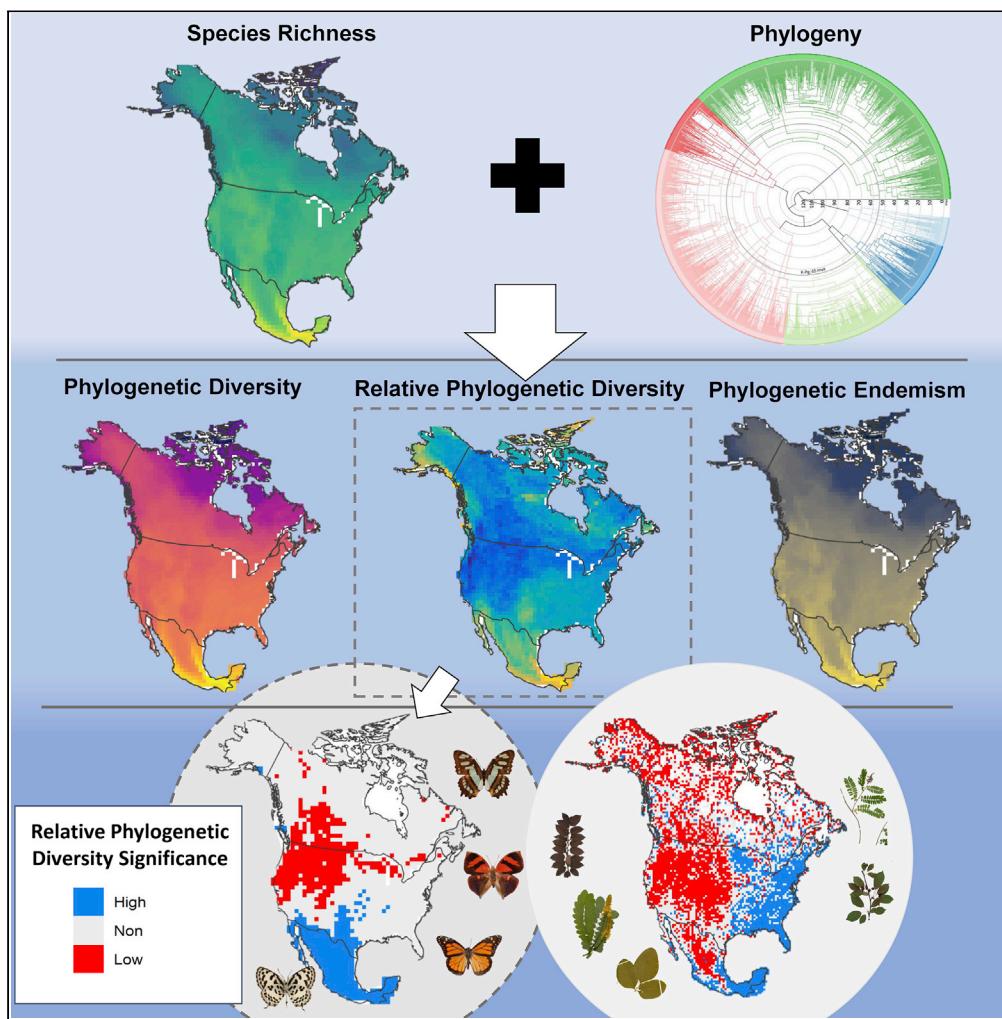


Article

Spatial phylogenetics of butterflies in relation to environmental drivers and angiosperm diversity across North America



Chandra Earl,
Michael W. Belitz,
Shawn W.
Laffan, ..., Brent D.
Mishler, Akito Y.
Kawahara, Robert
Guralnick

sunray1@ufl.edu

Highlights

Butterfly phylogenetic diversity (PD) is highest in the warmest, wettest areas

Areas with historically stable temperatures have high relative PD (RPD)

The relationship of plant and butterfly RPD is weak and spatially structured

North American warm deserts have high endemism, making it a conservation priority

Earl et al., iScience 24, 102239
April 23, 2021 © 2021 The Authors.

<https://doi.org/10.1016/j.isci.2021.102239>



Article

Spatial phylogenetics of butterflies in relation to environmental drivers and angiosperm diversity across North America

Chandra Earl,^{1,2,9,10,*} Michael W. Belitz,^{1,3,4,9} Shawn W. Laffan,⁵ Vijay Barve,¹ Narayani Barve,¹ Douglas E. Soltis,^{1,2,3,4} Julie M. Allen,⁶ Pamela S. Soltis,^{1,2,4} Brent D. Mishler,^{7,8} Akito Y. Kawahara,^{1,2,3} and Robert Guralnick^{1,2,3,4}

SUMMARY

Broad-scale, quantitative assessments of insect biodiversity and the factors shaping it remain particularly poorly explored. Here we undertook a spatial phylogenetic analysis of North American butterflies to test whether climate stability and temperature gradients have shaped their diversity and endemism. We also performed the first quantitative comparisons of spatial phylogenetic patterns between butterflies and flowering plants. We expected concordance between the two groups based on shared historical environmental drivers and presumed strong butterfly-host plant specializations. We instead found that biodiversity patterns in butterflies are strikingly different from flowering plants, especially warm deserts. In particular, butterflies show different patterns of phylogenetic clustering compared with flowering plants, suggesting differences in habitat conservation between the two groups. These results suggest that shared biogeographic histories and trophic associations do not necessarily assure similar diversity outcomes. The work has applied value in conservation planning, documenting warm deserts as a North American butterfly biodiversity hotspot.

INTRODUCTION

Insect biodiversity patterns are poorly understood across broad spatial, temporal, and phylogenetic scales. These shortfalls stand in stark contrast to our knowledge of vertebrate and flowering plant biodiversity, especially in North America, where rapid efforts to close phylogenetic and spatial information gaps (Davies and Buckley, 2011; Thornhill et al., 2017; Allen et al., 2019) have provided novel insights into the shorter and longer term processes structuring biodiversity and how best to preserve natural heritage for the future (Kling et al., 2019). One approach to quickly expand the knowledge base of insect biodiversity and preserve it in the face of accelerating terrestrial declines (van Klink et al., 2020) lies in focusing on clades where existing data are already dense but not yet fully integrated. Such efforts also provide a unique basis for direct, empirical comparisons with other lineages, such as various lineages of green plants, that are known to have strong evolutionary and ecological associations with herbivorous insects (Futuyma and Agrawal, 2009).

Butterflies (Papilioidea) serve as an ideal study group for researchers and naturalists due to their diurnal activity, often vibrant and showy colors, and specialized larval host plant associations (Brock and Kaufman, 2003; Grimaldi and Engel, 2005). Not only are they the most collected and photographed insects (Scoble, 1995), but many species and clades of butterflies have become models for studying diverse ecological and evolutionary processes, such as Batesian and Müllerian mimicry (e.g., butterflies in the genus *Heliconius* (Brower, 1996; Kronforst and Papa, 2015; Lewis et al., 2019)), genetics and migration (e.g., butterflies in the genus *Danaus*), and adaptation to agricultural systems (e.g., the common and widespread cabbage white, *Pieris rapae* (Shen et al., 2016)). Butterflies also serve as pollinators and bioindicators of change and are one of the few insect groups where conservation agencies such as the IUCN have made at least initial assessments of species endangered status (Bonelli et al., 2018).

Due to the interest of both professionals and amateurs, the natural history of North American butterflies is relatively well known with rich distributional and genetic data resources readily available. There are

¹Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA

²Genetics Institute, University of Florida, Gainesville, FL 32611, USA

³Department of Biology, University of Florida, Gainesville, FL 32611, USA

⁴Biodiversity Institute, University of Florida, Gainesville, FL 32611, USA

⁵School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, Australia

⁶Department of Biology, University of Nevada, Reno, Reno, NV 89557, USA

⁷University of Jepson Herbaria, University of California, Berkeley, CA 94720, USA

⁸Department of Integrative Biology, University of California, Berkeley, CA 94720, USA

⁹These authors contributed equally

¹⁰Lead contact

*Correspondence:
sunray1@ufl.edu

<https://doi.org/10.1016/j.isci.2021.102239>



approximately 1900 species of butterflies in North America (Lotts and Naberhaus, 2017), and natural history and genetic data exist for nearly 1500 species of them. This abundance of data positions butterflies as one of the best insect groups for asking broad-scale questions about the structure and drivers of diversity. These strong data sources enable moving beyond simple taxic summaries of diversity, such as species richness, and toward a more comprehensive, process-oriented understanding of how diversity is evolutionarily structured at the continental scale. Despite such potential, a synthetic, broad-scale phylodiversity analysis of butterflies (or any other insect group) and the drivers of that diversity has yet to be conducted. Even North-America-wide summaries of butterfly taxic diversity have been limited (Ricketts et al., 1999; Kocher and Williams, 2000; Luis-Martinez et al., 2002).

Butterflies are sensitive to changes in climate (Dennis, 1993), and a fundamental question is how current climate and historical changes in temperature and landscape across North America have shaped butterfly phylogenetic diversity and endemism. North America is characterized by a wide range of ecosystems, a dynamic geological history, and significant insect diversity (Danks, 1994; Godfray et al., 2000). Butterflies are distributed across 14 broad ecoregions, ranging from the eastern temperate forests to tundra and taiga in Northern Canada, tropical wet forests in southern Mexico, and the warm and cold deserts of the Southwest (Lotts and Naberhaus, 2017). Landscapes across the continent have dramatically changed during the Quaternary, especially in the west, due to long-term aridification and orogeny leading to formation of the Sierra Nevada mountains, and across northern portions of the continent through cyclic patterns of glaciations (Bintanja and van de Wal, 2008).

Butterflies also rely heavily on flowering plants, as sources for both adult nectar and larval food (Bronstein et al., 2006). A key question is whether butterflies and angiosperms show concordant broad biogeographic patterns, given these strong ecological associations and shared historical landscape and climate drivers. Recent efforts to document North American plant phylodiversity (Mishler et al., 2020) provide a data basis for direct, quantitative comparisons of butterflies with angiosperms. That recent analysis is the most comprehensive yet attempted, covering more than 19,500 plant species (out of more than 44,000 total species), found across the continent. This study is the first to directly compare spatial patterns and drivers of phylogenetic diversity between any group of insects and flowering plants at a continental scale.

Here we assembled and analyzed butterfly spatial phylogenetic diversity across North America and examined its connection to historical climate and flowering plant phylodiversity patterns. Phylogenetic approaches have two key advantages compared with traditional taxic approaches. First, phylodiversity metrics reduce reliance on species definitions; rather, branch lengths are used to calculate diversity metrics. Second, spatial phylogenetic approaches bring in evolutionary history and allow hypothesis testing, making it possible to assess, for example, whether communities are more distantly or closely related to each other than expected by chance.

We applied a set of spatial phylogenetic methods and metrics, including phylogenetic diversity (PD) (Faith, 1992), phylogenetic endemism (PE) (Rosauer et al., 2009), and relative phylogenetic diversity and endemism (RPD and RPE) (Mishler et al., 2014). We also employed CANAPE, which can differentiate between types of endemism found in a region (Mishler et al., 2014), namely between recent radiations leading to neoendemism and relictual endemism leading to range-restricted groups that were once more widespread, i.e. paleoendemism. Although these metrics are now commonly applied, we provide a short summary of those used here in Table 1.

We used these metrics of phylogenetic diversity and endemism to test hypotheses about a set of potential drivers and associations, including a unique, direct empirical comparison between butterflies and flowering plants. These same techniques also document centers of diversity and endemism that may differ from plant or vertebrate groups and inform conservation prioritization. Based on a recent analysis of North American plant phylodiversity, we made the following predictions:

1. Regions that are warmer and have remained more stable over time will have higher phylogenetic diversity (PD) (Rohde, 1992; Mittelbach et al., 2007). Stable areas, whether warm or cold, should have significantly higher than random PD because they have had the most time to accumulate lineages (Cowling and Lombard, 2002), along with specializations that may structure communities to avoid competition (Fine, 2015).

Table 1. Summary of phylodiversity metrics and tests used

Phylogenetic diversity (PD)	Measured as the sum of branch lengths connecting the terminal taxa present in each location (usually to the root of the tree)
Phylogenetic endemism (PE)	Like PD but measured on a tree where the branches are weighted by the inverse of their geographic range (a Range-Weighted Tree)
Relative phylogenetic diversity (RPD) and relative phylogenetic endemism (RPE)	Ratio of PD or PE measured on the original tree to PD or PE measured using a comparison tree with the same topology but where each branch is adjusted to be of equal length
Categorical analysis of neo- and paleo-endemism (CANAPE)	Geographic centers of endemism are identified, first as being significantly high in either the numerator or the denominator of RPE (or both) and then classified as paleo, neo, or mixed based on whether the RPE ratio is significantly high or low
Randomization tests	These metrics are tested for statistical significance using a spatially structured randomization that re-assigns terminal taxon occurrences on the map, subject to two constraints: the range size of each taxon and the richness of each locations are held constant

2. Relative phylogenetic diversity (RPD) will be higher than expected in areas that have been most stable, accumulating more long-surviving, older lineages (Fine, 2015). In North America this includes the eastern and southernmost portions of the continent, as seen in flowering plants. Areas with high topographic heterogeneity and that have been most climatically unstable, such as recently deglaciated areas in the north and portions of the west, will have significantly lower than expected RPD.
3. Butterfly phylogenetic endemism in North America will align with hotspots of high angiosperm endemism. Hotspots of neoendemism are more likely in younger areas with higher topographic relief, whereas areas of paleoendemism will be more likely where climate and landscapes have been more stable.
4. Continental-scale flowering plant and butterfly phylodiversity patterns will be congruent, due to co-evolutionary dynamics between the two groups and similarities in underlying landscape and climate drivers. Alternatively, the relative breadth of butterfly host preferences and the limited number of host plants compared with overall plant diversity may dilute and obscure congruence.

Materials and methods

Species name assembly

We consolidated a list of all North American butterfly species with their current valid names and known synonyms (see “SupDryad_specieslist.csv” in Earl (2021)). We defined North America as including Canada, Mexico, and the United States but excluded species that were endemic to islands near the North American landmass (e.g., the Caribbean), as these islands were not well included in field guides used for documenting ranges. Valid names were derived from a global checklist (Lamas, 2004) and were augmented via assembly of synonymies from the *Lepidoptera and other life forms* database (Funet (Savela, 2020)) and Wikipedia (Wikipedia, 2020) using the R package taxotools (Barve, 2020). The augmented master list was used to normalize names from resources that contained expert-assessed maps and assembled names used in field guide resources (Brock and Kaufman, 2003; Glassberg, 2018). Once names were normalized to a consistent, accepted name, we used those names and associated synonyms to (1) re-assign normalized names to those digitized species range maps (see below for range map assembly) where normalization was required and (2) search GenBank and other key resources for matching genetic data to construct a North America-specific butterfly phylogeny.

Range maps and digitization

Range maps were digitized from field guides covering the USA and Canada ([Brock and Kaufman, 2003](#)) and Mexico ([Glassberg, 2018](#)) for each species included in our species list. Digitizing steps included generating high-resolution scans, georeferencing the resulting images, and then manually tracing polygons based on those scans in QGIS version 3.2 ([QGIS Development Team, 2020](#)). Fewer than 1% of the species in field guides did not have an associated range map; in those cases we used occurrence records and descriptions in field guides to estimate the ranges. Range maps were combined into a single shapefile consisting of many spatial polygons that were clipped to only terrestrial areas within North America (see "SupDryad_fishnet.csv" in [Earl \(2021\)](#)). We provide details about range map digitization and a rigorous approach to quality control of maps in [supplemental information](#).

Many butterfly species have ranges that extend beyond the borders of North America. Calculations that involve range-weighting, such as phylogenetic endemism, should ideally rely on globally complete phylogenies and range estimates based on highly resolved maps. Here, we partially compensated for this currently unattainable goal by determining a coarse estimate of overall range extents using country-level range maps for every species in our list where needed. The ranges used in the endemism analyses (see below) were the sums of the areal extent of both the range maps from within North America and the total country level area extent outside North America. We generated these country-level estimates utilizing three separate resources: (1) country-level ranges from Funet ([Savela, 2020](#)); (2) GBIF data for all relevant species and extracting country-level data from these records; and (3) data from a trait database that was assembled from field guides and other published sources ([ButterflyNet, 2020](#)). [Supplemental information](#) describes more details on country list production and the extensive quality control that went into verifying a final country list.

A 100 km by 100 km resolution grid at the global scale (in order to include range areas from outside the continent) was projected to a North America Albers equal area conic coordinate system. A species was considered present in the cell if the cell centroid was within the species' range map or if the distance from a cell's centroid to the nearest edge of the species' range map was less than 1 km.

Sequence data acquisition and dataset construction

We compiled sequence data for 13 common markers (1 mitochondrial and 12 nuclear genes) used in butterfly phylogenetics (see [Table S1](#)). These sequences were obtained from GenBank ([Clark et al., 2016](#)), Barcode of Life Data System (BOLD) ([Ratnasingham and Hebert, 2007](#)), and by extracting DNA from tissues and sequencing with Sanger and target capture sequencing for species that lacked genetic data (See "SupDryad_lociloc.csv" in [Earl \(2021\)](#)). We accumulated existing marker data from GenBank using a python toolkit (<https://github.com/sunray1/GeneDumper>) developed by the lead author. This toolkit automatically fetches sequences from GenBank and uses a rigorous, automated cleaning workflow to choose the best matching sequences for further processing. Specifically, this workflow uses NCBI's BLAST toolkit to query GenBank for loci of interest across a list of species. Accounting for taxonomic updates and errors in species names, sequences are subject to various thresholds and metrics to select the ideal sequence for downstream analyses. The thresholds and metrics include length and content filters, a self-BLAST of the chosen sequences to check for correct species labeling, and a clustering analysis to account for differences in SNPs. We also queried for relevant loci and sequences in BOLD ([Ratnasingham and Hebert, 2007](#)), as some of these data records are not reflected in GenBank. BOLD data were assembled using its API (<http://www.boldsystems.org/index.php/resources/api>). Target capture sequencing utilized Anchored Hybrid Enrichment (AHE) ([Lemmon et al., 2012](#)) on 224 butterflies with the Butterfly 1.0 ([Espeland et al., 2018](#)) and 2.0 ([Kawahara et al., 2018](#)) target capture sets. Both of these kits include the 13 markers of interest.

We also generated new mitochondrial cytochrome c oxidase subunit I gene (*COI*) sequences by extracting DNA from dried pinned museum specimens in the Florida Museum of Natural History, McGuire Center for Lepidoptera, and Biodiversity at the University of Florida. These tissues were shipped to the Canadian Center for DNA Barcoding (CCDB; Guelph, Canada) for sequencing of the standard 658-bp region of *COI*. See [supplemental information](#) for additional information on sequence acquisition. We created FASTA files for each locus, which were aligned using MAFFT v.7.294b ([Katoh and Standley, 2014](#)) and concatenated into a single alignment with FASconCAT-G v.1.02 ([Kück and Longo, 2014](#)). The concatenated alignment was 12,361 bp in length and had 69% missing data. Most species (99%) were represented by *COI*

(See “SupDryad_alignment.fa” in [Earl \(2021\)](#)). Without COI, missing data increased to 76% across the remaining 12 loci.

Phylogeny construction

All loci included in this study were protein coding, and therefore the alignment was partitioned by gene and codon position, resulting in 39 partitions. PartitionFinder v.2.1 ([Lanfear et al., 2016](#)) was used to choose the best partitioning scheme and nucleotide substitution models. A phylogenetic analysis of North American butterflies was conducted in RAxML v.8.2.10 ([Stamatakis, 2014](#)) using a family-level constraint based on [Espeland et al. \(2018\)](#) and with the concatenated alignment and best partitioning scheme. We conducted 100 ML tree-searches with different random seeds, and the tree with the best log likelihood score was chosen as the final tree. We determined branch support by running 200 parametric bootstrap replicates in RAxML under the GTR+Γ+I model and a gradual “transfer” distance method implemented in BOOSTER ([Lemoine et al., 2018](#)). A final check was made to ensure that tip names were consistent with the species names from range products.

The interpretation of phylodiversity metrics depends on the units of the branch lengths in the phylogeny, e.g., the amount of “feature diversity” contained in a region when using a phylogram or the amount of “evolutionary history” in a region when using a chronogram ([Allen et al., 2019](#)). We focus here on chronograms, with their explicit focus on the age of lineages and communities, and therefore produced a time-calibrated tree. Divergence times were calculated using penalized likelihood in TreePL ([Sanderson, 2002; Smith and O’Meara, 2012](#)) following a congruification approach ([Eastman et al., 2013](#)). In particular, we obtained node calibrations from [Espeland et al. \(2018\)](#), extracting date ranges from each of the six family nodes that were concordant with our phylogeny and using those as estimates in TreePL (see [supplemental information](#) for more phylogeny reconstruction details).

Analysis of phylogenetic diversity and endemism

The spatial dataset and the phylogeny described above were imported into Biodiverse v.3.0 ([Laffan et al., 2010](#)). Tips on the tree were mapped to species in the spatial dataset to calculate species richness (SR), phylogenetic diversity (PD), and phylogenetic endemism (PE) metrics for equal-area square grid cells (100 × 100 km). The imported mapping extent was global (as described above) in order to calculate range size metrics, but our analysis region for phylodiversity metrics was constrained to North American grid cells using spatial constraints in Biodiverse. This approach ensured that PE metrics take into account overall range sizes of the terminals in the tree, including their extent outside of the continent. Relatives of the terminals that occur elsewhere in the world are not included, but this is the best approach possible to estimating PE until global analyses are feasible.

We also calculated relative phylogenetic diversity (RPD) and relative phylogenetic endemism (RPE) ([Mishler et al., 2014](#)). These are ratios of PD and PE on the original tree compared with a phylogeny with the same topology but with equal branch lengths. These metrics can provide useful information about areas with concentrations of significantly longer or shorter than expected phylogenetic branches. For example, areas that have more recently radiated taxa are expected to have lower chronogram-derived RPD. All cell-based values for PD, PE, RPD, and RPE were exported from Biodiverse for mapping and further analysis.

Randomization tests

Phylogenetic diversity and endemism measurements are expected to be highly correlated with taxic diversity (species richness), because each taxon added to a community must also add to the overall PD. If the co-occurring taxa are randomly distributed on the tree, the correlation should be tight, so a key step forward is to move beyond simply reporting summary measures, and test whether phylodiversity values are higher or lower than expected compared with null models of randomized communities ([Webb et al., 2002](#)). This is achieved through a randomization approach, where species occurrences within North America are randomly reassigned to grid cells while holding constant the richness of each cell and the range size of each species. Values for PD, PE, RPD, and RPE were then calculated for each randomization iteration, creating a null distribution for each grid cell. A two-tailed test was then applied to the PD, PE, and RPD randomizations to determine whether the observed values were significantly high or low when compared with the null distributions. We utilized a 40-core Dell Xeon PowerEdge standalone server to parallelize creation of 500 random realizations per grid cell across all cores. Only observations within the North American

study region were randomized, with the outside regions held constant. Output randomized results were merged, exported as GeoTIFF grids, and re-imported in R for downstream analysis.

RPE randomizations enable a means to categorize different types of phylogenetic endemism. This method, called *Categorical Analysis of Neo- And Paleo- Endemism* (CANAPE ([Mishler et al., 2014](#))), is a two-step approach that first selects grid cells that are significantly high (one-tailed test) in either the numerator or the denominator of RPE, then uses a two-tailed test of the RPE ratio to determine four possible outcomes per cell: higher than expected concentrations of range-restricted short branches (i.e. neoendemics); long branches (i.e. paleoendemics); a mixture of both types; or no significant endemism. Endemism measures, including randomizations, were calculated in Biodiverse, and the categorization method for CANAPE was run in R v.3.6.3 ([R Core Team, 2019](#)) to determine per-grid-cell phylogenetic endemism types and to plot those results spatially.

Drivers of phylodiversity

Assembly of explanatory variables for diversity patterns. We used seven variables to analyze the observed phylogenetic diversity patterns. These included four bioclimatic variables (annual mean temperature, annual precipitation, temperature seasonality [standard deviation * 100], and precipitation seasonality [coefficient of variation] ([Fick and Hijmans, 2017](#))), two climate stability variables (temperature stability and precipitation stability), and elevation. The climate stability variables represent the inverse of the mean standard deviation between equally spaced 1000-year time slices over the past 21,000 years and were provisioned from [Owens and Guralnick \(2019\)](#). These seven layers were chosen because they capture the geographic variation in climate stability over a significant transition from a full glacial to interglacial time-period, which likely is representative of similar transitions that occurred repeatedly in the Pleistocene ([Waltari et al., 2007](#)). Elevation values at this scale also provide a reasonable proxy of current topographic heterogeneity. All environmental variables were scaled to a mean of zero and SD of one and resampled to 100 km × 100 km.

Testing importance of climate and topographic drivers. Four diversity metrics (PD, RPD, and randomization tests for both measures) were utilized to test the most predictive explanatory variables of diversity and diversity significance. Phylogenetic significance analyses were derived from the randomized metrics and divided into binomial datasets, where significantly high values were scored with the value one and significantly low values assigned zero. Cells with non-significant values were excluded from the binomial logistic regressions.

We fit generalized linear models (GLMs) using the climatic and terrain variables described above as predictors and phylogenetic metrics as response variables. For the PD and RPD analyses, we used the Gaussian distribution; for models examining PD/RPD significance, we used the binomial distribution with the logit link function. We used the dredge function from the package MuMln ([Barton, 2009](#)) in R version 3.6.2 to examine all possible models. We used an information-theoretic approach using Akaike's Information Criterion (AIC) to rank models ([Burnham and Anderson, 2002](#)). Models that were a subset of another model examined were not considered to be competitive if within delta AIC ≤ 2. We examined the collinearity of variables of the models by calculating variance-inflation factors (VIF) using the car package, and models with VIF ≥ 5 were also not considered as competitive models. We used delta AIC values and Akaike weights (wi) to rank competing models.

Spatial autocorrelation can lead to an increase in type I errors when building simple linear models, because observations are effectively pseudoreplicated ([Bini et al., 2009](#); [Borcard et al., 2011](#)), and therefore it is challenging to properly assess the effect of predictor variables. To test for spatial autocorrelation in our top GLMs, we calculated Moran's I. Spatial autocorrelation was detected in all of our top GLMs, and we therefore fit spatial generalized linear mixed models (GLMMs) with spatially correlated random effects in R using the package spaMM ([Rousset and Ferdy, 2014](#)). We used Matern correlation models to generate the spatially correlated random effects included in the spatial model ([Rousset and Ferdy, 2014](#)). We present the results of the spatial GLMMs in the text below, but the results of our non-spatial GLMs can be found in the [supplemental information \(Table S2\)](#).

Comparison of butterfly and plant phylodiversity. We re-ran a recently published analysis of seed plant phylodiversity ([Mishler et al., 2020](#)) but excluded gymnosperms in order to compare our results

with spatial phylogenetic patterns for North American flowering plants. This re-analysis used the same methods as in Mishler et al. and resulted in inclusion of 19,173 angiosperm terminal taxa (and exclusion of 476 gymnosperms) across North America. We applied the same metrics for angiosperms as for butterflies and used nearly the same spatial extent, only excluding a small portion of the southern tip of Mexico, at 50-km resolution. We resampled gridded analysis products from Mishler et al. into the same 100-km resolution as the butterfly grids using bilinear interpolation for comparisons of associations between butterfly and plant phylodiversity metrics. We next generated univariate linear regression models, where plant PD, RPD, and PE values were the predictor variables and the corresponding butterfly PD, RPD, and PE values the response variables. The residuals of these models were then mapped spatially to display where butterfly diversity was higher or lower than predicted by the model. Although we did expect spatial autocorrelation in these simple linear models, we did not generate GLMMs with spatially autocorrelated random effects, because our main objective was to visualize the spatial pattern of the model residuals. Comparison of PD, RPD, and CANAPE significance for butterflies and flowering plants was done by visual inspection because a summary test was gauged to be superfluous given the striking regional differences between the two groups. We chose to compare butterfly phylodiversity with angiosperm phylodiversity instead of all seed plants, because far more butterfly host plants are angiosperms than gymnosperms (Narango et al., 2020). However, angiosperm phylodiversity was remarkably similar to seed plant phylodiversity (Figure S1), and similar comparisons were produced with both plant datasets.

RESULTS

A phylogeny for North American butterflies

A total of 1,437 (74.6%) known butterfly species had sequence data already available or were sequenced *de novo* for COI (Figure 1). *De novo* sequencing led to the addition of 140 species that otherwise lacked sequence data in public repositories. Of these 140 species, 96 (68.6%) were distributed only in Mexico. Although the backbone of the butterfly tree was constrained at the family level, subfamily and tribe-level relationships generally agreed with those of prior studies (supplemental information), and clade-based ages were largely congruent with recent butterfly-wide dating analyses (Table S3; See “SupDryad_treep1.tre” in Earl (2021)). We recovered a median bootstrap value across the entire tree of 86 using transfer bootstrapping (See “SupDryad_BOOST.tre” in Earl (2021)) (Lemoine et al., 2018).

Observed patterns of diversity and endemism

Maps of observed species richness (Figures 2A and S2A) and phylogenetic diversity (Figures 2B and S2B) both documented highest diversity primarily in the tropical dry and wet forests in Mexico and the lowest values across the arctic of Canada. Patterns of richness and PD are complex in western North America, likely reflecting the heterogeneous landscape, with peaks in areas adjoining the Sierras and Rocky Mountains. Relative phylogenetic diversity (RPD) peaked in wet and dry tropical forests in Mexico and remained uniformly high across the Eastern Temperate Forest, Great Plains, and southern deserts (Figures 2C and S2C). By comparison, observed RPD was lower across much of the temperate Intermountain West, the Mediterranean regions of California, and into northern ecosystems such as boreal forests and taiga. Phylogenetic endemism (PE; Figures 2D and S2D) showed the same general latitudinal gradient as PD and RPD but included areas of higher phylogenetic endemism along the temperate Sierra Madre mountain ranges in Mexico and in the coast ranges in the Pacific. PE was overall higher in the temperate west than in the east and associated with transition zones in the Rockies and Sierra Nevada.

Spatial randomization tests

We uncovered highly regionalized patterns of overdispersion and clustering based on PD randomizations (Figure 3B). All boreal, taiga, and tundra regions showed lower than expected PD, indicative of phylogenetic clustering. Most of the temperate regions in the west, including diverse ecoregions in cold deserts, west coast forests, and Mediterranean portions of California, also displayed clustering. In contrast, tropical wet and dry forests, the most phylodiverse areas in North America, showed higher than expected PD, or phylogenetic overdispersion, when compared with null models. We also note that portions of the south-central semi-arid prairies also showed higher than expected phylodiversity. The southern, warm deserts and eastern temperate forests did not show significantly high or low PD.

The RPD randomization indicated that southern portions of North America have communities containing longer branches than expected under null models (Figure 4B). This included not only tropical regions

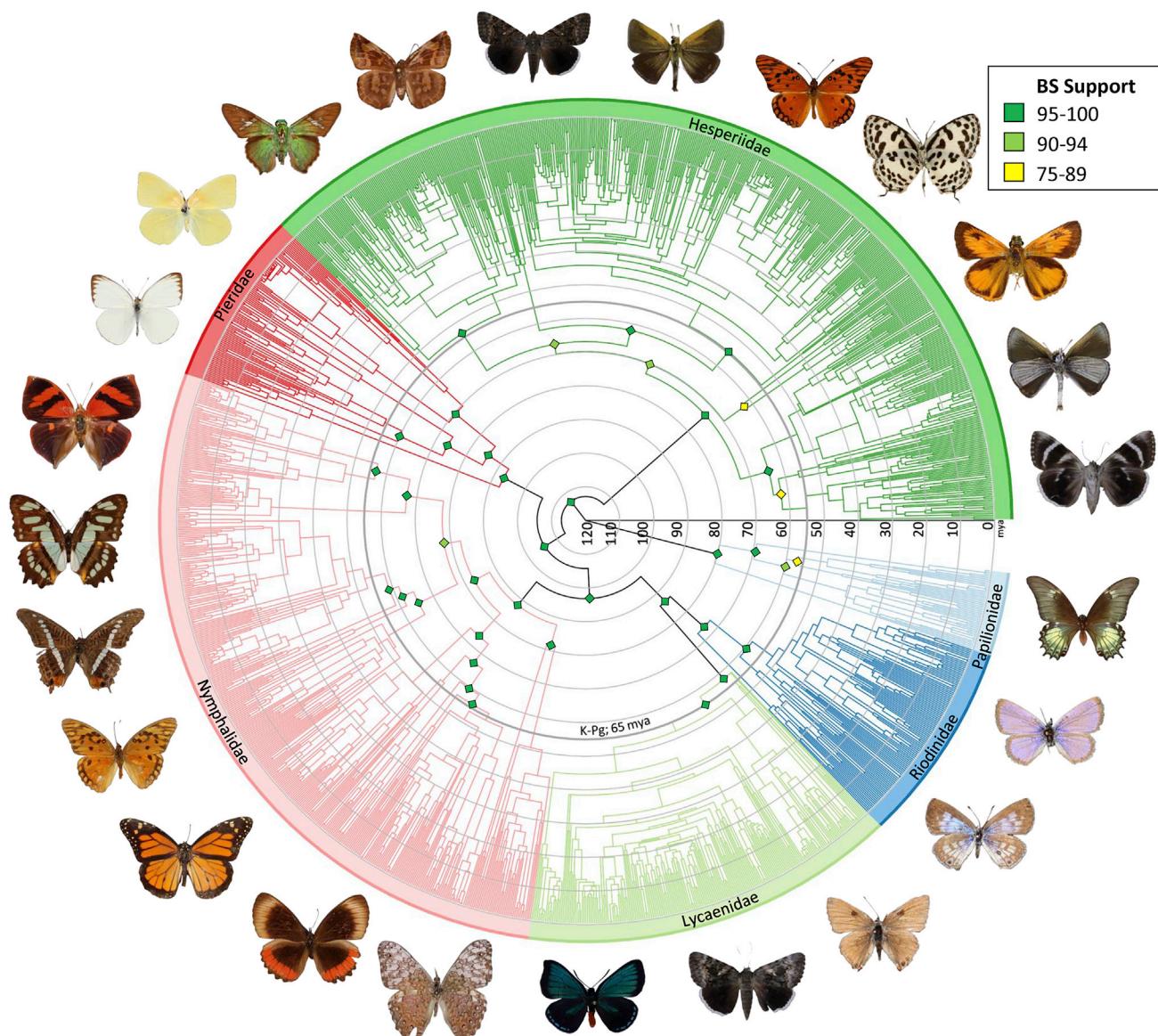


Figure 1. A time-calibrated tree of 1,437 North American butterflies with bootstrap support shown for 39 of the deepest nodes (before the K-Pg boundary)

but also semi-arid highlands and southern deserts into semi-arid plains and prairie. On the other hand, shorter than expected branch lengths were found across much of the Sierra Nevada, Rockies, and Intermountain West. We found no significant RPD in the Eastern Temperate Forest, northern Great Plains, and northernmost portions of North America.

CANAPE

Regions of significant neoendemism were located in the California Mediterranean region and western forests, including the Cascades, Coast Ranges, and Sierra Nevada (Figure 5B) and in transition zones across lower-elevation regions to the East. Mixed patterns of endemism with both paleo- and neo-endemics were found in predominantly warm deserts and the southeastern coastal plain and southern, subtropical portions of Florida. Sites dominated by paleoendemism were more rare, only indicated in some areas in tropical Mexico.

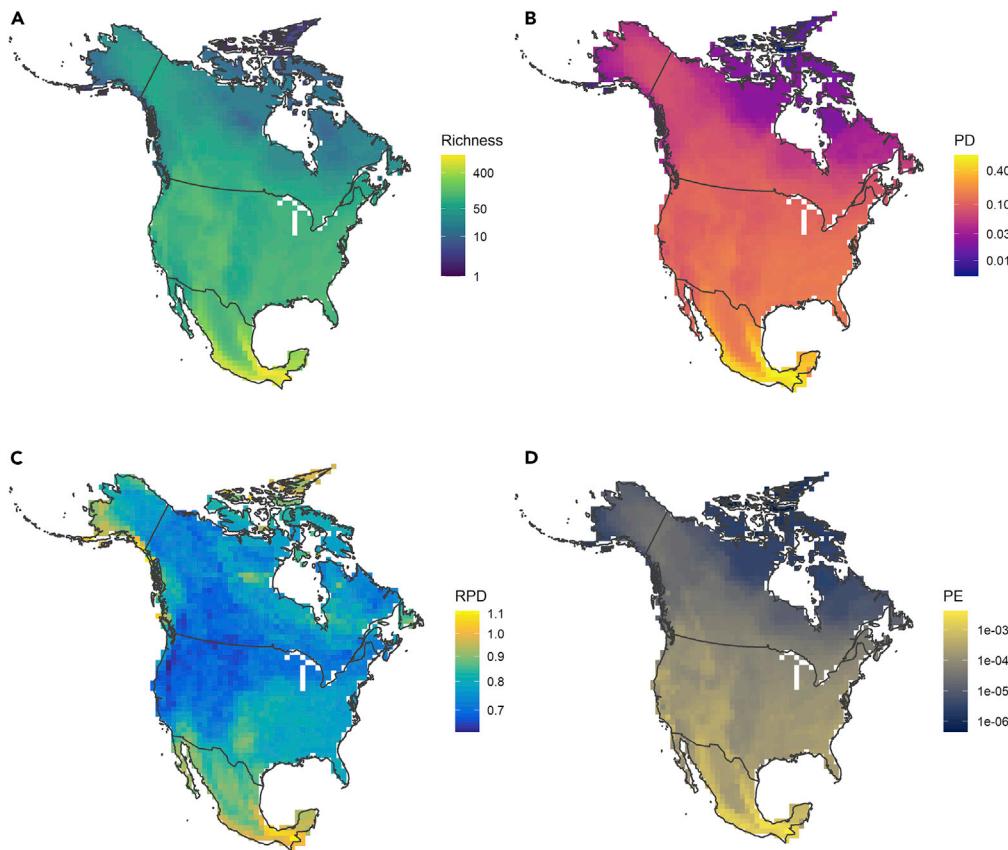


Figure 2. Diversity and endemism patterns

Observed values for North American butterflies for: (A) toxic richness, (B) phylogenetic diversity (PD), (C) relative phylogenetic diversity (RPD), and (D) phylogenetic endemism (PE). Maps without logarithmically scaled color palettes can be viewed in [Figure S2](#).

Drivers of phylodiversity

Annual mean temperature was the most important environmental variable in predicting PD, followed by mean annual precipitation and precipitation seasonality ([Table 2](#)). Areas that were warmer and wetter generally had the highest PD. As well, higher elevation areas and those with more seasonal precipitation generally had higher PD, although these are weaker effects. PD significance differed from results for observed PD, with temperature being the only covariate that had a coefficient estimate larger than its standard error when accounting for spatial autocorrelation ([Table 2](#)). Warmer areas were more likely to have significantly high PD. Lower RPD in an area indicates relatively short branches, potentially indicative of more recent radiations. Results from analysis of climate and terrain drivers showed high RPD in areas that have higher temperature stability and higher precipitation. As well, areas with low elevation had higher RPD ([Table 2](#)). Spatial GLMMs showed that areas with significantly low RPD are in colder areas with less precipitation; however, the coefficient estimates of all covariates were less than the standard error.

Similarities and differences between butterfly and plant phylodiversity

Butterflies and plants of North America displayed a similar pattern of PD ($r^2 = 0.34$), and areas of discordance displayed moderate spatial structuring of residuals in our simple linear models with butterfly PD as a response variable to plant PD. Butterfly PD was higher in the tropics and lower along the west coast of North America than predicted by angiosperm PD ([Figure 6](#)). Surprisingly, butterflies and plants did not have similar patterns of RPD ($r^2 = 0.01$), and RPD showed strong spatial structuring of linear model residuals, with butterfly RPD being much higher in the west and lower in the south than predicted based on angiosperm RPD ([Figure 6](#)). Butterflies and plants both showed a pattern of having the highest PE values in southern Mexico, but overall, the similarity across North America was relatively weak ($r^2 = 0.10$). The spatial

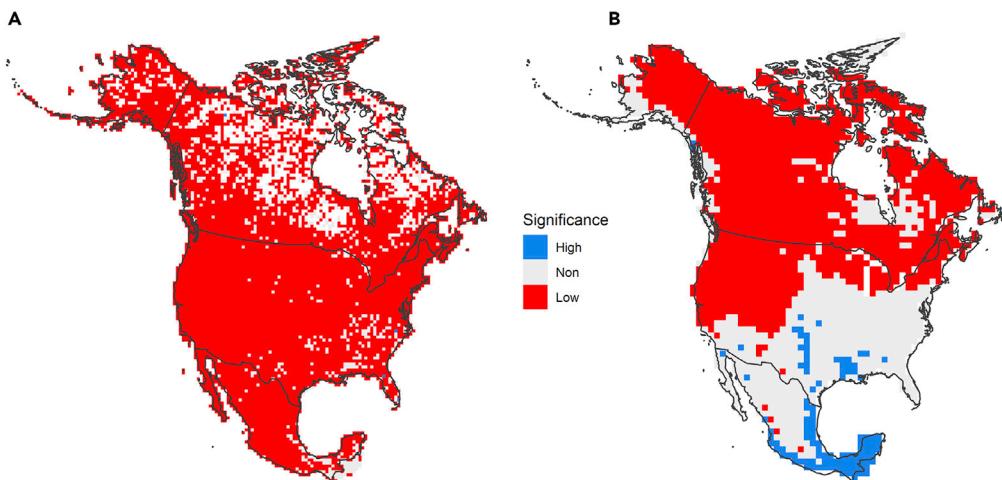


Figure 3. Statistical significance of PD

Statistical significance of phylogenetic diversity (PD) for (A) angiosperms and (B) butterflies across North America. Areas with significantly high values have taxa that are less closely related than expected by chance (blue), whereas areas with significantly low values have taxa that are more closely related than expected by chance (red).

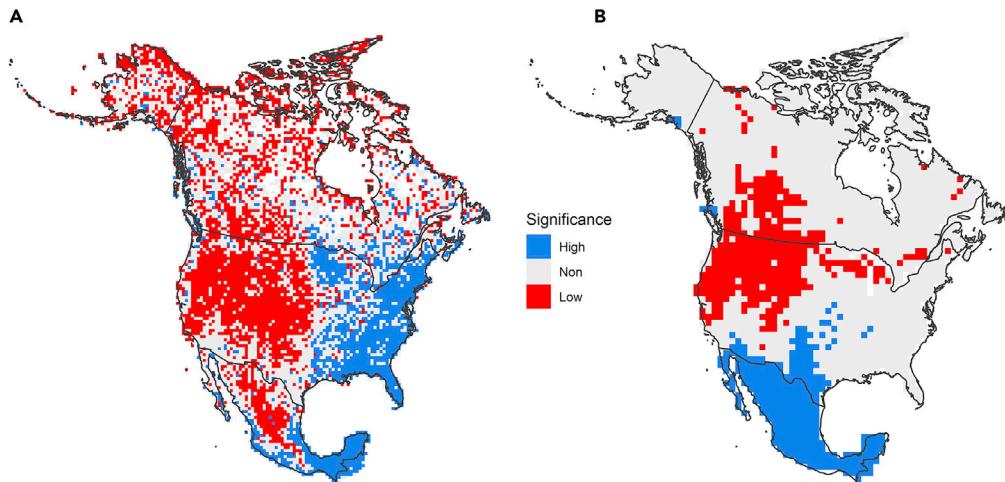
residuals of the PE linear models mirrored PD in the southern portions of the continent but without spatially structured error in temperate regions of the continent (Figure 6).

Butterflies showed a strikingly different pattern of PD significance compared with angiosperms, with higher-than-expected values in the south and lower-than-expected values in the north (Figure 3). Angiosperms showed a strong pattern of having significantly lower-than-expected values of RPD in western North America and significantly high RPD in southern Mexico and eastern North America (Figure 4). Butterflies also had significantly high areas of RPD in tropical wet and dry forests in southern Mexico, but unlike flowering plants, they also exhibited high RPD in Baja California and the American Southwest. Also unlike angiosperms, butterflies did not show significantly higher RPD in eastern temperate forests. Both groups showed significantly low RPD in much of western North America (Figure 4).

Flowering plants and butterflies showed generally discordant patterns of endemism. Centers of mixed paleo- and neo-endemism for angiosperms were found in Mexico, including the Baja California peninsula, as well as in Florida and the adjoining southern coastal plain. Although CANAPE results for butterflies also showed mixed phylogenetic endemism in Florida, the results for the groups were otherwise quite different. Butterflies showed strong patterns of mixed endemism north of Mexico, in the warm deserts and portions of the colder deserts of the Southwest, along with predominantly neoendemism in coastal regions of the West and limited paleoendemism in southern Mexico (Figure 5).

DISCUSSION

We present the first continental-scale phylogenetic diversity analyses for butterflies, focusing on North America. This analysis is notable for being relatively complete, with coarse-scale distribution data for all species and a phylogeny with ~75% sampling of North American species. This level of completeness provides, for the first time, a well-resolved, continental-scale view of phylogenetic diversity for an entire insect suborder. We also extended the range estimates beyond North America, by gathering very coarse country-level range maps for the ~25% of butterfly species that have ranges outside of our defined North America boundaries. We argue this approach is better than simply truncating ranges of terminals for any analyses relying on a range-weighted metric, such as PE. Here, not including full ranges would have led to many neotropical butterflies having much smaller ranges ending at the border of Mexico rather than properly extending into Central and South America. However, even when one extends the range estimates of terminal taxa, this does not fully solve the "edge-effect" problem, because the range sizes of related taxa occurring outside the study area are still being left out, which means ranges of deeper branches may be poorly estimated, affecting PE. Thus, any study incorporating less than a globally complete assessment is still

**Figure 4. Statistical significance of RPD**

Statistical significance of relative phylogenetic diversity (RPD) for (A) angiosperms and (B) butterflies. Areas in blue have significantly longer branches than expected; areas in red have significantly shorter branches than expected.

assessing primarily local endemism patterns (Daru et al., 2020). However, this effect will be relatively small for any branch with at least one wide-ranging terminal in the dataset, as the ranges of internal branches are calculated as the unions of the ranges of their terminal taxa, and such wide-ranging branches are strongly downweighted in the calculations.

Below we discuss how our results address key predictions regarding patterns of butterfly diversity across a continent with enormous habitat breadth, from the hot and dry deserts in the Southwest to the wet, tropical forests in the Yucatan, and the cold, arid environments of the taiga and tundra. In particular, we focus on processes that are likely to have shaped this diversity, based on both examination of climatic and topographic drivers, and via comparison with angiosperms. We explicitly predicted concordance of patterns and process because of strong associations between butterflies and their flowering plant hosts and because of the potential for similar responses to environmental changes given co-distribution across the same latitudinal and elevational gradients.

Patterns and drivers of phylogenetic diversity and endemism

Our results point especially to the importance of current climate drivers on phylogenetic diversity. We found that PD is highest in the warmest, wettest areas of the continent, along with regions of seasonal precipitation and along elevational gradients. RPD results point not to temperature, but rather temperature stability and precipitation as drivers. RPD was highest in regions exhibiting temperature stability over the past 21,000 years and high rainfall, perhaps representing conditions where highly divergent lineages could co-exist and where flowering plant diversity may also be unusually high. However, despite our predictions that temperature and precipitation stability would be a key predictor of PD and RPD significance, they were never a top predictor in any model. Rather, current annual temperature was the dominant driver of PD significance, and none of our predictor variables were particularly useful in determining drivers of RPD significance after including spatially correlated random effects. Below we summarize these results more thoroughly by major regions across North America, focusing on synthesis across geographic distance and environmental gradients.

Northern North America: we define Northern North America as regions that were mostly covered in ice during the Last Glacial Maximum 21,000 years ago. The northern portion of the continent showed low PD but not RPD. The former suggests the importance of environmental filtering due to cold and seasonal conditions and the latter orbitally forced range expansions and contractions, but limited cladogenesis, across glacial cycles (Jansson and Dynesius, 2002). We argue that environmental filtering is more likely in a volant group with high reproductive rates (Pellissier et al., 2013), such as butterflies, in comparison to clades where dispersal can often lag behind changing conditions (Alexander et al., 2018). This is further

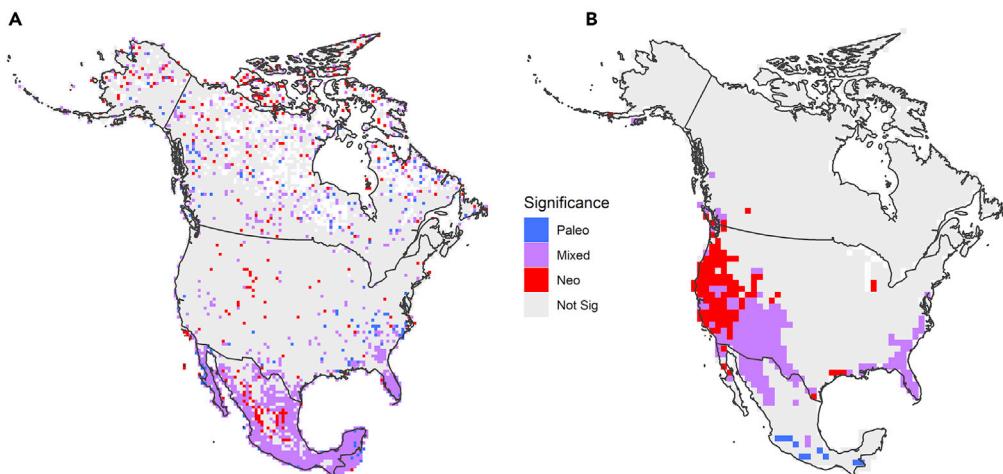


Figure 5. CANAPE results

CANAPE results showing statistically significant centers of phylogenetic endemism for (A) angiosperms and (B) butterflies. All cells that are colored have significantly high PE. Red cells have concentrations of rare short branches (neoendemism); blue cells have concentrations of rare long branches (paleoendemism), and purple cells have mixtures of neo- and paleoendemism.

supported by low PE and non-significant RPD and CANAPE results in this region (Figures 4 and 5), suggesting most species are wide-ranging and not recently radiating across the North.

Western United States: western areas south of past ice sheets and north of the warm deserts showed very strong patterns of significantly low PD and RPD. The former indicates potential environmental filtering, given steep elevational and climatic gradients that themselves were in flux during glacial-interglacial cycling during the Pleistocene (Thackray, 2008), whereas the latter might indicate that butterflies have undergone recent radiations in these areas. The potential for recent radiations would also suggest high levels of neoendemism in the region. This was not the case for the Intermountain West and Rocky Mountains; however, we did recover a strong signal of neoendemism in Mediterranean portions of California and western forests, extending into the western portions of the Great Basin.

Southern North America: the southern portion of North America showed particularly surprising results, especially in the warm deserts. PD and RPD were both significantly high in tropical regions of North America, consistent with the tropics as a museum for butterfly diversity (Farrera et al., 1999; Hostetler et al., 1999). In the Temperate Sierra and warm deserts, we found significantly high RPD but no indication of clustered or overdispersed PD. This novel result suggests phylogenetically old communities of butterflies in deserts, a climate that formed recently, during the mid-Pliocene (Axelrod, 1958). It has long been known that flowering plants found in this region are derived from related lineages in thornscrub and arid highlands that are phylogenetically much older (Axelrod, 1959). Butterflies in warm deserts may therefore also be connected to older lineages that persisted in subtropical, yet still seasonal, habitats in southern regions and less closely related to species in colder deserts in the Great Basin. Bioregionalizations and their associations derived from phylogenetic beta-diversity provide a next-step means to examine such questions (Mienna et al., 2020).

East of the Rockies: areas east of the Rockies, including the Great Plains and eastern temperate forests, are unremarkable in PD or RPD, which contrasts with spatial phylogenetic findings for flowering plants, as we discuss in detail below. We were particularly surprised that the Great Plains region, which became grassland-dominated during cooling in the Miocene and Pliocene, did not show accumulation of younger than expected lineages (i.e. significantly low RPD), as has been documented in other groups (Mishler et al., 2020). As well, tropical regions of Florida were not significantly higher in PD or RPD. However, tropical Florida and the nearby coastal plain do show significantly high levels of mixed PE, aligning with a known plant biodiversity hotspot (Myers et al., 2000).

Table 2. Summary of the top spatial GLMMs for PD, RPD, significant PD, and significant RPD

Model	Temp	Prec	Temp seas	Prec seas	Temp stab	Prec stab	Elev
PD	0.014 ± 0.005	0.010 ± 0.001		0.003 ± 0.001	-0.004 ± 0.005		0.003 ± 0.001
RPD		0.005 ± 0.002	0.003 ± 0.003	-0.002 ± 0.002	0.010 ± 0.005	-0.001 ± 0.003	-0.002 ± 0.001
PD Sig	53.1 ± 30.3	2.98 ± 18.82			-7.4 ± 21.64		1.55 ± 11.79
RPD Sig	57.1 ± 133	9.9 ± 28.8			-11.3 ± 72.4		7.6 ± 29.5

Numbers in the columns indicate changes in PD, RPD, and significant PD and RPD values ± standard error when variable values between locations increased by one standard deviation. Bolding denotes coefficients whose absolute values are greater than their standard error.

Comparisons of spatial phylodiversity patterns between butterflies and flowering plants

Given strong ecological associations and shared abiotic drivers that have played out over long evolutionary timeframes, we predicted butterflies and plants to have similar patterns of phylodiversity (Kumar et al., 2009), while also acknowledging alternate hypotheses. This prediction is generally borne out in the western and northern portions of the continent, both shaped by recent perturbations including glaciation and aridification. However, our analyses also revealed striking differences. For example, butterflies and flowering plants did not show similar patterns of RPD, suggesting that diversification timing and rates between butterflies and angiosperms may not be associated at the scale and extent of this analysis. The reasons for these differences may be partially methodological. However, these results also suggest that shared historical forces and strong ecological associations can still lead to divergent historical and current biogeographic outcomes. We discuss more about methodological and biological rationales for similarities and differences below.

Although we used consistent PD metrics across both studies, allowing for direct comparisons of outputs, sampling completeness varies dramatically between flowering plant and butterfly analyses. Sampling of flowering plants, which encompasses more than an order of magnitude more named species than butterflies (Mishler et al., 2020), is less complete in terms of both phylogenetic (ca. 44% of taxa included) and spatial distribution information. As problematic, both phylogenetic and spatial sampling is known to be biased. Some regions, especially in the North, are still nearly unsampled in terms of digitally accessible flowering plant specimen records, based on results in Mishler et al. This contrasts sharply with other regions, such as coastal California, where species sampling is mostly complete at this scale.

These differences in completeness of sampling and spatial bias make strong assessments of patterns more challenging. Mishler et al. recovered a pattern of significantly low PD across the continent for seed plants, and our more phylogenetically restricted analysis of flowering plants shows the same result. This suggests that co-occurring species are always more closely related than expected by chance compared with the full pool of species. This result, not seen in butterflies, might be affected to some extent by incomplete sampling, but it is such a strong and uniform result that it likely points to some fundamental differences in evolutionary ecology between butterflies and plants. Significantly low PD (phylogenetic clustering) is most often taken to indicate habitat filtering due to phylogenetically conserved habitat preferences that result in close relatives tending to occur together in communities. It may well be that habitat preference has a higher level of conservation in seed plants than in butterflies, a possibility in need of future research.

Although it has long been known that western North America has been dramatically reshaped by regional tectonism, orogeny, and climatic changes, the full magnitude of those impacts on flora and fauna besides vertebrates (Badgley, 2010) is just now starting to be understood (Pellissier et al., 2018). Plant spatial phylogenetic work (Mishler et al., 2020) has confirmed this in a spectacular fashion, with eastern temperate forests showing significantly older plant lineages than in the Great Plains and western portions of North America, both strongly shaped by cooling and aridification, showing more recent diversifications. We expected to find congruent results when examining RPD in butterflies. However, RPD between the two groups is not strongly correlated. Butterfly communities are comparatively younger in the West based on the spatial residual plots (Figure 6). As well, although some portions of the West show lower-than-expected RPD for both plants and butterflies, suggestive of more recent radiations there compared with other regions, we did not recover higher-than-expected butterfly RPD in the East as we did in plants. Much more of Mexico was significantly high in RPD in butterflies than in plants.

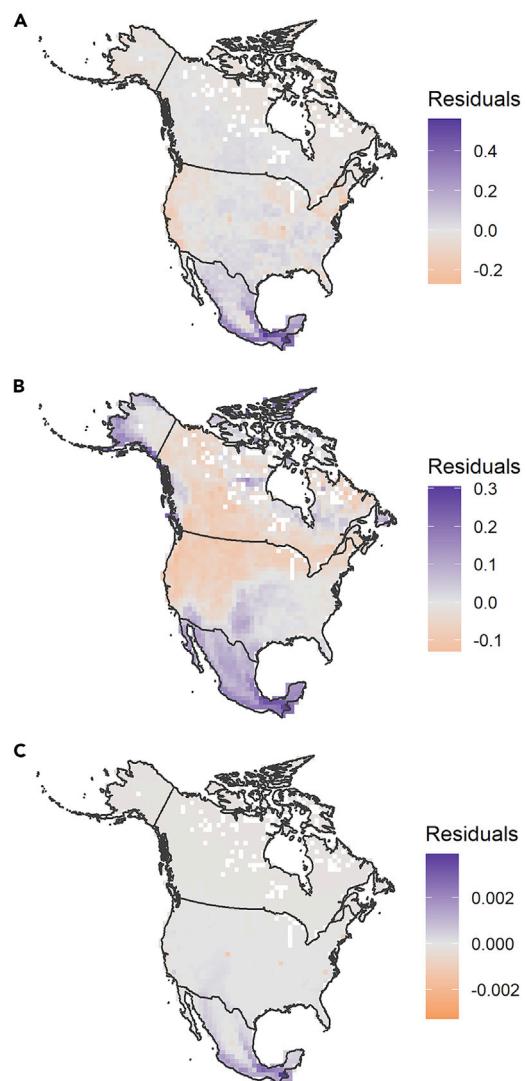


Figure 6. Spatial residuals of univariate linear regressions for observed PD (A), RPD (B), and PE (C), where angiosperm metrics were used to predict butterfly metrics. High residual values (blue) represent areas where butterfly values are higher than predicted values.

These results suggest that less stable areas such as northern and western portions of North America may show moderate discordance when comparing across groups at different trophic levels but in a consistent manner. Butterflies likely have diversified in the shadow of a persistent, highly diverse angiosperm-dominated forest in eastern temperate North America. Given massive inequality in numbers of butterfly to plant species in the region, providing ample opportunity for evolving new host relationships, the overall effect is likely equilibrium between the two groups. In the West, more extensive perturbation caused by loss of continuous forests and continuing cooling and drying likely drove stronger disequilibrium, with butterflies following bursts of new plant lineages forming in that region. Spatial residual plots for relative phylogenetic diversity are supportive of this scenario (Figure 6), but further examination, in other herbivores, is warranted to see if such ordering effects may be more general. We hypothesize that areas with more active geologic histories will show evidence of lags in community ages between hosts and consumers/pollinators.

Phylogenetic endemism patterns for plants and butterflies as seen in CANAPE were surprisingly discordant, with much stronger plant endemism found in southern and central Mexico compared with butterflies. Mishler et al. truncated ranges at southern edges of the region of interest, which might have led to more

artificial range restrictions and higher endemism. This is particularly likely given that many widespread species with wet tropical affinities have range edges in the Yucatan of Mexico. Still, the discordance between plant and butterfly endemism is notable and may point to fundamental differences in ecological, evolutionary, and biogeographic processes between butterflies and plants worthy of further study. We call particular attention to the strong pattern of mixed butterfly phylogenetic endemism in all of the warm deserts of North America. Although more work is needed, it would be unsurprising if these warm deserts were generally areas of diversification and endemism for many clades, based on continuing floristic (Sosa et al., 2020) and faunistic work (Riddle and Hafner, 2006).

Importance for conservation

Butterflies are under threat, perhaps represented most iconically by monarchs and their decline (Thogmartin et al., 2017; Agrawal and Inamine, 2018), but the entire fauna may be imperiled (Wepprich et al., 2019). Results here provide a needed step for better prioritizing areas of highest conservation need. In particular, we document areas high in PE, PD, and RPD, which collectively harbor more accumulated evolutionary history (Redding et al., 2008) and which also contain the most range-restricted lineages. These areas are likely to be high priorities for conservation (Davies and Cadotte, 2011; Kling et al., 2019). Within North America, there are four well-documented biodiversity hotspots: the California Floristic Province, North American Coastal Plain, Madrean Pine-Oak Woodlands, and Mesoamerica (Reid, 1998; Noss et al., 2015). These areas were recognized due to their high plant diversity and endemism and threat of extinction (Myers et al., 2000). Although some hotspots have been long recognized, understanding where diversity is highest, rarest, and most threatened is still a work in progress; for example, the North American Coastal Plain region was only recently recognized as a hotspot (Noss et al., 2015), and further work across the Tree of Life is still needed to discover areas harboring the most unique and rare diversity. Butterfly PD and PE patterns are also higher than expected within all four of these hotspots, strengthening arguments about protecting habitat in these areas. However, our results also uncovered a new hotspot showing significant endemism, relatively old lineages (based on significantly high RPD), and high PD in the warm deserts of North America. Our results make a strong case for habitat conservation across warm deserts in particular, especially because they are not already documented biodiversity hotspots.

Limitations of the study

This work points the way to still broader and more detailed future projects on insect and plant spatial phylogenetics. Key next steps include further closing spatial and phylogenetic knowledge gaps regionally and globally. On the spatial side, this work relied on coarse-grain range maps, limiting finer localizations along shorter spatial and environmental gradients such as in topographically heterogeneous regions (Bini et al., 2009). On the phylogenetic side, although our species sampling is relatively strong, covering ~75% of North American species, many tips are represented by a single locus, i.e. COI. This leads to limited locus coverage (69% missing data), which although not uncommon in supermatrix approaches (Sanderson et al., 2010), is still an area where sampling improvement is warranted. We attempted to mitigate issues with coverage, in part, by using a maximum likelihood framework because Bayesian frameworks tend toward more biased topologies due to interactions between missing data and priors (Lemon et al., 2009). We further opted to use a family-level backbone constraint from Espeland et al. (2018), although initial, unconstrained analyses resulted in the monophyly of all seven families. We also note issues with time calibration, and although congruification methods for dating trees are becoming commonplace, issues can arise when relationships at the analogous nodes between the target tree and the reference tree differ (Eastman et al., 2013). Because our custom-built phylogeny was constrained at the analogous nodes of the reference tree (Espeland et al., 2018), these issues are at least minimized here. Finally, we did not directly consider known butterfly host-plant associations in this work, which may be especially important, given recent work finding the majority of Lepidoptera are supported by a few important plant genera (Narango et al., 2020). That information, incorporated into a spatial phylogenetic framework, would deliver a stronger process-oriented understanding of spatial co-diversification that has shaped terrestrial ecosystems.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Chandra Earl (sunray1@ufl.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All sequence data custom generated from this project are available on the Barcode of Life Database (<http://dx.doi.org/10.5883/DS-56788>) and GenBank (Accession Numbers MW807620 - MW807751). Gridded summary products, DNA alignments, phylogenies and information about locus origin are available on Dryad (<https://doi.org/10.5061/dryad.00000002>). Maps and country level list data for taxa with ranges outside of North America are available on Dryad (<https://doi.org/10.5061/dryad.00000002>). The final phylogeny is available on Open Tree of Life (https://tree.opentreeoflife.org/curator/study/view/ot_2014).

METHODS

All methods can be found in the accompanying [transparent methods supplemental file](#).

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2021.102239>.

ACKNOWLEDGMENTS

We thank the Map of Life team for access to range products, especially for Mexican species. We also thank Caroline Storer for reviewing and providing valuable edits to the manuscript and Andrew Ryan, Minji Ku, and Lori Dever for helping with range map digitization and validation. We thank three peer reviewers for comments that improved the manuscript. All range products used here can be visualized at Map of Life (mol.org). CE, VB, AK, and RG were supported by the ButterflyNet project (DEB-1541500). MB was supported by a University of Florida Biodiversity Institute (UFBI) Fellowship.

AUTHOR CONTRIBUTIONS

CE, MB, SL, VB, and RG designed the experiment, gathered the data, and fit the models. RG, AK, CE, and MB wrote the manuscript with input from all authors. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 29, 2020

Revised: November 13, 2020

Accepted: February 23, 2021

Published: March 23, 2021

REFERENCES

- Agrawal, A.A., and Inamine, H. (2018). Mechanisms behind the monarch's decline. *Science* 360, 1294–1296.
- Alexander, J.M., Chalmandrier, L., Lenoir, J., Burgess, T.I., Essl, F., Haider, S., Kueffer, C., McDougall, K., Milbau, A., Nuñez, M.A., et al. (2018). Lags in the response of mountain plant communities to climate change. *Glob. Chang. Biol.* 24, 563–579.
- Allen, J.M., Germain-Aubrey, C.C., Barve, N., Neubig, K.M., Majure, L.C., Laffan, S.W., Mishler, B.D., Owens, H.L., Smith, S.A., Whitten, W.M., et al. (2019). Spatial phylogenetics of Florida vascular plants: the effects of calibration and Uncertainty on diversity estimates. *iScience* 11, 57–70.
- Axelrod, D.I. (1958). Evolution of the madro-tertiary geoflora. *Bot. Rev.* 24, 433–509.
- Axelrod, D.I. (1959). Late cenozoic evolution of the sierran bigtree forest. *Evolution* 13, 9.
- Badgley, C. (2010). Tectonics, topography, and mammalian diversity. *Ecography* 33, 220–231.
- Barton, K. (2009). Mu-MIn: Multi-Model Inference. R Package Version 0.12.2/r18, <https://cran.r-project.org/web/packages/MuMIn/index.html>.
- Barve, V. (2020). Taxotools: Tools to Handle Taxonomic Data (R Package V 0.0.43). <https://cran.r-project.org/web/packages/taxotools/index.html>.
- Bini, L.M., Diniz-Filho, J.A.F., de Marco, P., Jr, Terribile, L.C., Rangel, T.F.L.V.B., Akre, T.S.B., Albaladejo, R.G., Aparicio, A., Albuquerque, F.S., Montoya, D., et al. (2009). Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. *Ecography* 32, 193–204.
- Bintanja, R., and van de Wal, R.S. (2008). North American ice-sheet dynamics and the onset of 100,000-year glacial cycles. *Nature* 454, 869–872.
- Bonelli, S., Casacci, L.P., Barbero, F., Cerrato, C., Dapporto, L., Sbordoni, V., Scalercio, S., Zilli, A., Battistoni, A., Teofili, C., et al. (2018). The first red list of Italian butterflies. In *Insect Conservation and Diversity*, 11, S. Leather and P. DeVries, eds. (Insect Conserv Divers), pp. 506–521.
- Borcard, D., Gillet, F., and Legendre, P. (2011). *Numerical Ecology with R* (Springer).
- Brock, J., and Kaufman, K. (2003). *Field Guide to Butterflies of North America* (Houghton Mifflin).

- Bronstein, J.L., Alarcón, R., and Geber, M. (2006). The evolution of plant-insect mutualisms. *New Phytol.* 412–428.
- Brower, A.V.Z. (1996). Parallel race formation and the evolution and the evolution of mimicry in *Heliconius* butterflies: a phylogenetic hypothesis from mitochondrial DNA sequences. *Evolution* 50, 195–221.
- Burnham, K.P., and Anderson, D.R. (2002). *Model Selection and Multimodel Inference : A Practical Information-Theoretic Approach* (Springer-Verlag).
- ButterflyNet. (2020). <https://www.butterflynet.org/>.
- Clark, K., Karsch-Mirzachi, I., Lipman, D.J., Ostell, J., and Sayers, E.W. (2016). GenBank. *Nucleic Acids Res.* 44, 67–72.
- Cowling, R.M., and Lombard, A.T. (2002). Heterogeneity, speciation/extinction history and climate: explaining regional plant diversity patterns in the Cape Floristic Region. *Divers. Distrib.* 8, 163–179.
- Danks, H.v. (1994). Regional diversity of insects in North America. *Am. Entomol.* 40, 50–55.
- Daru, B.H., Farooq, H., Antonelli, A., and Faurby, S. (2020). Endemism patterns are scale dependent. *Nat. Commun.* 11, 2115.
- Davies, T.J., and Buckley, L.B. (2011). Phylogenetic diversity as a window into the evolutionary and biogeographic histories of present-day richness gradients for mammals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 2414–2425.
- Davies, T.J., and Cadotte, M.W. (2011). Quantifying biodiversity: does it matter what we measure? In *Biodiversity Hotspots*, F. Zachos and J. Habel, eds. (Berlin, Heidelberg: Springer), pp. 43–60.
- Dennis, R.L.H. (1993). *Butterflies and Climate Change* (Manchester United Press).
- Earl, C. (2021). Spatial phylogenetics of butterflies in relation to environmental drivers and angiosperm diversity across North America, Dataset (Dryad).
- Eastman, J.M., Harmon, L.J., and Tank, D.C. (2013). Congruification: support for time scaling large phylogenetic trees. *Methods Ecol. Evol.* 4, 688–691.
- 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.
- Faith, D.P. (1992). Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* 61, 1–10.
- Farrera, I., Harrison, S.P., Prentice, I.C., Ramstein, G., Guiot, J., Bartlein, P.J., Bonnefille, R., Bush, M., Cramer, W., von Grafenstein, U., et al. (1999). Tropical climates at the Last Glacial Maximum: a new synthesis of terrestrial palaeoclimate data. I. Vegetation, lake-levels and geochemistry. *Clim. Dyn.* 15, 823–856.
- 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.
- Fine, P.V.A. (2015). Ecological and evolutionary drivers of geographic variation in species diversity. *Annu. Rev. Ecol. Evol. Syst.* 46, 369–392.
- Futuyma, D.J., and Agrawal, A.A. (2009). Macroevolution and the biological diversity of plants and herbivores. In *Proceedings of the National Academy of Sciences of the United States of America* (National Academy of Sciences), pp. 18054–18061.
- Glassberg, J. (2018). *A Swift Guide to Butterflies of Mexico and Central America*, Second Edition. (Princeton University Press).
- Godfray, H.C.J., Lewis, O.T., and Memmott, J. (2000). Studying insect diversity in the tropics. In *Changes and Disturbance in Tropical Rainforest in South-East Asia* (Phil. Trans. R. Soc. Lond. B), pp. 87–100.
- Grimaldi, D., and Engel, M.S. (2005). *Evolution of the Insects* (Cambridge University Press).
- Hostetler, S.W., Clark, P.U., Bartlein, P.J., Mix, A.C., and Pisias, N.J. (1999). Atmospheric transmission of North atlantic Heinrich events. *J. Geophys. Res.* 104, 3947–3952.
- Jansson, R., and Dynesius, M. (2002). The fate of clades in a world of recurrent climatic change: Milankovitch oscillations and evolution. *Annu. Rev. Ecol. Syst.* 33, 741–777.
- Katoh, K., and Standley, D.M. (2014). MAFFT: iterative refinement and additional methods. *Methods Mol. Biol.* 1079, 131–146.
- 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.
- Kling, M.M., Mishler, B.D., Thornhill, A.H., Baldwin, B.G., and Ackery, D.D. (2019). Facets of phylodiversity: evolutionary diversification, divergence and survival as conservation targets. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 374, 20170397.
- van Klink, R., Bowler, D.E., Gongalsky, K.B., Swengel, A.B., Gentile, A., and Chase, J.M. (2020). Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science* 368, 417–420.
- Kocher, S.D., and Williams, E.H. (2000). The diversity and abundance of North American butterflies vary with habitat disturbance and geography. *J. Biogeogr.* 27, 785–794.
- Kronforst, M.R., and Papa, R. (2015). The functional basis of wing patterning in *Heliconius* butterflies: the molecules behind mimicry. *Genetics* 200, 1–19.
- Kück, P., and Longo, G.C. (2014). FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Front. Zool.* 11, 81.
- Kumar, S., Simonson, S.E., and Stohlgren, T.J. (2009). Effects of spatial heterogeneity on butterfly species richness in Rocky Mountain National Park, CO, USA. *Biodivers. Conserv.* 18, 739–763.
- Laffan, S.W., Lubarsky, E., and Rosauer, D.F. (2010). Biodiverse, a tool for the spatial analysis of biological and related diversity. *Ecography* 33, 643–647.
- Lamas, G. (2004). *Atlas of neotropical Lepidoptera: checklist Pt. 4a hesperioidea-papilionoidea*. In *Atlas of Neotropical Lepidoptera*, J.B. Heppner, ed. (Scientific Pub), p. 439.
- Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., and Calcott, B. (2016). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* 34, 772–773.
- Lemmon, A.R., Brown, J.M., Stanger-Hall, K., and Lemmon, E.M. (2009). The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and bayesian inference. *Syst. Biol.* 58, 130–145.
- Lemmon, A.R., Emme, S.A., and Lemmon, E.M. (2012). Anchored Hybrid enrichment for massively high-throughput phylogenomics. *Syst. Biol.* 61, 727–744.
- Lemoine, F., Domelevo Entfellner, J.-B., Wilkinson, E., Correia, D., Dávila Felipe, M., De Oliveira, T., and Gascuel, O. (2018). “Renewing Felsenstein’s phylogenetic bootstrap in the era of big data.” *Nature* 556, 452–456.
- Lewis, J.J., Geltman, R.C., Pollak, P.C., Rondem, K.E., Van Belleghem, S.M., Hubisz, M.J., Munn, P.R., Zhang, L., Benson, C., Mazo-Vargas, A., et al. (2019). Parallel evolution of ancient, pleiotropic enhancers underlies butterfly wing pattern mimicry. *Proc. Natl. Acad. Sci. U S A* 116, 24174–24183.
- Lotts, K., and Naberhaus, T. (2017). Butterflies and moths of North America. <http://www.butterfliesandmoths.org/>.
- Luis-Martinez, A., Llorente-Bousquets, J., Vargas-Fernandez, I., and Warren, A.D. (2002). Biodiversity and biogeography of Mexican butterflies (Lepidoptera: papilionoidea and hesperioidea). *Proc. Entomol. Soc. Wash.* 105, 209–224.
- Mienna, I.M., Speed, J.D.M., Bendiksby, M., Thornhill, A.H., Mishler, B.D., and Martin, M.D. (2020). Differential patterns of floristic phylogenetic diversity across a post-glacial landscape. *J. Biogeogr.* 47, 915–926.
- Mishler, B.D., Guralnick, R., Soltis, P.S., Smith, S.A., Soltis, D.E., Barve, N., Allen, J.M., and Laffan, S.W. (2020). Spatial phylogenetics of the north American flora. *J. Syst. Evol.* 58, 393–405.
- Mishler, B.D., Knerr, N., González-Orozco, C.E., Thornhill, A.H., Laffan, S.W., and Miller, J.T. (2014). Phylogenetic measures of biodiversity and neo- and paleo-endemism in Australian acacia. *Nat. Commun.* 5, 1–10.
- Mittelbach, G.G., Schemske, D.W., Cornell, H.V., Allen, A.P., Brown, J.M., Bush, M.B., Harrison,

- S.P., Hurlbert, A.H., Knowlton, N., Lessios, H.A., et al. (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol. Lett.* 10, 315–331.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858.
- Narango, D.L., Tallamy, D.W., and Shropshire, K.J. (2020). Few keystone plant genera support the majority of Lepidoptera species. *Nat. Commun.* 11, 5751.
- Noss, R.F., Platt, W.J., Sorrie, B.A., Weakley, A.S., Means, D.B., Costanza, J., and Peet, R.K. (2015). How global biodiversity hotspots may go unrecognized: lessons from the North American Coastal Plain. *Divers. Distrib.* 21, 236–244.
- Owens, H.L., and Guralnick, R. (2019). climateStability: an R package to estimate climate stability from time-slice climatologies. *Biodiv. Inf.* 14, 8–13.
- Pellissier, L., Alvarez, N., Espíndola, A., Pottier, J., Dubuis, A., Pradervand, J.-N., and Guisan, A. (2013). Phylogenetic alpha and beta diversities of butterfly communities correlate with climate in the western Swiss Alps. *Ecography* 36, 541–550.
- Pellissier, L., Heine, C., Rosauer, D.F., and Albouy, C. (2018). Are global hotspots of endemic richness shaped by plate tectonics? *Biol. J. Linn. Soc.* 123, 247–261.
- QGIS Development Team (2020). QGIS Geographic Information System (Open Source Geospatial Foundation Project).
- R Core Team (2019). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).
- Ratnasingham, S., and Hebert, P.D. (2007). BOLD: the Barcode of life data system: barcoding. *Mol. Ecol. Notes* 7, 355–364.
- Redding, D.W., Hartmann, K., Mimoto, A., Bokal, D., Devos, M., and Mooers, A.Ø. (2008). Evolutionarily distinctive species often capture more phylogenetic diversity than expected. *J. Theor. Biol.* 251, 606–615.
- Reid, W.V. (1998). Biodiversity hotspots. *Trends Ecol. Evol.* 13, 275–280.
- Ricketts, T.H., Dinerstein, E., Olson, D.M., and Loucks, C. (1999). "Who's where in North America? BioScience 49, 369–381.
- Riddle, B.R., and Hafner, D.J. (2006). A step-wise approach to integrating phylogeographic and phylogenetic biogeographic perspectives on the history of a core North American warm deserts biota. *J. Arid Environ.* 66, 435–461.
- Rohde, K. (1992). Latitudinal gradients in species diversity: the search for the primary cause. *Oikos* 65, 514.
- Rosauer, D., Laffan, S.W., Crisp, M.D., Donnellan, S.C., and Cook, L.G. (2009). Phylogenetic endemism: a new approach for identifying geographical concentrations of evolutionary history. *Mol. Ecol.* 18, 4061–4072.
- Rousset, F., and Ferdy, J.-B. (2014). Testing environmental and genetic effects in the presence of spatial autocorrelation. *Ecography* 37, 781–790.
- Sanderson, M.J. (2002). Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19, 101–109.
- Sanderson, M.J., McMahon, M.M., and Steel, M. (2010). Phylogenomics with incomplete taxon coverage: the limits to inference. *BMC Evol. Biol.* 10, 155.
- Savela, M. (2020). Lepidoptera and some other life forms. <https://www.nic.funet.fi/pub/sci/bio/life/intro.html>.
- Scoble, M.J. (1995). *The Lepidoptera: form, function, and diversity* (The Natural History Museum in association with Oxford University Press).
- Shen, J., Cong, Q., Kinch, L.N., Borek, D., Otwinowski, Z., and Grishin, N.V. (2016). Complete genome of *Pieris rapae*, a resilient alien, a cabbage pest, and a source of anti-cancer proteins. *F1000Res* 5, 2631.
- Smith, S.A., and O'Meara, B.C. (2012). treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* 28, 2689–2690.
- Sosa, V., Vásquez-Cruz, M., and Villarreal-Quintanilla, J.A. (2020). Influence of climate stability on endemism of the vascular plants of the Chihuahuan Desert. *J. Arid Environ.* 177, 104139.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313.
- Thackray, G.D. (2008). Varied climatic and topographic influences on Late Pleistocene mountain glaciation in the western United States. *J. Quat. Sci.* 23, 671–681.
- Thogmartin, W.E., Wiederholt, R., Oberhauser, K., Drum, R.G., Diffendorfer, J.E., Altizer, S., Taylor, O.R., Pleasants, J., Semmens, D., Semmens, B., et al. (2017). Monarch butterfly population decline in North America: identifying the threatening processes. *R. Soc. Open Sci.* 4, 170760.
- Thornhill, A.H., Baldwin, B.G., Freyman, W.A., Nosratinia, S., Kling, M.M., Morueta-Holme, N., Madsen, T.P., Ackerly, D.D., and Mishler, B.D. (2017). Spatial phylogenetics of the native California flora. *BMC Biol.* 15, 96.
- Waltari, E., Hijmans, R.J., Peterson, A.T., Nyári, A.S., Perkins, S.L., and Guralnick, R.P. (2007). Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *J. Chave*, ed. 2, e563.
- Webb, C.O., Ackerly, D.D., McPeek, M.A., and Donoghue, M.J. (2002). Phylogenies and community ecology. *Annu. Rev. Ecol. Syst.* 33, 475–505.
- Wepprich, T., Adrion, J.R., Ries, L., Wiedmann, J., and Haddad, N.M. (2019). Butterfly abundance declines over 20 years of systematic monitoring in Ohio, USA. *PLOS ONE* 14, e0216270.
- Wikipedia. (2020). <https://www.wikipedia.org/>.

Supplemental information

**Spatial phylogenetics of butterflies
in relation to environmental drivers
and angiosperm diversity across North America**

Chandra Earl, Michael W. Belitz, Shawn W. Laffan, Vijay Barve, Narayani Barve, Douglas E. Soltis, Julie M. Allen, Pamela S. Soltis, Brent D. Mishler, Akito Y. Kawahara, and Robert Guralnick

Figure S1

Observed PD (A & D), RPD (B & E), and PE (C & F) values for all seed plants (left panel) compared to angiosperms (right panel). Related to Figure 2.

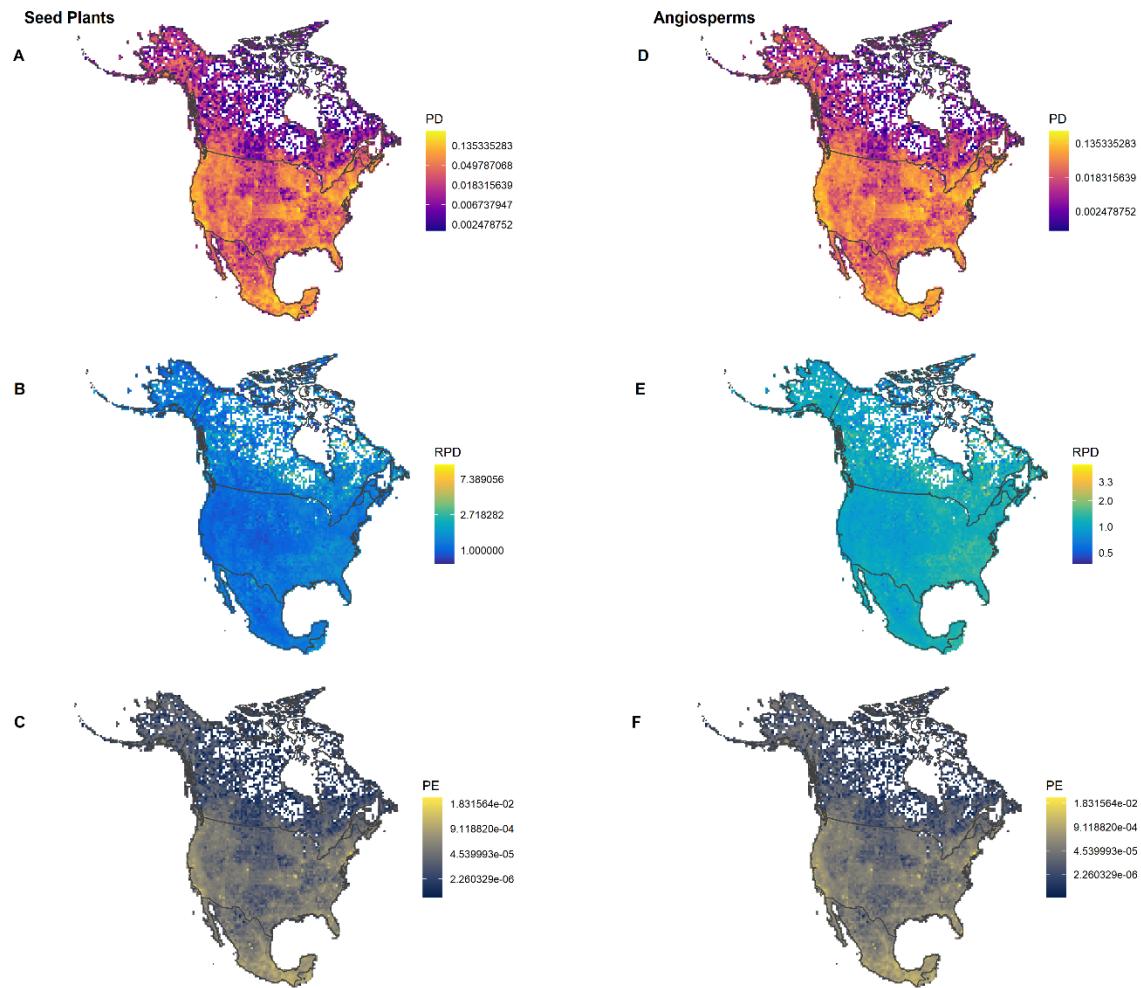


Figure S2

Observed values for North American butterflies for: (A) toxic richness, (B) phylogenetic diversity (PD), (C) relative phylogenetic diversity (RPD), and (D) phylogenetic endemism (PE). Related to Figure 2.

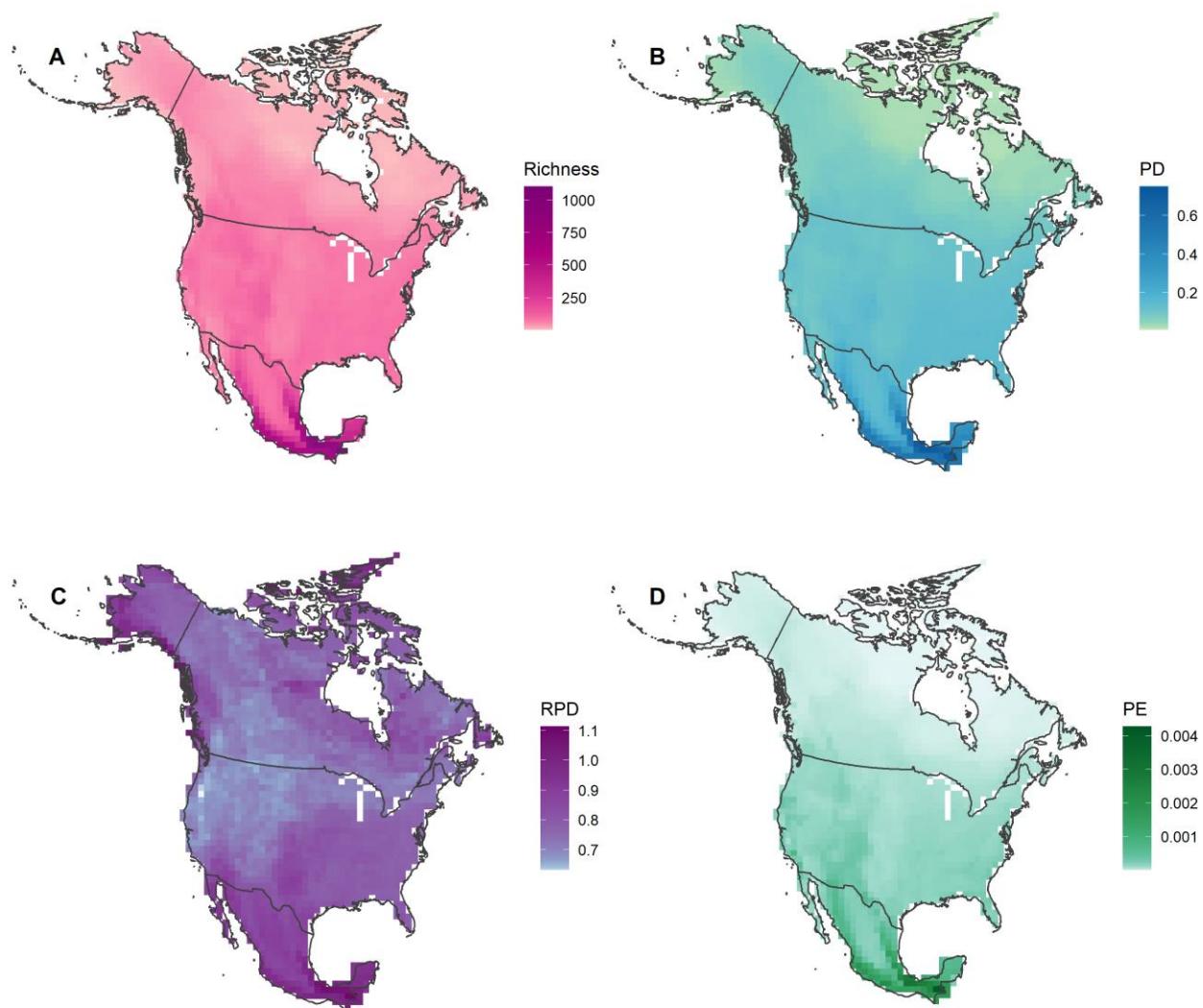


Table S1

The length of the 14 loci used as the starting sequence for input into GeneDumper. Related to Figure 1.

Locus Name	Locus Length
<i>ArgKin</i>	596
<i>CAD</i>	2,211
<i>CAT</i>	1,293
<i>COI_trnL_COII</i>	2,271 (COI: 1,531; <i>trnL</i> : 67; COII: 673)
<i>DDC</i>	957
<i>EF1a</i>	1,240
<i>GAPDH</i>	609
<i>HCL</i>	633
<i>IDH</i>	709
<i>MDH</i>	733
<i>RpS2</i>	474
<i>RpS5</i>	614
<i>Wgl</i>	453

Table S2

Summary of the top models for PD, RPD, significant PD, and significant RPD. Numbers in the columns indicate changes in PD, RPD, and significant PD and RPD values \pm standard error when variable values between locations increased by one standard deviation. Models were ranked based on the difference from the top model to the nearest competing model in Akaike's Information Criterion (Δ AIC) and Akaike weight (Wi). Models that were a subset of another model and within Δ AIC of 2 were not considered to be competitive. Additionally, models with variance inflation factors > 5 were not considered competitive. All variables contributed significantly ($p < 0.05$) to models. Related to Table 2.

Term	PD	RPD	PD Sig	RPD Sig
Temp	0.048 ± 0.001		7.82 ± 0.978	9.79 ± 1.03
Prec	0.023 ± 0.001	0.015 ± 0.002	1.25 ± 0.337	1.88 ± 0.327
Temp Seas		-0.027 ± 0.002		
Prec Seas	0.021 ± 0.002	0.024 ± 0.002		
Temp Stab	0.012 ± 0.01	0.007 ± 0.002	-1.63 ± 0.443	-1.81 ± 0.389
Prec Stab		0.01 ± 0.002		
Elev	0.008 ± 0.001	-0.024 ± 0.002	0.795 ± 0.309	1.39 ± 0.276
Δ AIC	-29.7	-7.91	-3.25	-27.7
Wi	1	0.981	0.836	1

Table S3

Ages of all families (in millions of years) in this study and others with 95% confidence intervals. Related to Figure 1.

Clade	Wahlberg et al. (2009)	Heikkilä et al. (2012)	Espeland et al. (2018) *	Condamine et al. (2018)	Chazot et al. (2019)	This study
Papilioidea	104 (93-116)	110 (92-128)	119 (91-143)	98 (66 - 189)	108 (89-129)	119.50
Papilionidae	63 (93 - 116)	75 (62-88)	84 (63-109)	86 (56-164)	68 (53-84)	79.67
Hesperiidae	N/A	65 (54-79)	79 (60-99)	76 (55-143)	65 (56-78)	72.49
Pieridae	73 (57-86)	80 (67-97)	87 (67-108)	71 (45-135)	77 (63-92)	92.29
Lycaenidae	75 (63-86)	73 (58-84)	78 (60-96)	62 (39-119)	71 (57-85)	57.40
Riodinidae	65 (56-76)	72 (57-83)	73 (56-92)	60 (36-116)	73 (60-88)	72.49
Nymphalidae	94 (84-104)	87 (74-101)	91 (71-112)	80 (52-151)	82 (68-98)	87.43

* Dates used in Espeland et al. (2018) were used to constrain the tree in the present study.

Transparent Methods

Range Map Production, Curation and Detailed Quality Control

All digitized range maps were joined into a single shapefile consisting of many spatial polygons which were clipped to only terrestrial areas within North America. We captured key attribute data for all range products including commonness and reported broad on-wing phenology. We did not digitize information about stray distributions for this work, since our key interest was range of source populations. We also therefore excluded species that stray into North America. We did not attempt in this first effort to further delineate breeding or non-breeding areas for long-distance migratory butterflies, which form only a very small percentage of species.

In order to produce highly credible range maps, we set up a rigorous review process for range map products, where maps were digitized by the authors and trained students, and a subset of the most difficult maps to digitize were reviewed by at least one member of the group to verify quality. We also instituted random spot-checks for a smaller subset of maps, especially when first starting the process. This rigor was essential because many maps included very fine-scale stippling of range locations that often required referring back to the original drawings and using a hand-lens to verify. More than a quarter of maps were carefully reviewed.

For species with ranges that extend outside of North America, we assembled country-level range estimates for those non-North American areas using a combination of different resources described fully in the main text. These initial estimates required further quality control checks and curation before use. We therefore implemented a rigorous review process for country-level estimates to assure their accuracy and to avoid issues with false positives or negatives. This curation and validation approach is as follows: 1) Country-level maps were compared to reference sources such as Wikipedia distribution descriptions and butterflies and moths of North America (Lotts and Naberhaus, 2017); 2) Maps with significant disjunctions were all checked by hand. In some cases, disjunctions represented systemic issues e.g., records in French Guiana that were reported as coming from France. In other cases, disjunctions represent either real patterns or knowledge gaps across a range. Careful checking of multiple resources (e.g., Butterflies of America (Warren *et al.*, 2016)) were used to make final validations. These country-level maps are coarse for regions outside North America, but provide a reasonable, albeit with some over-commission, approximation for range size needed for this work.

Sequence Acquisition

GenBank: GenBank was searched for 13 commonly sequenced butterfly loci across all North America species using a novel toolkit that utilizes an input taxonomy list, a list of known synonyms, and a list of loci for which to search, and returns matches. Its matching approach utilizes a probe sequence for each locus of interest. In cases where there are multiple returns for a species x locus combination, the toolkit chooses the best sequence using a set of well-defined rules, often simplifying to picking the longest sequence with the most unambiguous DNA content (See <https://github.com/sunray1/GeneDumper>). Of the 13 loci used for this work, 12 were nuclear genes: arginine kinase (*ArgKin*), ribosomal proteins S2 and S5 (*RpS2/RpS5*), carbamoyl phosphate synthetase (*CAD*), catalase (*CAT*), glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), elongation factor 1 alpha (*Ef1a*), dopa decarboxylase (*DDC*), malate dehydrogenase (*MDH*), hairy cell leukemia (*HCL*), isocitrate dehydrogenase (*IDH*), and wingless (*Wgl*). Three common mitochondrial genes were also included, but were treated as one to avoid duplicate sequences: cytochrome oxidase subunits 1 (*COI*) and 2 (*COII*) and the *trnL* intron located between them.

Once sequences were downloaded from GenBank, they were auto-cleaned using a set of criteria built into the GeneDumper toolkit. This approach separates full-length sequences and sequence fragments. Full-length sequences are first aligned, then chunks of sequence fragments are iteratively added to this alignment using the --addfragment and --adjustdirection commands of MAFFT (Katoh and

Standley, 2014). Since species x locus pairs frequently only have partial fragments of sequences that align to different parts of the whole locus, this approach decreases the number of misaligned fragments. Alignments were visualized using AliView v.1.26 (Larsson, 2014) and ultimately manually curated to remove flanking regions and to check that loci maintained the correct reading frame. GeneDumper uses a tiling approach for choosing sequences (in order to get the most amount of data across a locus), allowing some species x locus pairs to have slightly (0-20 bps) overlapping fragments generated from different sources. These fragments were merged together to create a 50% consensus sequence that contained data across the whole locus, where possible.

BOLD: Barcode of Life sequences were assembled for all species in the species list using a novel python script to query BOLD's API (<http://www.boldsystems.org/index.php/resources/api>) for locus information. Most of BOLD data is provisioned to Genbank but some genetic data remains unique to BOLD. If there were multiple sequences available for a species, all sequences were downloaded, aligned and a consensus sequence was created, choosing the most common nucleotide for each site. Only sequences that matched the 13 loci listed above were kept, and the vast majority of BOLD sequences are the barcode marker, cytochrome oxidase I (COI). Sequences were aligned using the same iterative approach described above and manually curated to remove inserts and flanking regions. We de-duplicated sequences present in both BOLD and GenBank.

Target Capture (AHE): We sequenced 224 samples using Anchored Hybrid Enrichment (AHE (Lemmon, Emme and Lemmon, 2012)) with the Butterfly1.0 target capture set of Espeland et al., (2018) (Espeland et al., 2018), and the Butterfly2.0 target capture kit of Kawahara et al. (2019) (Kawahara et al., 2018). Both sets include the 13 genes of interest and typically produced full-length, high quality sequences. When available, these sequences were prioritized over sequences found from other sources because their identification could be verified.

Novel sequencing of museum specimens: Collections in the McGuire Center for Lepidoptera & Biodiversity (MGCL), Florida Museum of Natural History, Gainesville, FL. USA, were searched for specimens for species without sequence data on public data portals. These specimens were pinned and dried, dried and papered, or stored in ethanol. One leg from each specimen was removed and transferred to a 96 well PCR microplate containing 30 uL 95% EtOH. Each well and specimen was labeled with a sample code before DNA extraction. For museum species with multiple specimens, the specimen with the most recent collection date was used to collect a tissue sample. Plates were shipped to the Canadian Centre for DNA Barcoding (CCDB) for COI barcoding (Sanger).

Digital voucher images of museum specimens were taken with their associated data whenever possible. For pinned specimens, we used a 12-megapixel rear-facing camera with quad-LED lights. For ethanol and papered specimens, we chose representative photos of species from Butterflies of America (Lotts and Naberhaus, 2017). Photographers were contacted for reuse permission and it was noted in the submission that the photos were not of the original specimen. COI barcodes were received and aligned together with all other data sources using MAFFT.

Phylogeny Reconstruction

Cleaned and aligned sequences for each locus from each source were concatenated together and realigned. Duplicate sequences were removed and those with 95% or more similarity across overlapping regions were merged into a 50% consensus sequence using a python script. Sequences that could not be merged were submitted to BLAST (Altschul et al., 1990) to determine similarity to other sequences. The sequence with the most closely related hits to the reference sequence was chosen.

Once there was one sequence per species per locus, all loci were concatenated across species into a supermatrix using FASconCAT-G v.1.02 (Kück and Longo, 2014).

Phylogenetic trees were built using maximum-likelihood phylogenetic analyses in RAxML v.8.2.10 (Stamatakis, 2014). A preliminary tree was built using an unpartitioned GTR+ Γ model of nucleotide evolution using a constraint tree based on family level relationships following Espeland et al., 2018 (Espeland et al., 2018). That analysis recovered seven monophyletic families, including Hesperiidae and Hedylidae. Long branches (greater than 99.5% longer than average branch lengths and thus significant outliers) from this preliminary tree were removed from the alignment. Singleton GenBank and BOLD sequences with species identifiers that did not group within the correct subfamily were removed under the assumption that these represented labeling or other errors during deposition. Only 17 sequences were removed based on these criteria.

The final alignment was partitioned in three different ways: by gene, by codon and by both gene and codon. PartitionFinder v.2.1 (Lanfear et al., 2016) was used to determine the best partitioning scheme and substitution model. Only models implemented within RAxML were searched and ranked under the AICc criterion. The best partitioning scheme was then used as input for model selection in RAxML to build 100 trees. The tree with the log likelihood score closest to 0 was chosen as the final tree.

We determined tree support by running 200 bootstraps via RAxML under the GTR+ Γ +I model and partitioning scheme. Bipartitions were drawn on the final maximum-likelihood tree using two methods: Felsenstein's binary bootstrap method and a gradual "transfer" distance method implemented in BOOSTER (Lemoine et al., 2018). The distance method accounts for the fact that datasets with large numbers of taxa may contain rogue taxa with unstable phylogenetic positions. Instead of removing these taxa, this method quantifies these errors as instability scores and uses these values to calculate BS scores. Phylogenies were visualized with FigTree v2.0 (Rambaut, 2010) and phytools (Revell, 2012).

Divergence times were estimated using treePL v.1.0 (Sanderson, 2002; Smith and O'Meara, 2012) using a congruification approach. We used dates from Espeland et al. (2018) as constraints on analogous nodes on the best-supported maximum likelihood tree, focusing on deeper branches of the tree. Espeland et al. (2018) describes two dated phylogenies; both phylogenies were built using nine fossil calibration points, the F48+ Γ substitution model and a lognormal, birth-death tree prior, but used differing values of the median age of Angiosperms for root calibration. We used the maximum overlap between date ranges of the six families as input into treePL. Analysis options were first optimized using the 'prime' command. A smoothing value of 0.00001 was chosen following five cross-validation analyses. All analyses determined this value to have the lowest chi-squared value.

Phylogenetic results and relationships: We retrieved 21,831 sequences recovered from GenBank across all 13 loci of interest. After sequence filtering, cleaning, and the best sequence(s) for a particular species/locus pair were chosen, 3,997 sequences across 1,019 (53%) species remained. From the BOLD sequence database, 1,696 sequences were harvested across 10 of the loci of interest, representing 1,218 (63%) species. Sequencing as part of the ButterflyNet project provided 2,735 sequences across 12 of the loci of interest were found, representing 224 (11%) species and adding 11 unique species.

Subfamily and tribe-level relationships within each family generally agree with those of Espeland et al. (2018). Papilionidae is thought to be the sister group to the remaining butterflies and has traditionally been divided into three extant subfamilies: Baroniinae, Parnassiinae, and Papilioninae (Condamine et al., 2018; Allio et al., 2020). The sole member of monospecific family Baroniinae, Baronia

brevicornis, is typically considered the sister to the other two subfamilies, but has also been grouped to Parnassiinae with Papilioninae as sister (Nazari, Zakharov and Sperling, 2007). All three subfamilies are generally accepted as monophyletic, but Papilioninae has been hypothesized to be polyphyletic (Espeland *et al.*, 2018). Our results recover Baroniinae as sister to the remainder of the family (BS = 99.5), a clade containing Parnassiinae + Papilioninae, rendering all subfamilies monophyletic (BS = 99.8 and 94.3 respectively).

Family Hesperiidae is one of the largest and most diverse families of butterflies. Although the monophyly of this group is well defined, there is disagreement on the relationships at subfamily and tribal levels (Yuan *et al.*, 2015; Zhang *et al.*, 2019). Only four of the eight described subfamilies were sampled in our phylogeny. Three were recovered as monophyletic (Hesperiinae, Eudaminae and Heteropterinae), with Pyrginae as paraphyletic. Within Pyrginae, tribes Carcharodini, Achlyodini, Erynnini, and Pyrgini formed a single clade (BS = 97) sister to tribes Pyrrhopygini and Celaenorrhiniini (BS = 89), followed by the other three subfamilies (Eudaminae + (Hesperiinae + Heteropterinae)) (BS = 97, 99 and 95, respectively).

Pieridae is currently divided into four subfamilies: Pseudopontiinae, Dismorphiinae, Coliadinae and Pierinae (Braby, Vila and Pierce, 2006). Pseudopontiinae is monotypic, containing a single genus found in Africa and was thus excluded from this study. Representatives from the other three subfamilies had the following relationships: (Dismorphiinae + (Coliadinae + Pierinae)), following recent studies (Cao *et al.*, 2016; Ding and Zhang, 2017) (BS = 100, 98 and 99). Subfamily Pierinae is usually divided into two tribes: Pierini and Anthocharidini. Our analysis supports the monophyly of these two tribes (BS = 100 and 98, respectively).

Relationships within Lycaenidae are currently unresolved and are the least well supported groups in our phylogeny (median BS = 68), but still generally follow the arrangement presented by Eliot, (1974) (Eliot, 1974). Eliot classified Lycaenidae into seven subfamilies, of which we have representatives from four (Miletinae, Theclinae, Polyommatinae and Lycaeninae). Like Espeland *et al.* (2018), we find Polyommatinae were nested within Theclinae. We also find Lycaeninae nested within Theclinae. The classification presented by Corbet *et al.* (1992) (Corbet, Pendlebury and Eliot, 1992) reduced Theclinae, Polyommatinae, and Lycaeninae to an inclusive “Lycaeninae”. Our phylogeny agrees with this lumping (BS = 99). Miletinae, although represented by only one species, was placed sister to these three subfamilies (BS = 100).

Riodinidae are mostly found in the Neotropics and are divided into three subfamilies: Riodininae, Euselasiinae and Nemeobiinae (Seraphim *et al.*, 2018). Nemeobiinae are strictly Old World riodinids and were therefore excluded from this study. The two remaining subfamilies, Riodininae and Euselasiinae, were both monophyletic with 100 and 98 BS support, respectively. Tribal relationships within Riodininae were still mixed and unsupported in all studies. This may be due to rapid radiations within this subfamily (Espeland *et al.*, 2015). Although there was some paraphyly between tribes, we agree with the suggestion made by Espeland *et al.* (2018) that an Emesis-Apodemia tribal group might need to be erected (BS = 94).

There is no phylogenetic consensus within Nymphalidae, the most speciose butterfly family, although recent studies have shed some light on its structure (Wahlberg *et al.*, 2009; Chazot *et al.*, 2019). Due to the sheer number of species (>6,000) and high amounts of diversity, Nymphalidae has been split into anywhere from 9 to 12 families (Brower, 2000; Wahlberg, Weingartner and Nylin, 2003; Freitas and Brown, 2004). All subfamilies in our analyses were monophyletic. We found Libytheinae to be sister to Danainae and Ithomiinae (lumped into Danainae by some) (BS = 98), which is in turn sister to the rest of

the family (BS = 96). This is also one of the topologies suggested by Espeland et al. (2018). Placement of the remaining nymphalids was largely concordant with the relationships in Espeland et al. (2018), with the exception of the placement of Apaturinae. Espeland et al. (2018) found Apaturinae to be sister to Biblidinae whereas we found Apaturinae sister to the clade containing Cyrestinae and Nymphalinae. This may be due to the lack of representation of subfamily Pseudergolinae in this study since it is found in Asia, not North America.

Supplemental Information References

- Allio, R. et al. (2020) "Whole Genome Shotgun Phylogenomics Resolves the Pattern and Timing of Swallowtail Butterfly Evolution," *Systematic Biology*, 69(1), pp. 38–60.
- Altschul, S. F. et al. (1990) "Basic local alignment search tool," *Journal of Molecular Biology*, 215(3), pp. 403–410.
- Braby, M. F., Vila, R. and Pierce, N. E. (2006) "Molecular phylogeny and systematics of the Pieridae (Lepidoptera: Papilionoidea): Higher classification and biogeography," *Zoological Journal of the Linnean Society*, 147(2), pp. 239–275.
- Brower, A. V. Z. (2000) "Phylogenetic relationships among the Nymphalidae (Lepidoptera) inferred from partial sequences of the wingless gene," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1449), pp. 1201–1211.
- Cao, Y. et al. (2016) "Molecular phylogenetic and dating analysis of pierid butterfly species using complete mitochondrial genomes," *Genetics and Molecular Research*, 15(4).
- Chazot, N. et al. (2019) "Priors and Posteriors in Bayesian Timing of Divergence Analyses: The Age of Butterflies Revisited," *Systematic Biology*, 68(5), pp. 797–813.
- Condamine, F. L. et al. (2018) "Mitochondrial phylogenomics, the origin of swallowtail butterflies, and the impact of the number of clocks in Bayesian molecular dating," *Systematic Entomology*, 43(3), pp. 460–480.
- Corbet, A. S., Pendlebury, H. M. and Eliot, J. N. (1992) *The butterflies of the Malay Peninsula*. Kuala Lumpur: Malayan Nature Society.
- Ding, C. and Zhang, Y. (2017) "Phylogenetic relationships of Pieridae (Lepidoptera: Papilionoidea) in China based on seven gene fragments," *Entomological Science*, 20(1), pp. 15–23.
- Eliot, J. N. (1974) "The higher classification of the Lycaenidae (Lepidoptera): a tentative arrangement," *Bull Br Mus Nat Hist Entomol*, 28, pp. 371–505.
- Espeland, M. et al. (2015) "Ancient Neotropical origin and recent recolonisation: Phylogeny, biogeography and diversification of the Riodinidae (Lepidoptera: Papilionoidea)," *Molecular Phylogenetics and Evolution*, 93, pp. 296–306.
- Espeland, M. et al. (2018) "A Comprehensive and Dated Phylogenomic Analysis of Butterflies," *Current Biology*, 28(5), pp. 770-778.e5.
- Freitas, A. V. L. and Brown, K. S. (2004) "Phylogeny of the Nymphalidae (Lepidoptera)," *Systematic Biology*, 53(3), pp. 363–383.
- Katoh, K. and Standley, D. M. (2014) "MAFFT: Iterative refinement and additional methods," *Methods in Molecular Biology*, 1079, pp. 131–146.
- Kawahara, A. Y. et al. (2018) "Phylogenetics of moth-like butterflies (Papilionoidea: Hedylidae) based on a new 13-locus target capture probe set," *Molecular Phylogenetics and Evolution*, 127, pp. 600–605.
- Kück, P. and Longo, G. C. (2014) "FASconCAT-G: Extensive functions for multiple sequence alignment preparations concerning phylogenetic studies," *Frontiers in Zoology*, 11(1).
- Lanfear, R. et al. (2016) "PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses," *Molecular Biology and Evolution*, p. msw260.
- Larsson, A. (2014) "AliView: a fast and lightweight alignment viewer and editor for large datasets," *Bioinformatics*, 30(22), pp. 3276–3278.
- Lemmon, A. R., Emme, S. A. and Lemmon, E. M. (2012) "Anchored Hybrid Enrichment for Massively High-Throughput Phylogenomics," *Systematic Biology*, 61(5), pp. 727–744.
- Lemoine, F. et al. (2018) "Renewing Felsenstein's phylogenetic bootstrap in the era of big data," *Nature*, 556(7702), pp. 452–456.
- Lotts, K. and Naberhaus, T. (2017) *Butterflies and Moths of North America*.
<http://www.butterfliesandmoths.org/>.

- Nazari, V., Zakharov, E. v. and Sperling, F. A. H. (2007) "Phylogeny, historical biogeography, and taxonomic ranking of Parnassiinae (Lepidoptera, Papilionidae) based on morphology and seven genes," *Molecular Phylogenetics and Evolution*, 42(1), pp. 131–156.
- Rambaut, A. (2010) "FigTree v2.0," *Institute of Evolutionary Biology, University of Edinburgh, Edinburgh*.
- Revell, L. J. (2012) "phytools: an R package for phylogenetic comparative biology (and other things)," *Methods in Ecology and Evolution*, 3(2), pp. 217–223.
- Sanderson, M. J. (2002) "Estimating Absolute Rates of Molecular Evolution and Divergence Times: A Penalized Likelihood Approach," *Molecular Biology and Evolution*, 19(1), pp. 101–109.
- Seraphim, N. et al. (2018) "Molecular phylogeny and higher systematics of the metalmark butterflies (Lepidoptera: Riodinidae)," *Systematic Entomology*, 43(2), pp. 407–425.
- Smith, S. A. and O'Meara, B. C. (2012) "treePL: divergence time estimation using penalized likelihood for large phylogenies," *Bioinformatics*, 28(20), pp. 2689–2690.
- Stamatakis, A. (2014) "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies," *Bioinformatics*, 30(9), pp. 1312–1313.
- Wahlberg, N. et al. (2009) "Nymphalid butterflies diversify following near demise at the Cretaceous/Tertiary boundary," *Proceedings of the Royal Society B: Biological Sciences*, 276(1677), pp. 4295–4302.
- Wahlberg, N., Weingartner, E. and Nylin, S. (2003) "Towards a better understanding of the higher systematics of Nymphalidae (Lepidoptera: Papilionoidea)," *Molecular Phylogenetics and Evolution*, 28(3), pp. 473–484.
- Warren, A. D. et al. (2016) *Illustrated Lists of American Butterflies*. Available at: <http://www.butterfliesofamerica.com/> (Accessed: April 28, 2020).
- Yuan, X. et al. (2015) "Phylogenetic relationships of subfamilies in the family Hesperiidae (Lepidoptera: Hesperioidea) from China," *Scientific Reports*, 5.
- Zhang, J. et al. (2019) "Genomes reveal drastic and recurrent phenotypic divergence in firetip skipper butterflies (Hesperiidae: Pyrrhopyginae)," *Proceedings of the Royal Society B: Biological Sciences*, 286(1903)