blue circles. Annual mean temperature (BIO1, WorldClim) for each species is indicated with colored bars in the outer ring. Butterfly images are not to scale. Bottom left inset shows a cladogram of the subfamily and tribal relationships in this figure.

Divergence time estimates varied depending on the calibration scheme, with calibrations taken from Kawahara et al.<sup>53</sup> resulting in younger dates (Table S1). The crown age of Pieridae was between 69.6 and 64.5 Ma with a stem age between 83.5 and 76.1 Ma. Subfamily crown ages were estimated to be 35.0–32.7 Ma for Dismorphiinae; 49.8–44.2 Ma for Coliadinae; and 49.9–48.9 Ma for Pierinae (Table S1). The five Afro-tropical Pseudopontiinae species are on a long branch with extant taxa diversifying only 5.1–3.8 Ma (Table S1).

### Diversification patterns of Pieridae

Pierid lineages have steadily increased since the family's origin around 70 Ma, according to the deterministic lineage through time plot (dLTT) (Figure S2). In contrast, fitted pulled speciation rates (PSR) gradually decreased from the stem lineage until 40 Ma, when rates started to increase, followed by two peaks between 20 Ma and the present (Figure S3).



**Figure 2. Summary of the inferred relationships among higher-level Pieridae taxa in recent molecular phylogenetic studies.**

The BAMM (Bayesian Analysis of Macroevolutionary Mixtures) estimation of net diversification rates identified five clades with increased diversification rates: *Colias*; a subclade of *Catantacta* (+*Archonias*); *Delias*; and clades within *Pieris* and *Mylothris* (Figures 1 and S4). These patterns were consistent with our RevBayes analysis (Figure S5). The distinct shift configuration of BAMM suggests a high (>0.75) marginal probability of shifts for *Colias*, *Pareronia*, and a subclade of *Catantacta* (+*Archonias*) (Figure S4).

### Pieridae diversity and temperature through time

An ancestral state reconstruction of annual mean temperature indicates that the ancestor of Pieridae was likely found in warm climates. Switches to colder, temperate climates occurred multiple times, such as in the genera *Colias* and *Pieris* (Figure S6), and the reconstruction of minimum temperature in the coldest month shows pronounced habitat shifts to colder temperatures in these genera (Figure S7). There was a clear expansion toward habitats with greater annual temperature ranges (seasonality) in *Aporia*, *Colias*, and *Pieris* (Figure S8). Reconstructions of mean diurnal range (Figure S9) and maximum temperature in the warmest month (Figure S10) did not show pronounced variation across the tree.

The best-fit model explaining Pieridae diversification dynamics was a positive exponential relationship for both speciation and extinction rates with (paleo)temperature (BEnv.VarDEnv.Var\_EXPO) (Table S2). Global temperature generally decreased during the past ~40 My during which most of the Pieridae crown group evolved. Thus, the exponential relationships of speciation and extinction rates with paleoclimate mean these rates also decreased (slowed) with time toward the present. The best model fit for the largest Pieridae subfamily, Pierinae, describes a pattern in which speciation rate was constant but extinction rate decreased with paleotemperature (BCSTDEnv.Var\_EXPO), meaning that net diversification remained positive in the group.

### Present day temperature and Pieridae diversity

Our phylogenetic generalized least squares (PGLS) analyses found significant relationships between current temperature and speciation/extinction for two variables: (1) minimum temperature of the coldest month (positive for speciation, negative for extinction) and (2) annual temperature range (negative for speciation, positive for extinction). We also found a significant negative relationship between speciation and mean diurnal range (Figure S11). The annual mean temperature and maximum temperature of the warmest month was not significantly related to speciation or extinction.

When QuaSSE (Quantitative State Speciation and Extinction) was used to fit different likelihood functions of speciation on the tree for different temperature-related variables, the only significant ( $\Delta AIC > 2$ ) variables were (1) mean diurnal range and (2) maximum temperature in the warmest month (Table S3). We found a negative sigmoidal relationship with drift between speciation and these two variables (<https://doi.org/10.5061/dryad.c59zw3rbh>). This means that speciation rates are higher in habitats with low diurnal temperature ranges and with mild summer temperatures, but there is a steep decline in speciation rates when the diurnal range exceeds 11°C and, separately, when summer temperatures are higher than 31°C.

PGLS and QuaSSE both indicated a negative relationship between speciation and mean diurnal temperature range (Table S3). The PGLS analysis of maximum temperature of the warmest month was not significant, but QuaSSE suggests a negative correlation as the best fit for this variable, indicating that extremely high temperatures negatively affect speciation rates. The direction of correlations between speciation and minimum temperature of the coldest month and temperature annual range differed between the two methods. However, PGLS p values were significant while the QuaSSE  $\Delta AIC$  was less than 2. These results suggest that milder temperatures in the coldest month and less variable temperature throughout the year are correlated with higher speciation.

The discrepancies between PGLS and QuaSSE are likely an algorithmic artifact. The PGLS used tip rate estimations extracted from BAMM, while QuaSSE evaluated correlations using its own lambda (speciation) and mu (extinction) estimates. Additionally, QuaSSE fits the data to five different models while PGLS is limited to a linear regression, which may not be the best model for the data.

## DISCUSSION

### Temperature stability and warmer climates are associated with higher diversification over time

Paleotemperatures partially explain current pierid diversification patterns. Throughout their 70 My evolutionary history, the diversification rate of Pieridae is positively associated with warmer temperatures. The MRCA of the family likely inhabited warm climates, and speciation and extinction rates remained positive but have slowed along with global cooling that began in the late Eocene. Thus, higher diversification rates are associated with warmer temperatures through time.

Interestingly, our investigations of diversification dynamics in the most species-rich subfamily, Pierinae, are different from those of the family as a whole. Net diversification has increased since the Cenozoic as a result of decreasing extinction rates combined with constant speciation. Future research could further clarify these dynamics by examining smaller clades within Pieridae but with complete or nearly complete species-level sampling.

### Temperate climates are associated with higher speciation across space

From their likely tropical origins, there have been numerous shifts from tropical to temperate areas with greater temperature variability throughout the year, and several of these shifts are associated with increased diversification, notably in the genera *Colias* and *Pieris*. *Delias*, the most species-rich genus of butterflies, is often associated with equatorial habitats, but many species are found in environments with cooler temperatures (Figure S6) including tropical and subtropical mountains.

### High variability in daily temperatures is associated with decreased diversification

We employed several analyses to identify the specific climatic factors associated with diversification rate shifts. Different analyses gave conflicting results about which aspects of temperature play an outsized role in determining extant diversity. PGLS analyses concluded that (1) diversification increases in cold areas (increased speciation and decreased extinction associated with lowest temperature in the coldest month); (2) diversification declines in areas with large temperature fluctuations throughout the year (lower speciation and increased extinction associated with annual temperature range); and (3) speciation declines in areas with large temperature fluctuation throughout the 24 h day (mean diurnal range). Annual mean temperature and hottest temperature of the warmest month were not significant in these analyses. On the other hand, the QuaSSE analysis identified this latter variable as significant, along with mean diurnal range. The QuaSSE analysis indicated that speciation rates are higher in habitats with low daily temperature variation (<11°C) and with cool summer temperatures that rarely exceed 31°C. We therefore tentatively conclude that large daily fluctuations in temperature—a feature of temperate and montane habitats—are associated with increased speciation, a conclusion on which both analyses agree.

Seasonality influences species' distributions<sup>54</sup> and is thought to be correlated with butterfly species richness.<sup>55</sup> In general, biodiversity is greater in stable, tropical environments, and these areas are often characterized by high temperatures that are generally more constant throughout the year.<sup>56</sup> In warm seasons, sunlight and ambient temperature are the main factors influencing butterfly activity, and if conditions are not ideal, butterflies will shift the timing and length of their activity.<sup>57</sup> The diversification of white and yellow butterflies is higher in these environments with stable circadian temperatures. The climatic stability hypothesis<sup>13</sup> predicts that the tropics are especially diverse in part because the climate does not fluctuate much throughout the day or year. We found no support for this hypothesis in Pieridae. However, one corollary of this idea was supported: large daily temperature fluctuations (lack of stability) are associated with decreased diversification (specifically, decreased speciation).

Although warmer environments are often linked to life history traits that accelerate speciation,<sup>10–12</sup> we found that speciation rates are elevated in temperate environments. Butterflies have an optimal internal temperature range for survival but can tolerate greater environmental ranges because they have some ability to control their thermal physiology.<sup>31,42,58</sup> Nonetheless, anything below or above the tolerable range can be lethal.<sup>31</sup> The optimal temperature range usually trends toward warmer temperatures, and most species are active during warm seasons. In some cases, the warmest month is the only time of the year when butterflies can reproduce. This is especially true for Arctic and alpine species. However, extremely warm temperatures can reduce activity in some species.<sup>57</sup> Recognizing these trends is crucial in light of the current, unprecedentedly rapid increase in global temperatures during the Anthropocene, particularly our evidence that speciation rates decrease with increasing temperature.

Franzén et al.<sup>57</sup> found that the lifespan of the Palearctic papilionid *Parnassius apollo* is negatively affected by increased median temperatures and that the lycaenid *Phengaris arion* lives longest at an optimal temperature of ca. 27°C, suggesting stabilizing selection on temperature. In general, if temperatures are extreme, behavior and morphology/wing coloration will adapt to regulate body temperature.<sup>40,59–61</sup> Butterflies use different strategies and mechanisms to maintain body temperature, which can disrupt their regular behavior.<sup>42</sup> They may, for example, find a shaded area to hide,<sup>31</sup> which then prevents them from feeding and finding mates.

There is a uniform distribution of taxa in warm regions with mild winters across the Pieridae phylogeny (Figures S6 and S7), which is unsurprising considering Pieridae are generally most diverse in warm climates (as in other butterfly families<sup>62</sup>). However, there are multiple independent origins of cold-adapted species (e.g., *Baltia* spp. and *Reliquia santamarta*) including clades that are sister to taxa inhabiting warm or mild climates (such as the clade that includes *Tatochila*, *Hyposchila*, *Theochila*, *Pierphulia*, *Phulia*, and *Infraphulia*). Some of these shifts are associated with adaptations to habitats at higher elevations, as shown by the convergence of morphology among groups that inhabit similarly cold, high-elevation habitats around the world.<sup>39,43</sup> Our results confirm the recent finding<sup>48</sup> that the monotypic genus *Reliquia*, endemic to a single mountain range in tropical Colombia is nested in the mostly Holarctic genus *Pontia* (including *Baltia*). *Pontia* inhabits colder environs at similarly high elevations in the northern hemisphere. This observation partially uncovers the intriguing interplay of Pieridae biogeography and

adaptation to high elevations around the world. Furthermore, our BAMM analysis highlighted heightened speciation rates in some temperate pierid clades.

### Phylogenetic relationships of Pieridae

Our study provides a robust phylogenetic framework for pierid butterflies and a foundation for testing many macroevolutionary questions. We show, with strong support (UFBoot and SH-aLRT = 100), that Dismorphiinae are sister to a clade containing Coliadinae, Pseudopontiinae, and Pierinae, in agreement with Wahlberg et al.<sup>44</sup> and Kawahara et al.<sup>25</sup> All tribal relationships of Pierinae were strongly supported and differ from previous studies.<sup>25,44,50</sup> All of the Pieridae sequence data used by Kawahara et al.<sup>25</sup> are incorporated into our dataset, but the inferred trees differ in tribal topology (Figures 2 and S1). This is likely a result of our considerably increased taxon sampling.

We recovered the subfamily Pseudopontiinae on an extremely long branch (as shown previously in the study of Mitter et al.<sup>63</sup>) of approximately 50 My in length (Table S1), which likely contributes to the difficulty of confidently inferring the phylogenetic placement of this subfamily. Topological uncertainty hinders inference of ancestral states of the group as well. The morphologically unique Afrotropical subfamily Pseudopontiinae was thought to only include a single species until 2011,<sup>43,63</sup> when molecular data demonstrated the existence of four additional cryptic species. Even though we incorporated additional molecular data for several species in the group, support for its position in relation to the other subfamilies is still low (as in Espeland et al.<sup>49</sup> and Chazot et al.<sup>50</sup>). We also find that several genera, including *Euchloe* and *Ganyra*, are not monophyletic (Figure 1), and will require further taxonomic investigations.

The Pieridae crown age of around 70 Ma overlaps the K-Pg (Cretaceous–Paleogene) boundary, a major extinction event. The lineage giving rise to the family pre-dated this event, and pierids started diversifying during the Paleogene (Figure S2).

### Limitations of the study

Divergence time estimation for butterflies has been historically difficult due to the notable underrepresentation of butterflies in the fossil record. We included an analysis with the most rigorously justified relevant butterfly fossils from recent works as well as a set of analyses using secondary calibrations but consider that the use of secondary calibrations themselves is a limitation on our downstream analyses. Future discovery of butterfly fossils will ideally be needed to help solidify our understanding of butterfly divergence times. Further, some of the analyses we conducted do not allow the fraction of missing species to be specified (the sampling fraction), and thus our ~50% sampling of Pieridae could have biased some results.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Molecular data
  - Distribution and bioclimatic data
  - Diversification and correlation analyses

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2024.109336>.

### ACKNOWLEDGMENTS

David Plotkin provided suggestions on phylogenetic dating, Caroline Storer helped molecular data assembly, and Martin Wiemers helped curate distribution data. Ana B.B. Morais and Yutaka Inayoshi provided samples. Fabien Condamine and Robert Guralnick provided suggestions on methods. Stilianos Louca assisted with the inhomogenous time grid in castor. Liam Revell provided the code for plotting temperature bars in Figure 1. We thank the NCBS Sequencing Facility for sequencing an Indian sample. We acknowledge Research Computing at the University of Florida for providing computational resources (<https://www.rc.ufl.edu>).

Funding came from the US National Science Foundation (NSF) GoLife “ButterflyNet” collaborative grant (DEB-1541500, 1541557, 1541560) to A.Y.K., D.J.L., and N.E.P. Specimen collection and preservation was funded by NSF DBI-1349345, 1601369, DEB-1557007, and IOS-1920895 (A.Y.K.), NSF DEB-1120380 (D.J.L.), Grants 9285-13 and WW-227R-17 from the National Geographic Society (D.J.L.), NSF DBI-1256742 (A.Y.K. and K.R.W.), NSF DEB-0639861 (K.R.W.) and NSF SES-0750480, DEB-0447244, and DEB-9615760 to NEP, grant 204308 from the Research Council of Norway (M.E.), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grants 200814/2015-0 (A.P.S.C.),

563332/2010-7 and 304273/2014-7 (H.P.R. and N.O.M.), and Ramanujan Fellowship from the Dept. of Science and Technology, Govt. of India from National Centre for Biological Sciences (K.K.).

## AUTHOR CONTRIBUTIONS

A.P.S.C., A.Y.K., D.J.L., H.L.O., and N.E.P. designed the research; J.W.B., A.P.S.C., K.M.D., C.E., M.E., A.G.K., D.J.L., R.L.M., H.L.O., and R.A.S. performed the research; K.A.P., M.F.B., S.C.C., M.E., N.T.H., S.H., Y.F.H., A.G.K., K.K., D.J.L., D.J.M., N.O.M., S.M., D.P., H.P.R., S.S., R.V., H.W., and K.R.W. conducted sampling and prepared specimens; K.A.P., M.F.B., A.P.S.C., S.C.C., C.E., M.E., N.T.H., S.H., Y.F.H., A.G.K., K.K., D.J.L., D.J.M., N.O.M., R.L.M., S.M., D.P., H.P.R., S.S., R.V., H.W., and K.R.W. identified specimens; A.P.S.C., C.E., H.L.O., and R.A.S. analyzed data; and M.F.B., A.P.S.C., C.E., A.Y.K., D.J.L., H.L.O., N.E.P., and R.A.S. wrote the paper with input from all authors.

## DECLARATION OF INTERESTS

The authors declare no conflict of interest.

Received: July 31, 2023

Revised: October 28, 2023

Accepted: February 21, 2024

Published: February 28, 2024

## REFERENCES

- Gaston, K.J. (2000). Global patterns in biodiversity. *Nature* 405, 220–227. <https://doi.org/10.1038/35012228>.
- Rohde, K. (1992). Latitudinal gradients in species diversity: the search for the primary cause. *Oikos* 65, 514–527. <https://doi.org/10.2307/3545569>.
- Hawkins, B.A., Porter, E.E., and Felizola Diniz-Filho, J.A. (2003). Productivity and history as predictors of the latitudinal diversity gradient of terrestrial birds. *Ecology* 84, 1608–1623. [https://doi.org/10.1890/0012-9658\(2003\)084](https://doi.org/10.1890/0012-9658(2003)084).
- Currie, D.J., Mittelbach, G.G., Cornell, H.V., Field, R., Guégan, J., Hawkins, B.A., Kaufman, D.M., Kerr, J.T., Oberdorff, T., O'Brien, E., and Turner, J.R.G. (2004). Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecol. Lett.* 7, 1121–1134. <https://doi.org/10.1111/j.1461-0248.2004.00671.x>.
- Schluter, D. (2016). Speciation, ecological opportunity, and latitude. *Am. Nat.* 187, 1–18. <https://doi.org/10.1086/684193>.
- Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guégan, J.F., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M., et al. (2003). Energy, water, and broad-scale geographic patterns of species richness. *Ecology* 84, 3105–3117. <https://doi.org/10.1890/03-8006>.
- Sanders, N.J., Lessard, J.-P., Fitzpatrick, M.C., and Dunn, R.R. (2007). Temperature, but not productivity or geometry, predicts elevational diversity gradients in ants across spatial grains. *Glob. Ecol. Biogeogr.* 16, 640–649. <https://doi.org/10.1111/j.1466-8238.2007.00316.x>.
- Zhang, Q.-G., Lu, H.-S., and Buckling, A. (2018). Temperature drives diversification in a model adaptive radiation. *Proc. Biol. Sci.* 285, 20181515. <https://doi.org/10.1098/rspb.2018.1515>.
- Cerezer, F.O., de Azevedo, R.A., Nascimento, M.A.S., Franklin, E., de Moraes, J.W., and de Sales Dambros, C. (2020). Latitudinal gradient of termite diversity indicates higher diversification and narrower thermal niches in the tropics. *Glob. Ecol. Biogeogr.* 29, 1967–1977. <https://doi.org/10.1111/geb.13167>.
- Allen, A.P., Brown, J.H., and Gillooly, J.F. (2002). Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science* 297, 1545–1548. <https://doi.org/10.1126/science.1072380>.
- Gillooly, J.F., Allen, A.P., West, G.B., and Brown, J.H. (2005). The rate of DNA evolution: Effects of body size and temperature on the molecular clock. *Proc. Natl. Acad. Sci. USA* 102, 140–145. <https://doi.org/10.1073/pnas.0407735101>.
- Wright, S., Keeling, J., and Gillman, L. (2006). The road from Santa Rosalia: a faster tempo of evolution in tropical climates. *Proc. Natl. Acad. Sci. USA* 103, 7718–7722. <https://doi.org/10.1073/pnas.0510383103>.
- Klopfer, P.H. (1959). Environmental determinants of faunal diversity. *Am. Nat.* 93, 337–342. <https://doi.org/10.1086/282092>.
- Owen, J.G. (1989). Patterns of herpetofaunal species richness: relation to temperature, precipitation, and variance in elevation. *J. Biogeogr.* 16, 141–150. <https://doi.org/10.2307/2845088>.
- Lennon, J.J., Greenwood, J.J.D., and Turner, J.R.G. (2000). Bird diversity and environmental gradients in Britain: a test of the species–energy hypothesis. *J. Anim. Ecol.* 69, 581–598. <https://doi.org/10.1046/j.1365-2656.2000.00418.x>.
- Kozak, K.H., and Wiens, J.J. (2010). Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecol. Lett.* 13, 1378–1389. <https://doi.org/10.1111/j.1461-0248.2010.01530.x>.
- Rolland, J., Condamine, F.L., Jiguet, F., and Morlon, H. (2014). Faster speciation and reduced extinction in the Tropics contribute to the mammalian Latitudinal Diversity Gradient. *PLoS Biol.* 12, e1001775. <https://doi.org/10.1371/journal.pbio.1001775>.
- Gómez-Rodríguez, C., Baselga, A., and Wiens, J.J. (2015). Is diversification rate related to climatic niche width? *Glob. Ecol. Biogeogr.* 24, 383–395. <https://doi.org/10.1111/geb.12229>.
- Condamine, F.L., Rolland, J., and Morlon, H. (2019). Assessing the causes of diversification slowdowns: temperature-dependent and diversity-dependent models receive equivalent support. *Ecol. Lett.* 22, 1900–1912.
- García-Porta, J., Irisarri, I., Kirchner, M., Rodríguez, A., Kirchoff, S., Brown, J.L., MacLeod, A., Turner, A.P., Ahmadzadeh, F., Albaladejo, G., et al. (2019). Environmental temperatures shape thermal physiology as well as diversification and genome-wide substitution rates in lizards. *Nat. Commun.* 10, 4077. <https://doi.org/10.1038/s41467-019-11943-x>.
- Kong, H., Condamine, F.L., Yang, L., Harris, A.J., Feng, C., Wen, F., and Wang, M. (2022). Phylogenomic and Macroevolutionary Evidence for an Explosive Radiation of a Plant Genus in the Miocene. *Syst. Biol.* 71, 589–609. <https://doi.org/10.1093/sysbio/syab068>.
- Condamine, F.L., Silvestro, D., Koppelhus, E.B., and Antonelli, A. (2020). The rise of angiosperms pushed conifers to decline during global cooling. *Proc. Natl. Acad. Sci. USA* 117, 28867–28875. <https://doi.org/10.1073/pnas.2005571117>.
- Sayle, M.H. (1928). *The metabolism of insects*. *Q. Rev. Biol.* 3, 542–553.
- Ehrlich, P.R., and Raven, P.H. (1965). *Butterflies and plants: a study of coevolution*. *Evolution* 18, 586–608.
- Kawahara, A.Y., Storer, C., Carvalho, A.P.S., Plotkin, D.M., Condamine, F.L., Braga, M.P., Ellis, E.A., St Laurent, R.A., Li, X., Barve, V., et al. (2023). A global phylogeny of butterflies reveals their evolutionary history, ancestral hosts and biogeographic origins. *Nat. Ecol. Evol.* 7, 903–913. <https://doi.org/10.1038/s41559-023-02041-9>.
- Wheat, C.W., Vogel, H., Wittstock, U., Braby, M.F., Underwood, D., and Mitchell-Olds, T. (2007). The genetic basis of a plant–insect coevolutionary key innovation. *Proc. Natl. Acad. Sci. USA* 104, 20427–20431. <https://doi.org/10.1073/pnas.0706229104>.
- Edger, P.P., Heide-Fischer, H.M., Bekaert, M., Rota, J., Glöckner, G., Platts, A.E., Heckel, D.G., Der, J.P., Wafula, E.K., Tang,



- M., et al. (2015). The butterfly plant arms-race escalated by gene and genome duplications. *Proc. Natl. Acad. Sci. USA*, 112, 8362–8366. <https://doi.org/10.1073/pnas.1503926112>.
28. Condamine, F.L., Sperling, F.A.H., Wahlberg, N., Rasplus, J.Y., and Kergoat, G.J. (2012). What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecol. Lett.* 15, 267–277. <https://doi.org/10.1111/j.1461-0248.2011.01737.x>.
  29. Turner, J.R.G., Gatehouse, C.M., and Corey, C.A. (1987). Does solar energy control organic diversity? Butterflies, moths and the British climate. *Oikos* 48, 195. <https://doi.org/10.2307/3565855>.
  30. Kristensen, N.P. (2003). In *Handbook of Zoology, Volume IV, Arthropoda: Insecta, Part 36. Lepidoptera, Moths and Butterflies, Volume 2: Morphology, Physiology, and Development*, N.P. Kristensen, ed. (De Gruyter).
  31. Clench, H.K. (1966). Behavioral thermoregulation in butterflies. *Ecology* 47, 1021–1034. <https://doi.org/10.2307/1935649>.
  32. Caporale, A., Romanowski, H.P., and Mega, N.O. (2017). Winter is coming: Diapause in the subtropical swallowtail butterfly *Euryades corethrus* (Lepidoptera, Papilionidae) is triggered by the shortening of day length and reinforced by low temperatures. *J. Exp. Zool. A Ecol. Integr. Physiol.* 327, 182–188. <https://doi.org/10.1002/jez.2091>.
  33. Klockmann, M., Kleinschmidt, F., and Fischer, K. (2017). Carried over: heat stress in the egg stage reduces subsequent performance in a butterfly. *PLoS One* 12, e0180968. <https://doi.org/10.1371/journal.pone.0180968>.
  34. Lamas, G. (2015). *Catalog of the Butterflies (Papilionoidea)*.
  35. Ackery, P.R., de Jong, R., and Vane-Wright, R.I. (1998). The butterflies: Hedyloidea, Hesperioidea and Papilionoidea. In *Handbook of Zoology, Volume IV, Arthropoda: Insecta, Part 35. Lepidoptera, Moths and Butterflies, Volume 1: Evolution, Systematics, and Biogeography*, 4. *Arthropoda: Insecta, Part 35, N.P. Kristensen, ed. (Walter de Gruyter)*, pp. 263–300.
  36. Shapiro, A.M., Forister, M.L., and Fordyce, J.A. (2007). Extreme high-altitude Asian and Andean pierid butterflies are not each others' closest relatives. *Arct. Antarct. Alp. Res.* 39, 137–142. <https://doi.org/10.1657/1523-0430>.
  37. Schmidt, B.C. (2018). Diagnosis and distribution of the Kivalliq sulphur (*Colias rankinensis* Verhulst), an enigmatic Arctic butterfly. *J. Lepid. Soc.* 72, 227–232. <https://doi.org/10.18473/lepi.v72i3.a8>.
  38. El-Gabbas, A., and Gilbert, F. (2016). The desert beauty *Calopieris eulimene*: a butterfly new to Egypt (Insecta: Lepidoptera). *Zool. Middle East* 62, 279–281. <https://doi.org/10.1080/09397140.2016.1202984>.
  39. Dufour, P.C., Willmott, K.R., Padrón, P.S., Xing, S., Bonebrake, T.C., and Scheffers, B.R. (2018). Divergent melanism strategies in Andean butterfly communities structure diversity patterns and climate responses. *J. Biogeogr.* 45, 2471–2482. <https://doi.org/10.1111/jbi.13433>.
  40. Watt, W.B. (1969). Adaptive significance of pigment polymorphisms in *Colias* butterflies, ii. Thermoregulation and photoperiodically controlled melanin variation in *Colias eurytheme*. *Proc. Natl. Acad. Sci. USA* 63, 767–774. <https://doi.org/10.1073/pnas.63.3.767>.
  41. Kingsolver, J.G. (1985). Thermoregulatory significance of wing melanization in *Pieris* butterflies (Lepidoptera: Pieridae): physics, posture, and pattern. *Oecologia* 66, 546–553.
  42. Tsai, C.-C., Childers, R.A., Nan Shi, N., Ren, C., Pelaez, J.N., Bernard, G.D., Pierce, N.E., and Yu, N. (2020). Physical and behavioral adaptations to prevent overheating of the living wings of butterflies. *Nat. Commun.* 11, 551. <https://doi.org/10.1038/s41467-020-14408-8>.
  43. Braby, M.F., Vila, R., and Pierce, N.E. (2006). Molecular phylogeny and systematics of the Pieridae (Lepidoptera: Papilionoidea): higher classification and biogeography. *Zool. J. Linn. Soc.* 147, 239–275. <https://doi.org/10.1111/j.1096-3642.2006.00218.x>.
  44. Wahlberg, N., Rota, J., Braby, M.F., Pierce, N.E., and Wheat, C.W. (2014). Revised systematics and higher classification of pierid butterflies (Lepidoptera: Pieridae) based on molecular data. *Zool. Scr.* 43, 641–650. <https://doi.org/10.1111/zsc.12075>.
  45. Braby, M.F., and Pierce, N.E. (2007). Systematics, biogeography and diversification of the Indo-Australian genus *Delias* Hübner (Lepidoptera: Pieridae): phylogenetic evidence supports an 'out-of-Australia' origin. *Syst. Entomol.* 32, 2–25. <https://doi.org/10.1111/j.1365-3113.2006.00349.x>.
  46. Braby, M.F., Pierce, N.E., and Vila, R. (2007). Phylogeny and historical biogeography of the subtribe Aporiina (Lepidoptera: Pieridae): implications for the origin of Australian butterflies. *Biol. J. Linn. Soc.* 90, 413–440. <https://doi.org/10.1111/j.1095-8312.2007.00732.x>.
  47. Warren-Gash, H., Aduse-Poku, K., Murillo-Ramos, L., and Wahlberg, N. (2020). Systematics and evolution of the African butterfly genus *Mylothris* (Lepidoptera, Pieridae). *Nota Lepidopterol.* 43, 1–14. <https://doi.org/10.3897/nl.43.46354>.
  48. Zhang, J., Cong, Q., Shen, J., Opler, P.A., and Grishin, N.V. (2021). Genomics-guided refinement of butterfly taxonomy. *Taxon. Rep. Int. Lepid. Surv.* 9, 1–54.
  49. Espeland, M., Breinholt, J., Willmott, K.R., Warren, A.D., Vila, R., Toussaint, E.F.A., Maunsell, S.C., Aduse-Poku, K., Talavera, G., Eastwood, R., et al. (2018). A comprehensive and dated phylogenomic analysis of butterflies. *Curr. Biol.* 28, 770–778.e5.
  50. Chazot, N., Wahlberg, N., Freitas, A.V.L., Mitter, C., Labandeira, C., Sohn, J.-C., Sahoo, R.K., Seraphim, N., de Jong, R., and Heikkilä, M. (2019). Priors and posteriors in Bayesian timing of divergence analyses: the age of butterflies revisited. *Syst. Biol.* 68, 797–813. <https://doi.org/10.1093/sysbio/syz002>.
  51. Braby, M.F., and Nishida, K. (2010). The immature stages, larval food plants and biology of Neotropical mistletoe butterflies (Lepidoptera: Pieridae). II. The *Catacticta* group (Pierini: Aporiina). *J. Nat. Hist.* 44, 1831–1928. <https://doi.org/10.1080/00222931003633227>.
  52. Mota, L.L., Silva, A.K., Freitas, A.V.L., and Kaminski, L.A. (2016). Immature stages of *Archonias brassolis tereas* (Godart) (Pieridae: Pierini), with notes on interspecific interactions between mistletoe butterflies. *J. Lepid. Soc.* 70, 289–294. <https://doi.org/10.18473/lepi.70i4.a7>.
  53. Kawahara, A.Y., Plotkin, D., Espeland, M., Meusemann, K., Toussaint, E.F.A., Donath, A., Gimmich, F., Frandsen, P.B., Zwick, A., dos Reis, M., et al. (2019). Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *Proc. Natl. Acad. Sci. USA* 116, 22657–22663. <https://doi.org/10.1073/pnas.1907847116>.
  54. Sheldon, K.S., and Tewksbury, J.J. (2014). The impact of seasonality in temperature on thermal tolerance and elevational range size. *Ecology* 95, 2134–2143. <https://doi.org/10.1890/13-1703.1>.
  55. Pyrz, T.W., Willmott, K., Garlacz, R., Boyer, P., and Gareca, Y. (2014). Latitudinal gradient and spatial covariance in species richness of tropical Lepidoptera in the Andes. *Insect Conserv. Divers.* 7, 355–364. <https://doi.org/10.1111/icad.12058>.
  56. Fischer, A.G. (1960). Latitudinal variations in organic diversity. *Evolution* 14, 64–81. <https://doi.org/10.1111/j.1558-5646.1960.tb03057.x>.
  57. Franzén, M., Francioli, Y., Askling, J., Kindvall, O., Johansson, V., and Forsman, A. (2022). Differences in phenology, daily timing of activity, and associations of temperature utilization with survival in three threatened butterflies. *Sci. Rep.* 12, 7534. <https://doi.org/10.1038/s41598-022-10676-0>.
  58. Bladon, A.J., Lewis, M., Bladon, E.K., Buckton, S.J., Corbett, S., Ewing, S.R., Hayes, M.P., Hitchcock, G.E., Knock, R., Lucas, C., et al. (2020). How butterflies keep their cool: physical and ecological traits influence thermoregulatory ability and population trends. *J. Anim. Ecol.* 89, 2440–2450. <https://doi.org/10.1111/1365-2656.13319>.
  59. Shapiro, A.M. (1976). Seasonal Polyphenism. In *Evolutionary Biology*, M.K. Hecht, W.C. Steere, and B. Wallace, eds. (Springer US), pp. 259–333. [https://doi.org/10.1007/978-1-4615-6950-3\\_6](https://doi.org/10.1007/978-1-4615-6950-3_6).
  60. Kingsolver, J.G., and Wiernasz, D.C. (1991). Seasonal polyphenism in wing-melanin pattern and thermoregulatory adaptation in *Pieris* butterflies. *Am. Nat.* 137, 816–830.
  61. Gautam, S., and Kunte, K. (2020). Adaptive plasticity in wing melanisation of a montane butterfly across a Himalayan elevational gradient. *Ecol. Entomol.* 45, 1272–1283. <https://doi.org/10.1016/j.een.12911>.
  62. Espeland, M., Hall, J.P.W., DeVries, P.J., Lees, D.C., Cornwall, M., Hsu, Y.-F., Wu, L.W., Campbell, D.L., Talavera, G., Vila, R., et al. (2015). Ancient Neotropical origin and recent recolonisation: Phylogeny, biogeography and diversification of the Riodinidae (Lepidoptera: Papilionoidea). *Mol. Phylogenet. Evol.* 93, 296–306. <https://doi.org/10.1016/j.ympev.2015.08.006>.
  63. Mitter, K.T., Larsen, T.B., De Prins, W., De Prins, J., Collins, S., Vande Weghe, G., Sáfián, S., Zakharov, E.V., Hawthorne, D.J., Kawahara, A.Y., and Regier, J.C. (2011). The butterfly subfamily Pseudopontiinae is not monobasic: marked genetic diversity and morphology reveal three new species of *Pseudopontia* (Lepidoptera: Pieridae). *Syst. Entomol.* 36, 139–163. <https://doi.org/10.1111/j.1365-3113.2010.00549.x>.

64. Katoh, K., and Standley, D.M. (2013). MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. <https://doi.org/10.1093/molbev/mst010>.
65. Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30, 3276–3278. <https://doi.org/10.1093/bioinformatics/btu531>.
66. Kück, P., and Meusemann, K. (2010). FASconCAT: convenient handling of data matrices. *Mol. Phylogenet. Evol.* 56, 1115–1118. <https://doi.org/10.1016/j.ympev.2010.04.024>.
67. Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274.
68. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534.
69. Smith, S.A., and O’Meara, B.C. (2012). TreePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* 28, 2689–2690.
70. Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.
71. Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. <https://doi.org/10.1093/molbev/mss075>.
72. Owens, H.L., Merow, C., Maitner, B.S., Kass, J.M., Barve, V., and Guralnick, R.P. (2021). occCite: Tools for querying and managing large biodiversity occurrence datasets. *Ecography* 44, 1228–1235. <https://doi.org/10.1111/ecog.05618>.
73. Chamberlain, S., Barve, V., Mcglinn, D., Oldoni, D., Desmet, P., Geffert, L., and Ram, K. (2022). rgbif: Interface to the Global Biodiversity Information Facility API.
74. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis.
75. Hijmans, R.J. (2021). raster: geographic data analysis and modeling. R package version 3, pp. 4–10.
76. Rabosky, D.L., Grundler, M., Anderson, C., Title, P., Shi, J.J., Brown, J.W., Huang, H., and Larson, J.G. (2014). BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.* 5, 701–707. <https://doi.org/10.1111/2041-210X.12199>.
77. Höhna, S., Landis, M.J., Heath, T.A., Boussau, B., Lartillot, N., Moore, B.R., Huelsenbeck, J.P., and Ronquist, F. (2016). RevBayes: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Syst. Biol.* 65, 726–736. <https://doi.org/10.1093/sysbio/syw021>.
78. Louca, S., and Doebeli, M. (2018). Efficient comparative phylogenetics on large trees. *Bioinformatics* 34, 1053–1055.
79. Pinheiro, J., and Bates, D.; R Core Team (2022). nlme: Linear and Nonlinear Mixed Effects Models.
80. FitzJohn, R.G. (2012). Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol. Evol.* 3, 1084–1092. <https://doi.org/10.1111/j.2041-210X.2012.00234.x>.
81. Krueger, F. (2015). Trim Galore Version 0.4.0.
82. Kawahara, A.Y., Breinholt, J.W., Espeland, M., Storer, C., Plotkin, D., Dexter, K.M., Toussaint, E.F.A., St Laurent, R.A., Brehm, G., Vargas, S., et al. (2018). Phylogenetics of moth-like butterflies (Papilionoidea: Hedylidae) based on a new 13-locus target capture probe set. *Mol. Phylogenet. Evol.* 127, 600–605. <https://doi.org/10.1016/j.ympev.2018.06.002>.
83. Wahlberg, N., and Wheat, C.W. (2008). Genomic outposts serve the phylogenomic pioneers: designing novel nuclear markers for genomic DNA extractions of Lepidoptera. *Syst. Biol.* 57, 231–242.
84. Breinholt, J.W., Earl, C., Lemmon, A.R., Lemmon, E.M., Xiao, L., and Kawahara, A.Y. (2018). Resolving relationships among the megadiverse butterflies and moths with a novel pipeline for Anchored Phylogenomics. *Syst. Biol.* 67, 78–93. <https://doi.org/10.1093/sysbio/syx048>.
85. Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>.
86. Chang, Z., Li, G., Liu, J., Zhang, Y., Ashby, C., Liu, D., Cramer, C.L., and Huang, X. (2015). Bridger: a new framework for de novo transcriptome assembly using RNA-seq data. *Genome Biol.* 16, 30.
87. Zhan, S., Merlin, C., Boore, J.L., and Reppert, S.M. (2011). The monarch butterfly genome yields insights into long-distance migration. *Cell* 147, 1171–1185. <https://doi.org/10.1016/j.cell.2011.09.052>.
88. Talavera, G., Lukhtanov, V., Pierce, N.E., and Vila, R. (2022). DNA Barcodes combined with multilocus data of representative taxa can generate reliable higher-level phylogenies. *Syst. Biol.* 71, 382–395. <https://doi.org/10.1093/sysbio/syab038>.
89. Minh, B.Q., Nguyen, M.A.T., and von Haeseler, A. (2013). Ultrafast approximation for phylogenetic Bootstrap. *Mol. Biol. Evol.* 30, 1188–1195.
90. St Laurent, R.A., Carvalho, A.P.S., Earl, C., and Kawahara, A.Y. (2021). Food plant shifts drive the diversification of sack-bearer moths. *Am. Nat.* 198, E170–E184. <https://doi.org/10.1086/716661>.
91. Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L.L., and Hernández-Hernández, T. (2015). A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytol.* 207, 437–453. <https://doi.org/10.1111/nph.13264>.
92. Jong, R.D. (2017). Fossil butterflies, calibration points and the molecular clock (Lepidoptera: Papilionoidea). *Zootaxa* 4270, 1–63.
93. GBIF: The Global Biodiversity Information Facility (2022). What is GBIF?. <https://www.gbif.org/what-is-gbif>.
94. Dincă, V., Lukhtanov, V.A., Talavera, G., and Vila, R. (2011). Unexpected layers of cryptic diversity in wood white *Leptidea* butterflies. *Nat. Commun.* 2, 324. <https://doi.org/10.1038/ncomms1329>.
95. Fick, S.E., and Hijmans, R.J. (2017). WorldClim 2: new 1 km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315.
96. Carnaval, A.C., Hickerson, M.J., Haddad, C.F.B., Rodrigues, M.T., and Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic Forest hotspot. *Science* 323, 785–789. <https://doi.org/10.1126/science.1166955>.
97. Höhna, S., Freyman, W.A., Nolen, Z., Huelsenbeck, J.P., May, M.R., and Moore, B.R. (2019). A bayesian approach for estimating branch-specific speciation and extinction rates. Preprint at bioRxiv. <https://doi.org/10.1101/555805>.
98. Louca, S., and Pennell, M.W. (2020). Extant timetrees are consistent with a myriad of diversification histories. *Nature* 580, 502–505. <https://doi.org/10.1038/s41586-020-2176-1>.
99. Zhang, X., Landis, J.B., Sun, Y., Zhang, H., Lin, N., Kuang, T., Huang, X., Deng, T., Wang, H., and Sun, H. (2021). Macroevolutionary pattern of *Saussurea* (Asteraceae) provides insights into the drivers of radiating diversification. *Proc. Biol. Sci.* 288, 20211575. <https://doi.org/10.1098/rspb.2021.1575>.
100. Condamine, F.L., Antonelli, A., Lagomarsino, L.P., Hoorn, C., and Liow, L.H. (2018). Teasing Apart Mountain Uplift, Climate Change and Biotic Drivers of Species Diversification. In *Mountains, Climate and Biodiversity* (John Wiley & Sons Ltd), pp. 257–272.
101. Morlon, H., Parsons, T.L., and Plotkin, J.B. (2011). Reconciling molecular phylogenies with the fossil record. *Proc. Natl. Acad. Sci. USA* 108, 16327–16332. <https://doi.org/10.1073/pnas.1102543108>.
102. Stadler, T. (2011). Mammalian phylogeny reveals recent diversification rate shifts. *Proc. Natl. Acad. Sci. USA* 108, 6187–6192. <https://doi.org/10.1073/pnas.1016876108>.
103. Condamine, F.L., Rolland, J., and Morlon, H. (2013). Macroevolutionary perspectives to environmental change. *Ecol. Lett.* 16, 72–85. <https://doi.org/10.1111/ele.12062>.
104. Gamsch, A. (2020). Technical comment on Condamine et al. (2019): a cautionary note for users of linear diversification dependencies. *Ecol. Lett.* 23, 1169–1171. <https://doi.org/10.1111/ele.13483>.
105. Morlon, H., Rolland, J., and Condamine, F.L. (2020). Response to technical comment ‘A cautionary note for users of linear diversification dependencies. *Ecol. Lett.* 23, 1172–1174. <https://doi.org/10.1111/ele.13513>.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Biological samples</b>		
Lepidoptera tissue samples	This paper	<a href="https://doi.org/10.5061/dryad.c59zw3rbh">https://doi.org/10.5061/dryad.c59zw3rbh</a>
<b>Deposited data</b>		
Alignments used for phylogenetic analyses	This paper	<a href="https://doi.org/10.5061/dryad.c59zw3rbh">https://doi.org/10.5061/dryad.c59zw3rbh</a>
Trees produced from all analyses	This paper	<a href="https://doi.org/10.5061/dryad.c59zw3rbh">https://doi.org/10.5061/dryad.c59zw3rbh</a>
Published sequence data	Various	<a href="https://doi.org/10.5061/dryad.c59zw3rbh">https://doi.org/10.5061/dryad.c59zw3rbh</a>
<b>Software and algorithms</b>		
GeneDumper v.0.8	This paper	<a href="https://github.com/sunray1/genedumper">https://github.com/sunray1/genedumper</a>
MAFFT v.7.0.1	Katoh and Standley <sup>64</sup>	<a href="https://mafft.cbrc.jp/alignment/software/">https://mafft.cbrc.jp/alignment/software/</a>
AliView v.1.27	Larsson <sup>65</sup>	<a href="https://ormbunkar.se/aliview/">https://ormbunkar.se/aliview/</a>
FASconCAT-G v.1.0.4	Kück and Meusemann <sup>66</sup>	<a href="https://bonn.leibniz-lib.de/en/research/research-centres-andgroups/fasconcat-g">https://bonn.leibniz-lib.de/en/research/research-centres-andgroups/fasconcat-g</a>
IQ-TREE v.2.0.3	Minh et al. <sup>67</sup> ; Nguyen et al. <sup>68</sup>	<a href="http://www.iqtree.org">http://www.iqtree.org</a>
TreePL	Smith and O'Meara <sup>69</sup>	<a href="https://github.com/blackrim/treePL">https://github.com/blackrim/treePL</a>
RAxML v.8.2.12	Stamatakis <sup>70</sup>	<a href="https://github.com/amkozlov/raxml-ng">https://github.com/amkozlov/raxml-ng</a>
TreeAnnotator v.1.10.4	Drummond et al. <sup>71</sup>	<a href="https://beast.community/treeannotator">https://beast.community/treeannotator</a>
occCite v.0. 4.9	Owens et al. <sup>72</sup>	<a href="https://cran.r-project.org/web/packages/occCite/index.html">https://cran.r-project.org/web/packages/occCite/index.html</a>
rgbif	Chamberlain et al. <sup>73</sup>	<a href="https://cran.r-project.org/web/packages/rgbif/index.html">https://cran.r-project.org/web/packages/rgbif/index.html</a>
ggplot2 v.3.3.5	Wickham <sup>74</sup>	<a href="https://ggplot2.tidyverse.org/">https://ggplot2.tidyverse.org/</a>
raster v.3.4–10	Hijmans <sup>75</sup>	<a href="https://cran.r-project.org/web/packages/raster/index.html">https://cran.r-project.org/web/packages/raster/index.html</a>
BAMMtools v.2.1.6	Rabosky et al. <sup>76</sup>	<a href="http://bamm-project.org/">http://bamm-project.org/</a>
RevBayes	Höhna et al. <sup>77</sup>	<a href="https://revbayes.github.io/">https://revbayes.github.io/</a>
castor v.1.5.7	Louca and Doebeli <sup>78</sup>	<a href="https://cran.r-project.org/web/packages/castor/index.html">https://cran.r-project.org/web/packages/castor/index.html</a>
nlme v.3.1–149	Pinheiro and Bates <sup>79</sup>	<a href="https://cran.r-project.org/web/packages/nlme/index.html">https://cran.r-project.org/web/packages/nlme/index.html</a>
diversitree v.0.9–15	FitzJohn <sup>80</sup>	<a href="https://cran.r-project.org/web/packages/diversitree/index.html">https://cran.r-project.org/web/packages/diversitree/index.html</a>
Trim Galore! V.0.40	Krueger <sup>81</sup>	<a href="https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/">https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/</a>

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests should be directed to the lead contact, Ana P. S. Carvalho ([apsdecarvalho@gmail.com](mailto:apsdecarvalho@gmail.com)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

- Nucleotide, sequencing, and associated datasets have been deposited on GenBank and Dryad and are publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- All original code has been deposited on Dryad and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## METHOD DETAILS

### Molecular data

#### Locus sampling

We obtained sequence data in 3 different ways. These were: 1) the anchored hybrid enrichment (AHE) BUTTERFLY2.0 kit,<sup>82</sup> which targets 13 genes, 2) AHE BUTTERFLY1.0<sup>49</sup>, which targets 425 genes, and 3) publicly available sequence data. Details of these approaches are outlined below.

#### Anchored hybrid enrichment

We sequenced DNA from 338 Pieridae species and two outgroups sourced from various alcoholic and dry insect collections (see full sample list on Dryad at <https://doi.org/10.5061/dryad.c59zw3rbh>) using methods described in Kawahara et al.<sup>25</sup> The BUTTERFLY2.0 AHE probe set captured up to 13 loci commonly used in Sanger sequencing studies of past decades.<sup>82,83</sup>

Raw sequencing reads underwent filtering and quality checks for base call quality (phred >20) and length (>30 bases) using Trim Galore! v.0.40<sup>81</sup>; and assembled via a custom Python v.2.7.6 script that performs an iterative baited assembly (IBA)<sup>84</sup> for each locus. IBA uses similarity searching in usearch<sup>85</sup> to assemble loci beginning with a database of reference loci from a closely related species (in this case *Danaus plexippus*) using bridger.<sup>86</sup> Second, assembled loci were filtered by blasting the original probe regions of the assembly against the *Danaus plexippus* reference genome<sup>87</sup> followed by mapping single hits to check for orthology. We screened the orthologous loci for sequence contamination and removed any sequences with 99% similarity to species in another family. The resulting sequences were aligned with MAFFT v.7.0.1.<sup>64</sup> Different copies of DNA were collapsed into consensus sequences using FASconCAT-G v.1.0.4.<sup>66</sup>

We augmented the dataset with 103 pierid species sequenced by Kawahara et al.<sup>25</sup> with the BUTTERFLY1.0 probe set<sup>49</sup> which have up to 425 AHE loci (including the same 13 loci in BUTTERFLY2.0).

#### Published sequences

We complemented the AHE molecular dataset with sequences of up to 13 loci (the same captured in BUTTERFLY2.0) from 134 pierid species in GenBank downloaded using the software GeneDumper v.0.8 (<https://github.com/sunray1/genedumper>).

We used the 13 loci captured by BUTTERFLY2.0 from *Colias erate* as a reference sequence for BLAST to find other available Pieridae sequences in GenBank. GeneDumper can integrate taxonomic name updates in the GenBank search and thus requires a synonymic list to determine name validity. We used the species list of Lamas.<sup>34</sup> After the original BLAST, GeneDumper filters and checks the sequence hits. When multiple sequences for the same locus of a species are available, GeneDumper considers the characteristics of the data and selects the longest sequence with the most unambiguous DNA content. The identification of the selected sequences are verified via a self-blast against GenBank's nucleotide collection. A tiling approach is also used, where sequences are tiled along the bait sequence used in the original BLAST and the least number of tiles providing the highest coverage across the baits are chosen. Therefore, one species may be represented by sequences from multiple specimens if there are no sequences available that fully cover the bait sequence (Further details of the method can be found at <https://github.com/sunray1/genedumper>).

The sequences of Warren-Gash et al.<sup>47</sup> were made available on GenBank after we used GeneDumper, so we downloaded those sequences from the database for species not already in our molecular dataset. We also incorporated sequences from other targeted studies,<sup>46,49</sup> and used genomes and transcriptomes assembled to the BUTTERFLY1.0 probe kit in Kawahara et al.<sup>25</sup> (including two as outgroups).

There were 15 species for which we created chimeras by adding cytochrome oxidase subunit 1 (COI) from one sample to another with more loci available but no COI (see full sample list on Dryad at <https://doi.org/10.5061/dryad.c59zw3rbh>).

#### Combining molecular datasets

All genetic sequences were then combined to create the molecular data matrix. This method of creating datasets where molecular coverage varies greatly has been shown to yield high-quality phylogenetic reconstructions if strategic taxon selection of species is taken into account for multilocus sequencing.<sup>88</sup> We undertook this process by first combining data from the two AHE datasets. The 13 loci of BUTTERFLY2.0 are included (as loci numbers 1 through 13) in BUTTERFLY1.0<sup>49</sup> and simplify concatenation and combination of datasets. We then manually added GenBank sequences to this genetic dataset, and subsequently aligned each locus with MAFFT following the gene alignment protocol of Breinholt et al.<sup>84</sup> After alignment, each locus was examined in AliView v.1.27<sup>65</sup> to check for indels and determine reading frames. Loci were concatenated with FASconCAT-G v.1.0.4<sup>66</sup> to create a concatenated dataset of 425 loci.

The final dataset had fifteen outgroups with at least two species from each of the six other butterfly families. The taxonomy of butterfly names follows Lamas.<sup>34</sup> In total, 593 pierid species and 15 outgroups are represented in our final concatenated dataset with up to 425 loci and 180,381 base pairs per species. Missing data is variable due to the inclusion of taxa limited to 13 loci and to 425 loci, see Dryad for a complete breakdown of loci recovered on a per-species basis.

#### Partitioning, tree inference, and branch support

The best nucleotide partition schemes were identified using ModelFinder as implemented in IQ-TREE v.2.0.3.<sup>67,68</sup> For this analysis, we allowed partitions to be identified and merged with '-m TESTNEWMERGEONLY' to limit overparameterization. To reduce computation times, we used the rcluster method (set to 10, with a maximum of 1000) and limited the model set to GTR. We consolidated the original 425 locus

partitions into 69 partitions but did not further partition these by codon positions to avoid extremely high numbers of *a priori* partitions for such a large taxon set. Separately, we ran ModelFinder to identify the best models of nucleotide evolution for our partitions using all available models. With these 69 partitions and models, we then performed 100 independent IQ-TREE tree searches with 1000 Ultrafast bootstraps (UF-Boot) and 1000 Shimodaira-Hasegawa approximate likelihood ratio test (SH-aLRT) replicates to test for branch support. We implemented the ‘-bnni’ command when running these analyses to alleviate concerns about model violation inherent in the Ultrafast bootstrap method.<sup>89</sup> All phylogenetic analyses were run on the University of Florida HiPerGator2 Cluster (<https://www.rc.ufl.edu/about/hipergator/>).

### Divergence time estimation

We used two different secondary calibration schemes obtained from the trees of Espeland et al.<sup>49</sup> and Kawahara et al.<sup>53</sup> to estimate pierid divergence times (Table S4). Ages, based on 95% confidence intervals in these studies, were treated as hard minimum/maximum ages and we employed the penalized likelihood tree dating with TreePL.<sup>69</sup> TreePL requires an input tree, for which we used the most likely tree (highest likelihood of 100 runs) inferred with IQ-TREE. TreePL outputs divergence times as single dates per node. In order to provide a range of dates for each divergence event, we performed TreePL independently 100 times and summarized results, following the approach of St Laurent et al.<sup>90</sup> This method uses several custom Python scripts (available at <https://github.com/sunray1/treepL>) and creates 100 bootstrap replicates in RAxML<sup>70</sup> v. 8.2.12 using the 425-locus data matrix, partitioned according to the same 69 partitions inferred with ModelFinder. These 100 bootstrap alignments were then used to infer branch lengths on the constrained input phylogeny. We generated 100 bootstrap trees, utilizing optimization parameters that were determined with the ‘prime’ command and random subsample and replicate cross-validation with the ‘randomcv’ command for identifying the best smoothing parameter. This optimization and smoothing step was performed three times on each tree to identify the best combination of parameters and assess consistency in the smoothing parameter. All analyses used the ‘thorough’ command in TreePL. These analyses resulted in 100 dated trees varying only in branch lengths, which were summarized as a maximum clade credibility (MCC) tree in TreeAnnotator v.1.10.4 in the BEAST package.<sup>71</sup> The MCC tree was used in all downstream analyses that required a dated phylogeny.

In addition to making use of secondary calibrations in TreePL, we conducted an analysis using fossil calibrations as the minimum ages. To calibrate the root, a median age of Angiosperms (139.4 Ma) from Magallón et al.<sup>91</sup> was used (as done by Espeland et al.<sup>49</sup>). In total, eight fossil calibrations were applied to the following places in the tree (fossil names and their age noted in parentheses): 1) the Parnassiinae stem age (*Thaites ruminiana*, 23 Ma); 2) the Hesperidae stem (*Protocoeliades kristenseni*, 54 Ma); 3) an internal Hesperidae crown age (*Pamphilites abdita*, 23 Ma); 4) Nymphidiini (Riodinidae) stem (*Theope* sp., 15 Ma); 5) the crown node of Satyrinae + Heliconiinae (Nymphalidae) (*Vanessa amerindica*, 33.7 Ma); 6) *Aporia* (Pieridae) crown age (*Aporia* cf. *crataegi*, 2.6 Ma); 7) *Belenois* (Pieridae) crown age (*Belenois crawshayi*, 0.02 Ma); 8) split between Coliadinae and (Pseudopontiinae + Pierinae) (*Vanessa pluto*, 16 Ma). Each of these fossils were selected because they can be confidently identified as belonging to a particular lineage, making them well-suited to calibrate dating analyses.<sup>92</sup> TreePL was run on this dataset as noted above, with the exception that the root was a hard maximum age and all fossils were hard minimum ages. For our discussion, we focus on estimated dates obtained using calibrations from Kawahara et al.,<sup>53</sup> because this is the most comprehensive, fossil-calibrated dated phylogeny of Lepidoptera to date.

### Distribution and bioclimatic data

We used the R library occCite v.0.4.9<sup>72</sup> to assemble distribution data. It provides a platform to compile georeferenced data from global databases while also generating a list of primary data sources for each record from institutional collections. We extracted occurrence data using the command ‘occQuery’ and followed the taxonomy of Pieridae according to Lamas.<sup>34</sup> This command checks species names against the Global Names Index ([gmi.globalnames.org](http://gmi.globalnames.org)), then uses rgbif<sup>73</sup> to send a query to GBIF,<sup>93</sup> requesting all records with geographic coordinates. We removed records that did not have decimal places to increase accuracy and removed records with identical coordinates to the first four decimal places because they were too precise to provide meaningful additional information. Original occurrence data sources with digital object identifiers (DOIs) were gathered using ‘occCitation’. We also extracted additional occurrence data from the literature and directly from GBIF for 37 species that were not obtained using the method described above (see full reference list on Dryad at <https://doi.org/10.5061/dryad.c59zw3rbh>).

We mapped all records by species in an atlas using the package ggplot2 v.3.3.5,<sup>74</sup> with a custom R script (<https://github.com/hannahlowens/PieridaeDiversity>). The atlas was inspected for occurrence points that appeared as outliers or errors based on published distribution maps and opinions of Pieridae experts (authors and collaborators - see Acknowledgments) and we subsequently removed these putatively erroneous records. We also combined the records of the cryptic species *Leptidea sinapis*, *L. juvenica* and *L. reali* as the “sinapis complex” due to difficulty in delimiting distribution in these species.<sup>94</sup> The final dataset had over 800,000 records for 541 pierid species, representing 91% of species in the phylogeny (<https://doi.org/10.5061/dryad.c59zw3rbh>). We wanted to have present day representation of distribution for a more informative dataset on abiotic requirements, therefore we did not remove records that could be associated with established introductions and migration.

We extracted bioclimatic data for the geographic locations of these records from the WorldClim dataset (2.5 arc-minute resolution<sup>95</sup>) using the package raster v.3.4–10.<sup>75</sup> Since climatic stability can refer to different timescales including paleoclimates,<sup>96</sup> we use the phrase “temperature stability” to refer to annual variability (seasonality) and daily variability, depending on the context. Considering our hypotheses regarding the roles of extreme conditions and seasonality, we focused on the following variables: annual mean temperature (average of BIO1), mean diurnal temperature range (average of BIO2, which is the mean of (daily max temp - daily min temp) within a month), maximum

temperature in the warmest month (95% quantile of BIO5), minimum temperature in the coldest month (5% quantile BIO6), and temperature annual range (average of BIO7). The averages (BIO1, BIO2, BIO7), or 5% (BIO6) or 95% (BIO5) quantiles were calculated for each species that had at least one record. The summarized data can be found on Dryad (<https://doi.org/10.5061/dryad.c59zw3rbh>).

A maximum likelihood-based ancestral state reconstruction (ASR) was conducted in Phytools using the 'contMap' function ('anc.ML' model) of all temperature-related variables using the dated Pieridae tree without outgroups.

### Diversification and correlation analyses

We conducted a BAMM (Bayesian Analysis of Macroevolutionary Mixtures) analysis to infer diversification rates on the pierid phylogeny. The input topology was the TreePL tree, dated with secondary calibrations from Kawahara et al.,<sup>53</sup> excluding outgroups. We estimated prior parameters using the R package BAMMtools v.2.1.6<sup>76</sup> with the command 'setBAMMpriors' (lambdaInitPrior ~3.33; lambdaShiftPrior ~0.01; mulnitPrior ~3.33). We ran a reversible-jump MCMC for 100 million generations, sampling every 100,000 generations with six different expected shifts (5, 10, 15, 20, 25, and 30). We used a fraction file to account for taxon sampling bias across the tree (<https://doi.org/10.5061/dryad.c59zw3rbh>). BAMM output files were analyzed in BAMMtools with a 10% burn-in. We plotted shift posterior probabilities for all six analyses and determined that they converged around 20 shifts (Figure S12). Thus, the result based on 20 shifts was used for diversity analyses described below.

We also conducted a lineage-specific birth-death-shift (LSBDS) analysis in RevBayes<sup>77</sup> with an MCMC chain of 10,000 steps; following the parameters employed in Höhna et al.<sup>97</sup> Rate variables were set to lognormal distributions for both mean speciation and extinction rates, the rate of rate-shift events was set to uniform, and six rate categories were set as a global parameter. LSBDS does not have an option for fine level sampling bias, and therefore a global sampling fraction of 51% of 1,159 described species of Pieridae was applied.

Considering recent criticisms of estimating diversification rates using phylogenies of extant taxa (e.g., Louca and Pennell<sup>98</sup>), we also used the castor v.1.5.7 R package,<sup>78</sup> to estimate pulled speciation rates (PSR) and deterministic lineage through time plots (dLTT). To ensure time slices are biased more toward the present where extant time trees are expected to be more reliable, we first used the function 'fit\_hbd\_psr\_on\_grid' with arguments 'Ngrid' set to 15, and 'Ntrials' to 100 in order to get density values for X and Y axes. These density values were used as arguments in the castor function 'castor::get\_inhomogeneous\_grid\_1D' to create an inhomogeneous 'age\_grid' which could then be used as an argument in the function 'fit\_hbd\_psr\_on\_grid' to calculate pulled speciation rate (PSR) and dLTTs on the inhomogeneous grid.

To assess a possible correlation between temperature and rates of diversification, we calculated phylogenetic generalized least squares (PGLS) in the R package nlme v.3.1–149.<sup>79</sup> Speciation and extinction rates were extracted from the results from the BAMM analysis using the 'getTipRates' command in BAMMtools for all species in the phylogeny. With the command 'gls', we set the correlation function to Brownian and method to 'ML'. We performed ANOVA tests on resultant PGLS models to check for significance and identify the relationship of diversification rates and variable (slopes) as positive or negative.

As another means of testing for the relationship between temperature and speciation rates, we used Quantitative State Speciation and Extinction (QuaSSE) as implemented in diversitree v.0.9–15.<sup>80</sup> We used a pruned tree including only species with available trait data (541 species). Our methods largely followed Zhang et al.,<sup>99</sup> whereby we tested increasingly complex QuaSSE models to create likelihood functions where speciation is constant, linear, sigmoidal, or hump-shaped, with and without a drift parameter. The Akaike information criterion (AIC) was used to identify the best fit QuaSSE model for each of the BIO variables.

Climatic BIO variables used in the analysis above stem from present day temperatures. To evaluate how global temperature changes through geological time affected diversity in the family as well as its most diverse subfamily Pierinae, we fitted four temperature-dependent models to our phylogeny in the R. We also considered the possibility of time influencing diversity by fitting six time-dependent models. To run these analyses, we used the pipeline and custom R scripts developed by Condamine et al.<sup>100</sup> Specifically, time-dependent models were derived from Morlon et al.<sup>101</sup> and Stadler<sup>102</sup> and the paleoclimate models from Condamine et al.<sup>103</sup> The original script provided by the authors included linear and exponential models, but due to criticism directed toward linear diversification dependency models by Gamisch<sup>104</sup> (but also see Morlon et al.<sup>105</sup>), we commented out code for linear models. The time-dependent models were fitted with maximum likelihood using the 'fit\_bd' function and the paleoclimate models used the 'fit\_env' function, both functions available in the pipeline of Condamine et al.<sup>100</sup> Sampling fractions of 0.5 for Pieridae and 0.44 for Pierinae were assigned to account for the proportion of taxon coverage in the phylogenies.