Environmental and geographic drivers of global bat phylogenetic diversity

Abigail E. Green[1], Camilo A. Calderón-Acevedo[3], J. Angel Soto-Centeno[4], Tara A. Pelletier[2]

1. Department of Mathematics & Statistics, Radford University, Radford, VA 24142
2. Department of Biology, Radford University, Radford, VA 24142
3. Smithsonian Tropical Research Institute, Balboa, Ancón, Panamá
4. Department of Mammalogy, Division of Vertebrate Zoology, American Museum of Natural History, New York, NY 10024

Biosketch

Our team is interested in exploring biodiversity patterns in a wide-range of taxa using methods from population genetics, phylogeography, and systematics.

AEG: www.linkedin.com/in/abigail-abbi-green

TAP: https://sites.google.com/site/taraapelletier/

CCA: https://mammalbiogeo.wordpress.com/

ASC: https://www.mormoops.com/

Abstract

Aim
Understanding patterns of biodiversity is vital to conserving species. We aimed to combine open-source genetic, environmental, and geographic information to analyze patterns of bat phylogenetic diversity (PD) in continuous ecoregions across the globe. This information is important for developing bat conservation strategies, and our methodology can work with any taxa that has sufficient georeferenced genetic data available.

Location
Global.

Time period
Last Interglacial to Present Day.

Major taxa studied
Chiroptera.

Methods
After curating a global dataset containing 14,037 COI DNA sequences from 343 described species of bats, we calculated PD for two sets of continuous ecoregions at different spatial scales. To avoid the possible difficulties of using current species names, we used genetic OTUs identified by a single-locus species delimitation method to calculate PD. We then estimated a

lineage through time plot and used random forest predictive modeling to identify the environmental and geographic predictors of PD.

Results
We found that temperature variables were most closely associated with PD. In addition to temperature during the present, temperature during the last glacial maxima and temperature changes since the last interglacial were deemed important. However, when looking at different spatial scales, top variables differed slightly. Southeast Asia and South America had the highest levels of PD, along with some parts of Africa and the Himalayas. We demonstrate that, regardless of spatial scale and uneven sampling across the globe, single-locus genetic data is enough to reflect species diversity gradients.

Main conclusions
We show that publicly available, single-locus data can be used to analyze large-scale evolutionary patterns and inform effective and efficient conservation efforts. Additionally, choice of biodiversity measure and spatial scale matter when assessing species patterns. When looking at bats, temperature variables with a historical component are most important for predicting PD.

Keywords

Chiroptera, phylogenetic diversity, COI, species delimitation, ecoregions, random forest

Introduction

Considering multiple components of biodiversity is necessary for understanding the relative roles of evolutionary processes in shaping communities at different spatial scales (Stevens & Gavilanez, 2015). For example, species richness and within-species genetic diversity are not always evenly or similarly distributed across geographic space and can be influenced by different evolutionary processes (Amador, Arroyo-Torres, Barrow, 2024; Brown et al., 2020; Lino, Ferreira, Fonseca, Fischer, Ramos Pereira, 2021). Within-species genetic diversity and phylogeographic breaks follow a similar pattern to the latitudinal species diversity gradient (da Silva Fonseca, Soto-Centeno, Simmons, Ditchfield, Leite, 2024; Smith et al., 2014) while patterns of isolation-by-distance or -environment are better explained by range size (Pelletier & Carstens, 2018; Amador et al. 2025). Documenting these differing biodiversity components (i.e., species richness or genetic diversity) is also important for structuring conservation efforts that are effective and efficient (Harper et al., 2021; Hosegood et al., 2020). Increased anthropogenic land use and habitat fragmentation can negatively affect the composition of biodiverse communities (Newbold et al, 2015; López-Baucells et al., 2022). Habitat degradation reduces ecosystem stability and productivity which endangers important ecosystem services, like food production and pollination, at local and regional scales (Cardinale et al., 2012; Moreira-Hernández, Calderón-Acevedo, Muchhala, 2021). Additionally, the increased rate of the rise in global temperatures caused by human emissions of greenhouse gases and pollution is making many habitats inhospitable by humans, wildlife, and vegetation alike (Grimm et al., 2013;

Weiskopf et al., 2020). A better understanding of the key environmental and geographic factors that both foster and discourage biodiversity long-term can help improve conservation decision making and land use management. Strong conservation initiatives are now more pertinent than ever given the rapid increase in biodiversity and habitat loss that has led to the endangerment and extinction of many species in recent millennia (Ceballos, Ehrlich, Raven, 2020; Newbold et al, 2015).

Most data available in open-source repositories can lead to new ways of documenting and exploring biodiversity, especially by integrating environmental, geographic, and genetic data. For instance, efforts to create global maps of biodiversity can be useful for tracking changes over time (Perreira 2016), identifying areas of high diversity in need of conservation (Pelletier, Carstens, Tank, Sullivan, Espíndola, 2018), and understanding evolutionary processes (Guralnick & Hill, 2009; Theissinger et al., 2023). Recent studies aimed at documenting biodiversity provide information to facilitate conservation action and may help unravel the processes that shape highly speciose regions (e.g., phylogeographic breaks in birds: Smith, Seeholzer, Harvey, Cuervo, Brumfield, 2017; response to Pleistocene glaciations in bats: Carstens, Morales, Field, Pelletier, 2018; hidden diversity in mammals: Parsons, Pelletier, Wieringa, Duckett, Carstens, 2022). Conservation initiatives aimed at including more complex levels of biodiversity that incorporate evolutionary history and capture functional diversity can help maintain the stability of a wide range of taxonomic groups and minimize species loss. For example, when space for protected areas is limited, targeting phylogenetic diversity can more effectively conserve a larger proportion of the mammal tree of life (Rosauer, Pollock, Linke, Jetz, 2017; Rosauer et al., 2018).

Phylogenetic diversity takes into account branch lengths on a phylogenetic tree (Faith, 1992) or genetic distance (i.e., divergence) among species. The most common ways to measure phylogenetic diversity include Faith's index, mean pairwise distance (MPD), and mean nearest taxon distance (MNTD) (Faith, 1992; Warwick & Clarke, 1995; Webb, Ackerly, McPeek, Donoghue, 2002; Webb, 2000). Faith's index sums of all the branch lengths in a phylogenetic tree. Mean pairwise distance (MPD) calculates average distance among all species in a phylogenetic tree. Mean nearest taxon distance (MNTD) calculates the average distance between a species and its closest evolutionary neighbor. Assemblages of species that contain more distantly related species have been shown to consist of and produce higher levels of biomass (Cadotte, 2013). More distantly related groups also capture more diversity in species traits, which increases future potential for adaptation (Forest et al., 2007). Phylogenetic diversity is also listed as an important consideration by the International Union for Conservation of Nature (www.pdtf.org).

Many studies have examined the effects of environmental factors on phylogenetic diversity across multiple taxonomic and geographic scales (Grimshaw & Higgins, 2017; de Carvalho, Martins, Esbérard, Palmeirim, 2019; Paz et al., 2021; Montaño-Centellas, Baiser, McGrew, Trotta, Li, 2023; Hu et al., 2021). The importance of different environmental factors changes with the measure of phylogenetic diversity used (Grimshaw & Higgins, 2017; Paz et al., 2021; Montaño-Centellas et al., 2023). For example, Paz et al. (2021) found that, in plants, MNTD and

MPD were more closely associated with temperature while Faith's index was more closely associated with precipitation. Nonetheless, there is limited information about the effects of environmental factors on biodiversity globally because most studies are conducted at regional geographic scales (Grimshaw & Higgins, 2017; Paz et al., 2021; de Carvalho et al., 2019; Hu et al., 2021).

Bats consist of 1482 species, which constitute about 22% of crown mammal diversity (Mammal Diversity Database, 2024; Simmons & Cirranello, 2024). Given their unique ability to fly, most bats are able to disperse across large areas providing important ecosystem services like insect pest control, pollination, and seed dispersal (Kunz, Braun de Torrez, Bauer, Lobova, Fleming, 2011; Moréira-Hernández et al., 2021; Ramírez-Fráncel et al., 2022). Over one-third of bat species assessed by the IUCN are considered threatened or data deficient, and 80% have unknown or decreasing population trends (Frick, Kingston, Flanders, 2020). Human derived activities at local, regional, and global scales are the main source of bat species decline and population fragmentation (Frick et al., 2020; Soto-Centeno & Calderón-Acevedo, 2023); however, we have limited knowledge on how phylogenetic diversity relates to habitat and environmental change across space and time in most taxonomic groups. We examined the impact of environmental and geographic factors on bat phylogenetic diversity on a global scale. First, we estimated phylogenetic diversity using georeferenced DNA sequences and species delimitation methods. Then, we used a random forest predictive algorithm to determine which geographic, environmental, and historical variables had the largest impact on phylogenetic diversity measures in continuous ecoregions across the globe.

Methods

DNA sequence data and OTUs

We obtained data from public databases, such as Genbank and the Global Biodiversity Information Facility (GBIF), through the *phylogatR* platform (Pelletier et al. 2022; https://phylogatr.org/) using the search term 'Chiroptera' on February 22, 2022. The total dataset represented 17 out of 21 (80%) families, 429 out of 1482 (29%) species, and 18 loci with a total of 20,020 DNA sequences. *PhylogatR* provides DNA sequence alignments on a species-by-species basis that are almost analysis ready. In this study, we focused on the mitochondrial gene Cytochrome oxidase I (COI) as this locus was the best represented in the data. We first checked for potentially misidentified sequences. Each species sequence alignment was run through BLAST via the NCBI database. The top hit for each DNA accession was compared to the species name associated with the DNA sequence in *phylogatR*. Sequences that were considered not plausible data (i.e., too short; 530 sequences) and two that did not BLAST to bats (i.e. *Mus musculus* and *Apodemus draco*) were removed from the alignments. Cleaned alignments from this step were merged for all species and re-aligned using MAFFT v7.5 (Katoh & Standley, 2013) with the default settings and –adjustdirection. We visually inspected the concatenated COI sequence alignment and trimmed the ends of the sequences to remove large areas with missing data resulting in a 658 base-pair region of this barcoding gene. Our final dataset consisted of 14,037 COI DNA sequences found globally (Fig. 1; Table S1). Of these, 1215 (8%) had species names in GenBank that did not match records existing in the

GBIF or the Barcode of Life Data System (BOLD). Because we used OTUs as units of analysis, we chose to retain these taxa without adjusting taxonomy (details below).

Mammal taxonomy is constantly changing (Burgin, Colella, Kahn, Upham, 2018; Parsons et al., 2022; Reeder, Helgen, Wilson, 2007; Solari & Martínez-Arias, 2014; Simmons & Cirranello, 2024; Mammal Diversity Database, 2024) and species names are often incorrect in online databases (Mulcahy et al., 2022). Therefore, we used OTUs estimated from the DNA sequence data itself, rather than species names, as terminal taxa in our phylogeny. By using OTUs, we captured biodiversity above and below currently names species diversity, which can have useful conservation implications (Rosauer et al., 2018). This is crucial, particularly for groups with complicated taxonomy, such as bats (Calderón-Acevedo, Rodríguez-Posada, Muchhala, 2021; Calderón-Acevedo, Bagley, Muchhala, 2022; da Silva Fonseca, 2024; Mônico & Soto-Centeno, 2024; Rodríguez-Posada, Morales-Martínez, Ramírez-Chaves, Martínez-Medina, Calderón-Acevedo, 2021; Yi et al., 2024) and small mammals in general (Parsons et al. 2022). To estimate genetic lineages and define OTUs, we used the distance-based species delimitation method Assemble Species by Automatic Partitioning (ASAP; Puillandre, Brouillet, Achaz, 2020), as tree-based methods (e.g., GMYC [Pons et al., 2006] or PTP [Zhang, Kapli, Pavlidis, Stamatakis, 2013]) would be prohibitive given the size of our dataset. All default settings in ASAP were used. This method establishes a set of ten different partitions ranked by an ASAP score. Each partition is a list of genetically distinct groups defining which individuals belong to each group. The ASAP score is a combination of the probability that each ASAP group is a single species along with the width of the barcode gap between the current and previous partition. Generally, the smaller the ASAP score, the better the partition is (Puillandre et al., 2020), so we used the partition with the lowest score (i.e., best ranked).

Phylogeny reconstruction and estimates of divergence

One DNA sequence was randomly sampled from each ASAP genetic group (i.e., OTU) to estimate a phylogenetic tree representing each COI lineage present in our dataset. This phylogeny (Supporting files S1 & S2) was then used in subsequent estimates of divergence times, lineage through time plots, and phylogenetic diversity (see below). We first used ModelFinder (Kalyaanamoorthy, Minh, Wong, von Haeseler, Jermiin, 2017) to select the best model of nucleotide evolution. Then, we implemented the resulting model (GTR+F+R8) in RAxML-NG (Kozlov, Darriba, Flouri, Morel, Stamatakis, 2019). This analysis ran for an initial search of 10 Maximum Likelihood (ML) trees followed by a 1000 bootstrap tree search (see commands and other related Supporting Files in Supporting Information). In this analysis, we used Yinpterochiroptera as an outgroup to all other bats following well-supported previously published bat relationships that combined mitochondrial and nuclear sequences (Shi & Raboski, 2015; Teeling et al., 2005). Given the limited information in a single-locus dataset, we wanted to ensure that we followed major accepted bat relationships (Teeling et al., 2005; Baker, Solari, Cirranello, Simmons, 2016; Cirranello, Simmons, Solari, Baker, 2016; Rojas, Warsi, Dávalos, 2016). A configuration file (Supporting file S3) for the RaxML-NG analysis included constrained relationships at the level of superfamily and family, but allowing the inference of subfamily, genera, and species/OTU (all intermediate tree construction outputs can be found in supplemental files S1-S12).

We used the rooted constrained RaxML-NG phylogeny (Supporting files S2 and S9) to compute divergence time estimates with the penalized likelihood semi-parametric method available in TreePL (Smith & O'Meara, 2012). This method allows for the estimation of different rates on different branches on a previously estimated phylogeny with branch lengths. To estimate divergences, we used the node ages of Yingochiroptera and Yangochiroptera (71–58 my), Noctilionoidea (37–47 my), the node of Pteropodidae and Rhinolophoidea (53–63 my), the node between *Pteronotus* and *Mormoops* (36.2–30.8 my), the internal node of Stenodermatinae and Phyllostominae within Phyllostomidae (36.2–27.33), and the first node within the subfamily Desmodontinae (23.9–16.7 my) as calibration points previously available from the literature (see Teeling et al. 2005; Shi & Raboski, 2015; Rojas et al., 2016). The TreePL analysis then followed an initial priming run where optimization parameters were obtained, followed by a cross-validation step to select the best smoothing parameter for the penalized likelihood estimates (Supporting file S10, Smith & O'Meara, 2012). The resulting divergence time estimates (Supporting file S11) were used to construct a lineage through time (LTT) plot with the R packages 'ape' and 'phytools' (Paradis & Schliep, 2019; Revell, 2024). The LTT plot was produced using the family Phyllostomidae as a reference because this clade of bats shows a remarkable level of species diversification in the Neotropics.

Phylogenetic diversity and geographic units, ecoregions

We calculated three measures of phylogenetic diversity (PD) using the R package 'picante' (Kembel et al., 2010): Faith's Index (Faith, 1992), MPD (Warwick & Clarke, 1995; Webb et al., 2002), and MNTD (Webb, 2000). For each ecoregion (see below), all three of these measures and OTU richness were calculated. PD estimates were used as the response variable in the predictive modeling framework described below. To estimate levels of PD across the globe, we chose to analyze data by ecoregion rather than using grid cells that may be arbitrary from a biological perspective (Paz-Vinas et al., 2021). We used the 846 RESOLVE Ecoregions 2017 dataset (Olson et al., 2001; Dinerstein et al., 2017). There were 161 ecoregions containing DNA sequences in our dataset. However, many of those had very few species, limiting our ability to make comparisons of OTU level diversity. For example, the 161 ecoregions ranged from 2 – 85 OTUs per ecoregion and had an average of 8.96 OTUs per region. There were 109 out of 161 ecoregions that had more than two OTUs represented and could be used to estimate PD. We also created a second set of larger continuous ecoregions by grouping RESOLVE Ecoregions that contained DNA sequences from our dataset, were adjacent to each other, and belonged to the same biome. This dataset contained 25 continuous ecoregions, from 185 RESOLVE Ecoregions, as some regions contained small fragmenting ecoregions that were included to create a continuous space (Table S2). There were 2 – 116 OTUs per ecoregion with an average of 21.48 OTUs. We used both datasets, a 109-eco set that used the smaller single ecoregions and a 25-eco set with combined ecoregions, to conduct subsequent analyses.

We also conducted linear regressions to test whether species richness explained the different measures of PD to determine whether PD captured a different feature of biodiversity other than simply species richness, especially given that our bat dataset is not complete (i.e., open-source databases rarely contain all species). Species richness explained almost all of the variation in

Faith's index (109-eco: $R^2 = 0.9767$, $p < 0.001$; 25-eco: $R^2 = 0.9815$, $p < 0.001$), so we moved forward with predictive modeling on PD measurements that were not correlated with species richness: MNTD (109-eco: $R^2 = 0.04689$, $p = 0.02372$; 25-eco: $R^2 = 0.03237$, $p = 0.1925$) and MPD (109-eco: $R^2 = 0.1139$, $p < 0.001$; 25-eco: $R^2 = 0.1218$, $p = 0.0488$).

Predictor variables

We curated a set of predictor variables using 96 environmental data layers (Table S3) and geographic information associated with each ecoregion. We took the mean and standard deviation of each data layer in each region by randomly sampling 1000 points per region. The data layers included 19 BIOCLIM layers for present time from the Chelsa database (Karger et al., 2017; Karger et al., 2018; Karger, Nobis, Normand, Graham, Zimmermann, 2021) at 1km resolution, elevation (Tachikawa et al., 2011), population density (CIESIN, 2017), terrestrial habitat heterogeneity (Tuanmu & Jetz, 2015), gross domestic product (GIS processing World Bank DECRG, 2010), global land cover classification (European Space Agency, 2009), global river classification (Dallaire, Lehner, Sayre, Thieme, 2019), disaster risk (Peduzzi, 2019), anthropogenic biome (Ellis, Klein Goldewijk, Siebert, Lightman, Ramankutty, 2010), global lakes and wetlands database (Lehner & Döll, 2004), and various indicators of seasonal growth Karger et al., 2017; Karger et al., 2018; Karger et al., 2021). We also included BIOCLIM layers for three different historical time periods: Last Glacial Maximum (LGM, ca. 21 ka), Last Interglacial (LIG, ca. 130 ka), and the Pliocene (ca. 3.3 ma) (Karger et al., 2021; Otto-Bliesner, et al., 2006; Dolan, et al., 2015)**.** We used the R packages 'raster' (v3.6-14; Hijmans, 2023), 'tools' (R Core Team, 2022)**,** and 'sf' (Pebesma, 2018) to extract ecoregion specific information for each of these layers. We also calculated the mean, standard deviation, maximum, and minimum values for latitude and longitude of each ecoregion and used the st_area function to calculate the area of each ecoregion (Pebesma, 2018).

To capture change of climate variables over time, we calculated—in addition to the mean and standard deviation of each layer— the difference between the mean in the present day and the mean in the historical period for each of the historical layers. This gave us a total of 253 variables that represent a variety of environmental and geographic traits for each of our ecoregions (Tables S4 & S5). Fourteen of the 253 variables for the 25-eco dataset contained NA values that complicated downstream analysis and were removed from the dataset, as the random forest algorithm does not handle missing data. These fourteen variables included the mean and standard deviation for three growing season temperature variables (Karger et al., 2017; Karger et al., 2018; Karger et al., 2021), snow water equivalent (Karger et al., 2017; Karger et al., 2021), global river classification (Dallaire et al., 2019), and two global lakes and wetlands classifications (Lehner & Döll, 2004). There were six additional variables that contained NA values and were removed for the 109-eco dataset: the mean and standard deviation for the last day of the growing season, the first day of the growing season, and global drought events.

Predictive modeling

To reduce the number of variables in our predictive models, we used different sets of predictor variables based on correlation coefficients and variable categories. We first calculated pairwise

Pearson correlation coefficients for all environmental and geographic variables using the R package 'Hmisc' (Harrell Jr., 2023) to remove highly correlated variables (Tables S6 & S7). We started with three correlation cutoffs, 0.7, 0.8, and 0.9, where we removed predictor variables with a correlation coefficient greater than or equal to the cutoff. We then split the variables into six groups based on category of variable and ran a Random Forest (RF) for each group: geographic, current BIOCLIM, LGM BIOCLIM, LIG BIOCLIM, Pliocene BIOCLIM, and miscellaneous (including things like flood risk, growing season, anthropogenic biomes, etc.). We ran these category analyses using all variables and those removed at 0.9. In total, we ran 15 RF analyses with different sets of predictor variables for both MPD and MNTD as response variables with the R package 'randomForest' using the default settings, 1000 trees, and retaining the importance values (Table 1).

Based on the predictive power of the RF models (Table 1), we focus on MPD as our phylogenetic diversity response variable. Percent increase in mean squared error (MSE) is calculated for each variable when removed from the model to determine which variables contribute the most to the prediction. We used this value to estimate the relative importance of each variable. To further reduce the variables in our data and assess which were most important, we took the top 1-6 predictors from the six random forest models (each category listed above with 0.9 or higher correlated variables removed). Variables with a 10% increase in MSE or higher were used to run a final RF with only the top predictors. There were 27 top predictors for the 109-eco set and 14 top predictors for the 25-eco set across the six different categories of variables. For the top predictors based on the final RF models, we conducted linear regressions to test the overall significance of the predictor variables.

Results

Genetic data and OTUs

After data cleaning and trimming the sequence alignment, there were 658 base-pairs of COI from 343 species and 14,037 individual sequences (Fig. 1; Table S1). ASAP identified 372 genetic partitions or OTUs. With >14,000 sequences, the ASAP analysis took less than 20 minutes on iMac (3.6 GHz 8-Core Intel Core i9 16 GB 2667 MHz DDR4). The ASAP results largely matched species names but there were cases where nominal species were split, merged, and sometimes both (Fig. S1; Table S8). 155 nominal bat species matched the ASAP groups exactly, where all individual sequences from each species only belonged to one ASAP group, 81 nominal species were split into multiple ASAP groups, and 161 nominal species were combined in 69 ASAP groups. A total of 54 nominal species that were split into multiple ASAP groups also had individuals merged with other nominal species. The average value for OTU richness across the 25-eco set was 21.48 with East Asia having the highest (116) and four different regions (Mountainous region of Asia (Himalayas), Northwest Asia, Northwest North America, Philippines) having the lowest (3) (Table S4). The average value for OTU richness across the 109-eco set was 12.28 with the Guianan lowland moist forests having the highest (85), followed by the Northern Annamites rain forests (53), and 23 regions with the lowest (3) (Table S5).

Bat phylogeny reconstruction and lineages through time

The estimated phylogeny (Fig. 2 & S2; Supporting file S2) accurately inferred most of the currently accepted relationships of bats both at the levels of subfamily and genus. Estimates of lineages through time based on the reconstructed phylogeny demonstrated a pattern congruent with rapid species diversification in Phyllostomidae, the most speciose Neotropical bat family, compared to more gradual lineage accumulation for other more deeply or recently divergent families (e.g., Molossidae and Hipposideridae; Fig. 3). However, we noted that some deep nodes (i.e., the sister relationship of Emballonuroidea with and Noctilionoidea) in the phylogeny were recovered with short branches (Fig. 2). Furthermore, the divergence time estimates produced by TreePL show deeper (i.e., older) divergences than expected based on previously published studies (e.g., Teeling et al., 2005; Rojas et al., 2016). These observations were likely due to our use of a single mitochondrial marker with a fast mutation rate compared to other studies that use a combination of nuclear and mitochondrial data. Despite that, our choice of using a single mitochondrial marker allowed us to include a wide sampling breadth across 70% of all bat families worldwide, which facilitated the analysis of drivers of diversity globally (read below).

Phylogenetic diversity and predictive modeling

Average values for the 25-eco dataset for MNTD and MPD across the ecoregions were 0.38 and 0.62, respectively. There were no obvious global geographic patterns in PD as measured by MPD, except that the southern hemisphere contained higher levels of MPD (Fig. 1). Middle Africa and Southeast Asia had the highest levels of MPD, followed by regions in Central America, Northern South America, and Southern Africa. Western North America had the lowest levels of MPD. Average values for the 109-eco dataset for MNTD and MPD across the ecoregions were 0.42 and 0.60, respectively. Ecoregions in Asia had the highest MPD values, followed by some areas in Africa (Fig. S4). Similarly, the southern, more tropical regions of the globe had higher levels of MPD.

MNTD as the response variable had low predictive power. For the 25-eco set, at most, 16.48% of the variation was explained by Pliocene variables. For the 109-eco set, at most, 19.36% of the variation was explained when using all variables and a 0.8 correlation cut-off. The models using MPD as a response variable explained anywhere from 42.3-68.23% [25-eco] or 34.27-56.86% [109-eco] of the variation depending on the model; thus, we focus our attention on the results using MPD as our measure of PD (i.e., our response variable for predictive modeling) (Table 1). While the percent variation explained differed among all our MPD RF models, the geographic (25-eco: 65.18%; 109-eco: 50.45%) and miscellaneous (25-eco: 65.41%; 109-eco: 51.87%) variables had the most explanatory power for both sets of ecoregions, though all had moderate predictive power.

For each set of ecoregions, we conducted an RF analysis with the top 1-6 predictors chosen based on a cutoff of 10% increase in MSE or higher from each of the six previous RF categories (Fig. 4). We further examined the variables with 10% or higher increase in MSE from this overall top variable random forest analysis. For the 25-eco set, these variables included mean temperature of the wettest quarter (Bio8) in present day, mean temperature of the wettest

quarter (Bio8) during the last glacial maximum, and mean difference in diurnal temperature range (Bio2: mean of monthly (max temp. – min temp.)) between the last interglacial and present day. This top variable RF model had high predictive power, explaining 78.28% of the variation in MPD. For the 109-eco set, there were 6 variables with a 10% or higher increase in MSE. We chose to run additional linear regressions on the top two variables, which had ~4% higher increase in MSE compared to the other 4 variables, similar to the patterns of the top three variables in 25-eco set (i.e., a large gap in MSE). These variables were mean latitude and mean temperature of the wettest quarter (Bio8) in present day. This model explained 57.23% of the variation in MPD. This suggests that our iterative variable reduction technique increased predictive power allowing us to hone in on the most important variables.

For the 25-eco set, all three variables were significant after implementing a Bonferroni correction (0.05/3=0.016) (Fig. 5). Mean temperature of the wettest quarter had an R^2 value of 0.886 and a p-value <0.001, mean temperature of the wettest quarter during the last glacial maximum had an R^2 value of 0.7712 and a p-value <0.001, and mean difference in diurnal temperature range between the last interglacial and present day had an R^2 value of 0.482 and a p-value <0.001. For the 109-eco set, both variables were significant after implementing a Bonferroni correction (0.05/2=0.025) (Fig. 5). Mean latitude had an R^2 value of 0.1634 and a p-value <0.001 and mean temperature of the wettest quarter had an R^2 value of 0.454 and a p-value <0.001.

Discussion

Terrestrial ecosystems have been rapidly altered by anthropogenic pressures (Ellis et al., 2021). The degradation and fragmentation of habitats and overexploitation of species globally has culminated in the current biodiversity crisis. Various conservation initiatives have proposed targets of 30 to 50% protection of the terrestrial biosphere to mitigate species loss (Dinerstein et al., 2019; Watson & Venter, 2017). Deciphering the environmental and spatial factors that shape biodiversity is imperative to resolve regional and global patterns of diversity and promote conservation. We leveraged open-source, single-locus genetic data and spatial layers to examine drivers of bat phylogenetic diversity at a global scale using random forest predictive modeling.

We produced a global bat phylogeny and defined OTUs using the Assemble Species by Automatic Partitioning (ASAP) method. While ASAP genetic groups generally matched current bat taxonomy, we encountered various groups in the dataset where merging and/or splitting of nominal species was common. These groups fit into different categories. For example, some merged groups included OTUs within genera with cryptic species that have undergone recent taxonomic revisions (e.g., Jarrín-V & Kunz, 2008; Parlos, Timm, Swier, Zeballos, Baker, 2014; Velazco, Soto-Centeno, Fleck, Voss, Simmons, 2017; Calderón-Acevedo & Muchhala, 2018; Calderón-Acevedo & Muchhala, 2020; Calahorra-Oliart, Ospina-Garcés, León-Paniagua, 2021; Calderón-Acevedo et al., 2022; Calahorra-Oliart et al., 2021; da Silva Fonseca et al. 2024; Mônico & Soto-Centeno 2024), particularly small Stenodermatine frugivores (Rodríguez-Posada, Morales-Martínez, Gutiérrez-Sanabria, Calderón-Capote, Fernández-Rodríguez, 2018; Morales-Martínez, Rodríguez-Posada, Ramírez-Chaves, 2021). Other mismatched groups

included widely distributed insectivores that occur from Africa to Southeast Asia (Foley et al., 2015; Foley, Goodman, Whelan, Puechmaille, Teeling, 2017; Patterson et al., 2019; Taylor et al., 2012; Csorba, Son, Saveng, Furey, 2011). These examples highlight some of the many recent efforts to provide more stable taxonomy based on species delimitation (Calahorra-Oliart et al., 2021; da Silva Fonseca et al. 2024; Mônico & Soto-Centeno 2024). Our ASAP results provide a baseline to guide future taxonomic work in multiple clades.

Southeast Asia is known for its high levels of bat species richness and our analysis revealed that it also exhibits some of the highest levels of PD. Despite having the second highest bat species richness, the Neotropics showed lower phylogenetic diversity than parts of Africa and the Himalayas that had the next highest levels of phylogenetic diversity in our analysis. This matches previously found patterns in species diversity and other measures of PD in bats (López-Aguirre, Hand, Laffan, Archer, 2018). The fact that the Himalayas had relatively higher levels of PD but lower species richness, a pattern observed previously, might be particularly important because of this region's high sensitivity to climate change (Chakravarty, Mohan, Voigt, Krishnan, Radchuk, 2021). Furthermore, in birds, the only other group of flying vertebrates, PD was higher in mountainous regions while being lower in overall species richness (Jarzyna, Quintero, Jetz, 2021; Montaño-Centellas et al., 2023). Similar to the latitudinal gradient of species diversity, we did detect relatively high PD around the tropics compared to temperate areas (Fig. 1 & S4) and latitude was identified as an important predictor in our 109-eco set (a smaller spatial scale). However, it is important to note that there are large areas of South America and parts of central Africa that did not have enough data to estimate PD, and overall sampling was lower in the global south, potentially limiting some ability to detect all PD.

At both spatial scales, the 25- and 109-ecoregion sets, the mean temperature of the wettest quarter (Bio8) was a significant contributor to PD in bats. Mean temperature of the wettest quarter during the Last Glacial Maximum and change in diurnal air temperature range since the Last Interglacial were identified as important in our analyses when looking at larger spatial scales (25-eco set). This could be key to understanding patterns of diversity given that climate oscillations during the Quaternary (i.e., Pleistocene and Holocene) have been associated with extinction or shaping genetic structure and distribution of many small mammals in temperate regions (Blois, McGuire, Hadly, 2010; Shafer, Cullingham, Côté, Coltman, 2010). Changes in climate during this time likely also played a significant role in the historical demography of bats (Carstens et al., 2018). For example, the warmer Pliocene climate was a driver of speciation in African birds (Voelker, Outlaw, Bowie, 2010); it is possible the same pattern holds true for bats, explaining the positive relationship between PD and temperature change and temperature during wet months from the past to the present day.

These results are congruent with similar studies analyzing the relationship between environmental variables and various measures of phylogenetic diversity in bats and birds. Phylogenetic diversity measures like MNTD and MPD have been associated with temperature when looking at both Mexican bat communities and global raptor diversity (Grimshaw & Higgins, 2017; Montaño-Centellas et al., 2023). There are also indications that considering historical components is important for explaining phylogenetic diversity (Montaño-Centellas et al., 2023;

Paz et al., 2021), though these variables are not included as often in studies of biodiversity. Speciation rates in most bat families have slowed down over time, as shown in our lineage through time plots and estimated elsewhere (Shi & Rabosky, 2015). This indicates that historical processes should be considered when the goal is to preserve PD. The measure of phylogenetic diversity being used also seems to influence which predictors are important. In contrast to analyses using MNTD and MPD, studies looking at Faith's index saw precipitation variables as top predictors (Grimshaw & Higgins, 2017; Paz et al., 2021). We did not consider Faith's index in our predictive modeling because it was highly correlated with species richness in our dataset, and we did not want our results to be driven by species richness.

We attempted to explore the top variables, those with a 10% increase in MSE or higher, in predicting PD at different spatial scales. While the 25-eco worldwide dataset had top variables that were primarily temperature related, similar to analyses from smaller regions like the Atlantic Forest of Brazil and Mexico (Grimshaw & Higgins, 2017; Paz et al., 2021), the 109-eco data, although it had lower predictive power, also identified latitude as important. We believe it could be worthwhile to more systematically explore if the same patterns apply within and between ecoregions and what differences arise. For example, similar to our results, López-Baucells et al. found that at small spatial scales environmental variables do not seem to explain PD in aerial insectivorous bats, but PD was higher in larger forest fragments (2022). This could influence how we think about conservation practices for species at different spatial scales in conjunction with evolutionary history. Furthermore, protected areas do not always maintain within-species genetic diversity (Thompson, Pelletier, Carstens, 2021) or fully capture phylogenetic diversity (Saraiva, dos Santos, Overbeck, Giehl, Jarenkow, 2018). Future studies at local scales should examine the usefulness of protected areas for preserving phylogenetic diversity in bats (e.g., Leon-Alvarado & Miranda-Esquivel, 2023).

Conclusions

When it comes to biodiversity assessments needed to justify and inform conservation efforts, we have shown that open-source, single-locus data can be a valuable tool. Global bat phylogenetic diversity is highest around the tropics of Asia, Central America, South America, and Africa which aligns well with patterns of species richness, but assessments at different spatial scales will provide slightly different results. Phylogenetic diversity in bats is best explained by temperature variables while also considering a historical component. This speaks to the idea of protecting large swaths of geographic space to efficiently protect phylogenetic diversity moving forward, particularly if the potential for future speciation events is deemed important.

Acknowledgements

Funding provided by National Science Foundation DBI-1911293 to TAP. We thank Radford University's Office of Undergraduate Research for supporting AEG through the Highlander Research Rookies program.

References

Amador, L., Arroyo-Torres, I., & Barrow, L. N. (2024). Machine learning and phylogenetic models identify predictors of genetic variation in Neotropical amphibians. *Journal of Biogeography, 51*(5), 909-923. https://doi.org/10.1111/jbi.14795

Amador, L., Wiley, D. L., Arroyo-Torres, I., McDaniels, C. X., Rosario-Sanchez, E. O., Farmer, H., ... Barrow, L. N. (2025). Caudata macrogenetics: Species range size predicts intraspecific genetic variation in global salamanders. *bioRxiv*, 2025-01.

Baker, R. J., Solari, S., Cirranello, A., & Simmons, N. B. (2016). Higher level classification of phyllostomid bats with a summary of DNA synapomorphies. *Acta Chiropterologica, 18*(1), 1–38. https://doi.org/10.3161/15081109acc2016.18.1.001

Blois, J. L., McGuire, J. L., & Hadly, E. A. (2010). Small mammal diversity loss in response to late-Pleistocene climatic change. *Nature, 465*(7299), 771-774.

Brown, J. L., Paz, A., Reginato, M., Renata, C. A., Assis, C., Lyra, M., ... Carnaval, A. C. (2020). Seeing the forest through many trees: Multi-taxon patterns of phylogenetic diversity in the Atlantic Forest hotspot. *Diversity and Distributions, 26*(9), 1160-1176. https://doi.org/10.1111/ddi.13116

Burgin, C. J., Colella, J. P., Kahn, P. L., & Upham, N. S. (2018). How many species of mammals are there? *Journal of Mammalogy, 99*(1), 1–14. https://doi.org/10.1093/jmammal/gyx147

Cadotte, M. W. (2013). Experimental evidence that evolutionarily diverse assemblages result in higher productivity. *Proceedings of the National Academy of Sciences, 110*(22), 8996-9000. https://doi.org/10.1073/pnas.1301685110

Calahorra-Oliart, A., Ospina-Garcés, S.M. & León-Paniagua, L. (2021) Cryptic species in *Glossophaga soricina* (Chiroptera: Phyllostomidae): do morphological data support molecular evidence? *Journal of Mammalogy*, 102, 54–68.

Calderón-Acevedo, C. A., Bagley, J. C., & Muchhala, N. (2022). Genome-wide ultraconserved elements resolve phylogenetic relationships and biogeographic history among Neotropical leaf-nosed bats in the genus Anoura (Phyllostomidae). *Molecular Phylogenetics and Evolution*, 167, e107356. https://doi.org/10.1016/j.ympev.2021.107356

Calderón-Acevedo, C.A. & Muchhala, N. (2020) First report of the Broad-toothed Tailless Bat, *Anoura latidens* Handley, 1984 (Chiroptera, Phyllostomidae), in Bolivia. *Check List*, 16, 1545–1550.

Calderón-Acevedo, C.A. & Muchhala, N.C. (2018) Identification and diagnosis of *Anoura fistulata* with remarks on its presumed presence in Bolivia. *Journal of Mammalogy*, 99, 131–137.

Calderón-Acevedo, C. A., Rodríguez-Posada, M. E., & Muchhala, N. (2021). Morphology and genetics concur that *Anoura carishina* is a synonym of *Anoura latidens* (Chiroptera, Glossophaginae). *Mammalia*, 85, 471–481. https://doi.org/10.1515/mammalia-2020-0183

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., ... Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature, 486*(7401), 59-67. https://doi.org/10.1038/nature11148

Carstens, B. C., Morales, A. E., Field, K., & Pelletier, T. A. (2018). A global analysis of bats using automated comparative phylogeography uncovers a surprising impact of

Pleistocene glaciation. *Journal of Biogeography, 45*(8), 1795-1805. https://doi.org/10.1111/jbi.13382

Ceballos, G., Ehrlich, P. R., & Raven, P. H. (2020). Vertebrates on the brink as indicators of biological annihilation and the sixth mass extinction. *Proceedings of the National Academy of Sciences, 117*(24), 13596-13602. https://doi.org/10.1073/pnas.1922686117

Center for International Earth Science Information Network - CIESIN - Columbia University. (2017). Gridded population of the world, version 4 (GPWv4): Population density, revision 11. *NASA Socioeconomic Data and Applications Center (SEDAC)*. https://doi.org/10.7927/H49C6VHW

Chakravarty, R., Mohan, R., Voigt, C. C., Krishnan, A., & Radchuk, V. (2021). Functional diversity of Himalayan bat communities declines at high elevation without the loss of phylogenetic diversity. *Scientific Reports, 11*(1), 22556.

Cirranello, A., Simmons, N. B., Solari, S., & Baker, R. J. (2016). Morphological diagnoses of higher-level phyllostomid taxa (Chiroptera: Phyllostomidae). *Acta Chiropterologica, 18*(1), 39–71. https://doi.org/10.3161/15081109acc2016.18.1.002

Csorba, G., Son, N. T., Saveng, I., & Furey, N. M. (2011). Revealing cryptic bat diversity: three new *Murina* and redescription of *M. tubinaris* from Southeast Asia. *Journal of Mammalogy, 92*(4), 891-904.

da Silva Fonseca, B., Soto-Centeno, J. A., Simmons, N. B., Ditchfield, A. D., & Leite, Y. L. R. (2024). A species complex in the iconic frog eating bat *Trachops cirrhosus* (Chiroptera, Phyllostomidae) with high variation in the heart of the Neotropics. *American Museum Novitates*, *4021*, 1–27.

Dallaire, C. O., Lehner, B., Sayre, R., & Thieme, M. (2019). A multidisciplinary framework to derive global river reach classifications at high spatial resolution. *Environmental Research Letters, 14*(2), 024003.

de Carvalho, W. D., Martins, M. A., Esbérard, C. E. L., & Palmeirim, J. M. (2019). Traits that allow bats of tropical lowland origin to conquer mountains: Bat assemblages along elevational gradients in the South American Atlantic Forest. *Journal of Biogeography, 46*(2), 316–331. https://doi.org/10.1111/jbi.13506

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., ... Noss, R. (2017). An ecoregion-based approach to protecting half the terrestrial realm. *BioScience*, *67*(6), 534–545. https://doi.org/10.1093/biosci/bix014

Dinerstein, E., Vynne, C., Sala, E., Joshi, A. R., Fernando, S., Lovejoy, T. E., ... Wikramanayake, E. (2019). A global deal for nature: guiding principles, milestones, and targets. *Science advances, 5*(4), eaaw2869.

Dolan, A. M., Haywood, A. M., Hunter, S. J., Tindall, J. C., Dowsett, H. J., Hill, D. J., & Pickering, S. J. (2015). Modelling the enigmatic late Pliocene glacial event—Marine Isotope Stage M2. *Global and Planetary Change, 128*, 47-60. https://doi.org/10.1016/j.gloplacha.2015.02.001

Ellis, E. C., Klein Goldewijk, K., Siebert, S., Lightman, D., & Ramankutty, N. (2010). Anthropogenic transformation of the biomes, 1700 to 2000. *Global ecology and biogeography, 19*(5), 589-606. https://doi.org/10.1111/j.1466-8238.2010.00540.x

Ellis, E.C., Gauthier, N., Klein Goldewijk, K., Bliege Bird, R., Boivin, N., Díaz, S., ... Watson, J.E.M. (2021). People have shaped most of terrestrial nature for at least 12,000 years. *Proceedings of the National Academy of Sciences, 118*, e2023483118.

European Space Agency (ESA). 2009. ESA GlobCover project.

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological conservation, 61*(1), 1-10. https://doi.org/10.1016/0006-3207(92)91201-3

Foley, N.M., Goodman, S.M., Whelan, C. V., Puechmaille, S.J. & Teeling, E. (2017). Towards navigating the minotaur's labyrinth: Cryptic diversity and taxonomic revision within the speciose genus hipposideros (Hipposideridae). *Acta Chiropterologica*, 19, 1–18.

Foley, N.M., Thong, V.D., Soisook, P., Goodman, S.M., Armstrong, K.N., Jacobs, D.S., ... Teeling, E.C. (2015). How and why overcome the impediments to resolution: Lessons from rhinolophid and hipposiderid bats. *Molecular Biology and Evolution*, 32, 313–333.

Forest, F., Grenyer, R., Rouget, M., Davies, T. J., Cowling, R. M., Faith, D. P., ... Savolainen, V. (2007). Preserving the evolutionary potential of floras in biodiversity hotspots. *Nature, 445*, 757-760. https://doi.org/10.1038/nature05587

Frick, W. F., Kingston, T., & Flanders, J. (2020). A review of the major threats and challenges to global bat conservation. *Annals of the New York Academy of Sciences, 1469*(1), 5–25. https://doi.org/10.1111/nyas.14045

GIS processing World Bank DECRG. (2010). Washington, DC, extrapolation UNEP/GRID-Geneva. https://doi.org/10.57966/edbt-r040

Grimm, N. B., Chapin III, F. S., Bierwagen, B., Gonzalez, P., Groffman, P. M., Luo, Y., ... Williamson, C. E. (2013). The impacts of climate change on ecosystem structure and function. *Frontiers in Ecology and the Environment, 11*(9), 474-482. https://doi.org/10.1890/120282

Grimshaw, J. R., & Higgins, C. L. (2017). Environmental correlates of phylogenetic structure in Mexican bat communities. *Journal of Mammalogy, 98*(6), 1657-1666. https://doi.org/10.1093/jmammal/gyx101

Guralnick, R., & Hill, A. (2009). Biodiversity informatics: Automated approaches for documenting global biodiversity patterns and processes. *Bioinformatics, 25*(4), 421-428. https://doi.org/10.1093/bioinformatics/btn659

Harper, M., Mejbel, H. S., Longert, D., Abell, R., Beard, T. D., Bennett, J. R., ... Cooke, S. J. (2021). Twenty-five essential research questions to inform the protection and restoration of freshwater biodiversity. *Aquatic Conservation: Marine and Freshwater Ecosystems, 31*(9), 2632-2653. https://doi.org/10.1002/aqc.3634

Harrell Jr., F. (2023). Hmisc: Harrell Miscellaneous. R package version 5.0-1, https://CRAN.R-project.org/package=Hmisc.

Hijmans R (2023). _raster: Geographic Data Analysis and Modeling_. R package version 3.6-14, <https://CRAN.R-project.org/package=raster>.

Hosegood, J., Humble, E., Ogden, R., De Bruyn, M., Creer, S., Stevens, G. M., ... Carvalho, G. (2020). Phylogenomics and species delimitation for effective conservation of manta and devil rays. *Molecular ecology, 29*(24), 4783-4796. https://doi.org/10.1111/mec.15683

Hu, Y., Fan, H., Chen, Y., Chang, J., Zhan, X., Wu, H., ... Wei, F. (2021). Spatial patterns and conservation of genetic and phylogenetic diversity of wildlife in China. *Science Advances, 7*(4), eabd5725. https://doi.org10.1126/sciadv.abd5725

Jarrín-V., P. & Kunz, T.H. (2008). Taxonomic history of the genus Anoura (Chiroptera: Phyllostomidae) with insights into the challenges of morphological species delimitation. *Acta Chiropterologica, 10*, 257–269.

Jarzyna, M. A., Quintero, I., & Jetz, W. (2021). Global functional and phylogenetic structure of avian assemblages across elevation and latitude. *Ecology Letters, 24*(2), 196-207. https://doi.org/10.1111/ele.13631

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589. https://doi.org/10.1038/nmeth.4285

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., ... Kessler, M. (2017). Climatologies at high resolution for the Earth land surface areas. *Scientific Data, 4.* 170122. https://doi.org/10.1038/sdata.2017.122

Karger D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., ... Kessler, M. (2018). Data from: Climatologies at high resolution for the earth's land surface areas. *Dryad Digital Repository*. http://dx.doi.org/doi:10.5061/dryad.kd1d4

Karger, D. N., Nobis, M. P., Normand, S., Graham, C. H., & Zimmermann, N. E. (2021). CHELSA-TraCE21k v1. 0. Downscaled transient temperature and precipitation data since the last glacial maximum. *Climate of the Past Discussions, 1-27*.

Katoh, K. & Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution, 30*, 772–780.

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., ... Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics, 26*(11), 1463-1464. https://doi.org/10.1093/bioinformatics/btq166

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35(21)**, 4453–4455. https://doi.org/10.1093/bioinformatics/btz305

Kunz, T.H., Braun de Torrez, E., Bauer, D., Lobova, T. & Fleming, T.H. (2011). Ecosystem services provided by bats. *Annals of the New York Academy of Sciences, 1223*, 1–38.

Lehner, B., & Döll, P. (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of hydrology, 296*(1-4), 1-22. https://doi.org/10.1016/j.jhydrol.2004.03.028

Leon-Alvarado, O.D. & Miranda-Esquivel, D.R. (2023) Phylogenetic diversity and North Andean block conservation. *PeerJ, 11*, e16565.

Lino, A., Ferreira, E., Fonseca, C., Fischer, E. & Ramos Pereira, M.J. (2021). Species–genetic diversity correlation in phyllostomid bats of the Bodoquena plateau, Brazil. *Biodiversity and Conservation*, 30, 403–429.

López-Aguirre, C., Hand, S. J., Laffan, S. W., & Archer, M. (2018). Phylogenetic diversity, types of endemism and the evolutionary history of New World bats. *Ecography, 41*(12), 1955-1966. https://doi.org/10.1111/ecog.03260

López-Baucells, A., Rowley, S., Rocha, R., Bobrowiec, P. E., Palmeirim, J. M., Farneda, F. Z., & Meyer, C. F. (2022). Interplay between local and landscape-scale effects on the taxonomic, functional and phylogenetic diversity of aerial insectivorous neotropical bats. *Landscape Ecology, 37*(11), 2861-2875. https://doi.org/10.1007/s10980-022-01493-x

Mammal Diversity Database. (2024). Mammal Diversity Database (Version 1.11) [Data set]. Zenodo. https://doi.org/https://doi.org/10.5281/zenodo.7830771

Mônico, P. I. & Soto-Centeno, J. A. (2024). Phylogenetic, morphological and niche differentiation unveil new species limits for the big brown bat (*Eptesicus fuscus*). *Royal Society Open Science*, *11* (2), 231384.

Montaño-Centellas, F., Baiser, B., McGrew, A., Trotta, L., & Li, D. (2023). Global patterns and drivers of raptor phylogenetic and functional diversity. Global Ecology and Biogeography, 32(2), 281-294. https://doi.org/10.1111/geb.13619

Morales-Martínez, D.M., Rodríguez-Posada, M.E. & Ramírez-Chaves, H.E. (2021). A new cryptic species of yellow-eared bat *Vampyressa melissa* species complex (Chiroptera: Phyllostomidae) from Colombia. *Journal of Mammalogy, 102*, 90–100.

Moreira-Hernández, J.I., Calderón-Acevedo, C.A. & Muchhala, N. (2021). Fur, wings, and flowers: Development and progress on nectarivorous bat research in the last 50 years. 50 Years of Bat Research. In B.K. Lim, M.B. Fenton, R.M. Brigham, S. Mistry, A. Kurta, E.H. Gillam, A. Russell, and J. Ortega, & (Eds.), pp. 135–149. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-54727-1_9

Mulcahy, D. G., Ibáñez, R., Jaramillo, C. A., Crawford, A. J., Ray, J. M., Gotte, S. W., ... de Queiroz, K. (2022). DNA barcoding of the National Museum of Natural History reptile tissue holdings raises concerns about the use of natural history collections and the responsibilities of scientists in the molecular age. *Plos one, 17*(3), e0264930. https://doi.org/10.1371/journal.pone.0264930

Newbold, T., Hudson, L. N., Hill, S. L., Contu, S., Lysenko, I., Senior, R. A., ... Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature, 520*(7545), 45-50. https://doi.org/10.1038/nature14324

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., ... Kassem, K. R. (2001). Terrestrial ecoregions of the world: A new map of life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, *51*(11), 933–938. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Otto-Bliesner, B. L., Marshall, S. J., Overpeck, J. T., Miller, G. H., Hu, A., & CAPE Last Interglacial Project members. (2006). Simulating arctic climate warmth and icefield retreat in the last interglaciation. *Science, 311*(5768), 1751-1753.

Paradis, E. & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35, 526-528. doi:10.1093/bioinformatics/bty633

Parlos, J.A., Timm, R.M., Swier, V.J., Zeballos, H. & Baker, R.J. (2014) Evaluation of the paraphyletic assemblages within Lonchophyllinae, with description of a new tribe and genus. *Occasional Papers, Museum of Texas Tech University, 320*, 1–23.

Parsons, D. J., Pelletier, T. A., Wieringa, J. G., Duckett, D. J., & Carstens, B. C. (2022). Analysis of biodiversity data suggests that mammal species are hidden in predictable places. *Proceedings of the National Academy of Sciences, 119*(14), e2103400119. https://doi.org/10.1073/pnas.2103400119

Patterson, B.D., Webala, P.W., Kerbis Peterhans, J.C., Goodman, S.M., Bartonjo, M. & Demos, T.C. (2019). Genetic variation and relationships among Afrotropical species of *Myotis* (Chiroptera: Vespertilionidae). *Journal of Mammalogy, 100,* 1130–1143.

Paz, A., Brown, J.L., Cordeiro, C.L.O., Aguirre-Santoro, J., Assis, C., Amaro, R.C., ... Carnaval, A.C. (2021). Environmental correlates of taxonomic and phylogenetic diversity in the Atlantic Forest. *Journal of Biogeography, 48*, 1377–1391. https://doi.org/10.1111/jbi.14083

Paz-Vinas, I., Jensen, E. L., Bertola, L. D., Breed, M. F., Hand, B. K., Hunter, M. E., ... & Hoban, S. (2021). Macrogenetic studies must not ignore limitations of genetic markers and scale. *Ecology Letters, 24*(6), 1282-1284. https://doi.org/10.1111/ele.13732

Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. *The R Journal 10*(1), 439-446. https://doi.org/10.32614/RJ-2018-009

Peduzzi, P. (2019). The disaster risk, global change, and sustainability nexus. *Sustainability, 11*(4), 957. https://doi.org/10.3390/su11040957

Pelletier, T. A., & Carstens, B. C. (2018). Geographical range size and latitude predict population genetic structure in a global survey. *Biology Letters, 14*(1), 20170566. https://doi.org/10.1098/rsbl.2017.0566

Pelletier, T. A., Carstens, B. C., Tank, D. C., Sullivan, J., & Espíndola, A. (2018). Predicting plant conservation priorities on a global scale. Proceedings of the National Academy of Sciences, 115(51), 13027-13032. https://doi.org/10.1073/pnas.1804098115

Pelletier, T. A., Parsons, D. J., Decker, S. K., Crouch, S., Franz, E., Ohrstrom, J., & Carstens, B. C. (2022). phylogatR: Phylogeographic data aggregation and repurposing. *Molecular Ecology Resources*, *22*(8), 2830– 2842. https://doi.org/10.1111/1755-0998.13673

Pereira, H. M. (2016). A latitudinal gradient for genetic diversity. *Science, 353*(6307), 1494-1495. https://doi/10.1126/science.aah6730

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., ... Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology, 55*(4), 595–609. https://doi.org/10.1080/10635150600852011

Puillandre, N., Brouillet, S., & Achaz, G. (2020). ASAP: assemble species by automatic partitioning. *Molecular Ecology Resources, 21*(2), 609– 620. https://doi.org/10.1111/1755-0998.13281

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Ramírez-Fráncel, L.A., García-Herrera, L.V., Losada-Prado, S., Reinoso-Flórez, G., SÁánchez-Hernández, A., Estrada-Villegas, S., ... Guevara, G. (2022). Bats and their vital ecosystem services: a global review. *Integrative Zoology, 17*, 2–23. https://doi.org/10.1111/1749-4877.12552

Reeder, D. M., Helgen, K. M., & Wilson, D. E. (2007). Global trends and biases in new mammal species discoveries. Occasional Papers, Museum of Texas Tech University.

Revell, L. J. (2024). phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ*, 12, e16505.

Rodríguez-Posada, M.E., Morales-Martínez, D.M., Gutiérrez-Sanabria, D.R., Calderón-Capote, M.C. & Fernández-Rodríguez, C. (2018). New records of the rare bat *Vampyressa melissa* Thomas, 1926 (Chiroptera: Phyllostomidae: Stenodermatinae) from the eastern slope of the Andes in the Colombian Orinoco basin. *Mammalia, 82*, 298–302.

Rodríguez-Posada, M. E., Morales-Martínez, D. M., Ramírez-Chaves, H. E., Martínez-Medina, D., & Calderón-Acevedo, C. A. (2021). A new species of Long-eared Brown Bat of the genus *Histiotus* (Chiroptera) and the revalidation of *Histiotus colombiae*. *Caldasia*, *43*(2), 221–234. https://doi.org/10.15446/caldasia.v43n2.85424

Rojas, D., Warsi, O. M. & Dávalos, L. M. Bats (Chiroptera: Noctilionoidea) challenge a recent origin of extant neotropical diversity. (2016). *Syst. Biol. 65*, 432–448. https://doi.org/10.1093/sysbio/syw011

Rosauer, D. F., Byrne, M., Blom, M. P., Coates, D. J., Donnellan, S., Doughty, P., ... Moritz, C. (2018). Real-world conservation planning for evolutionary diversity in the Kimberley, Australia, sidesteps uncertain taxonomy. *Conservation Letters, 11*(4), e12438. https://doi.org/10.1111/conl.12438

Rosauer, D. F., Pollock, L. J., Linke, S., & Jetz, W. (2017). Phylogenetically informed spatial planning is required to conserve the mammalian tree of life. *Proceedings of the Royal*

Society B: Biological Sciences, 284(1865), 20170627. https://doi.org/10.1098/rspb.2017.0627

Saraiva, D. D., dos Santos, A. S., Overbeck, G. E., Giehl, E. L. H., & Jarenkow, J. A. (2018). How effective are protected areas in conserving tree taxonomic and phylogenetic diversity in subtropical Brazilian Atlantic Forests? *Journal for Nature Conservation, 42*, 28-35. https://doi.org/10.1016/j.jnc.2018.02.002

Shafer, A. B., Cullingham, C. I., Côté, S. D., & Coltman, D. W. (2010). Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular ecology, 19*(21), 4589-4621.

Shi, J. J., & D. L. Rabosky. (2015). Speciation dynamics during the global radiation of extant bats. *Evolution*, *69*:1528–1545.

Simmons, N. B., & Cirranello, A. L. (2024). Bat Species of the World: A taxonomic and geographic database. Version 1.6. Accessed on 10/3/2024.

Smith, B. T., McCormack, J. E., Cuervo, A. M., Hickerson, M. J., Aleixo, A., Cadena, C. D., ... Brumfield, R. T. (2014). The drivers of tropical speciation. Nature, 515(7527), 406-409.

Smith, B. T., Seeholzer, G. F., Harvey, M. G., Cuervo, A. M., & Brumfield, R. T. (2017). A latitudinal phylogeographic diversity gradient in birds. *PLoS biology, 15*(4), e2001073. https://doi.org/10.1371/journal.pbio.2001073

Smith, S.A. & O'Meara, B.C. (2012). treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics*, **28**, 2689–2690.

Solari, S., & Martínez-Arias, V. (2014). Cambios recientes en la sistemática y taxonomía de murciélagos Neotropicales (Mammalia: Chiroptera). *Therya*, *5*(1), 167–196. https://doi.org/10.12933/therya-14-180

Soto-Centeno, J. A., & Calderón-Acevedo, C. A. (2023). Global change and the conservation of Caribbean bat communities. In A. Kurta & A. Rodríguez-Durán (Eds.), *Bats of the West Indies, A Natural History and Field Guide* (pp. 72–86). Cornell University Press.

Stevens, R. D., & Gavilanez, M. M. (2015). Dimensionality of community structure: phylogenetic, morphological and functional perspectives along biodiversity and environmental gradients. *Ecography, 38*(9), 861-875. https://doi.org/10.1111/ecog.00847

Tachikawa, T., Kaku, M., Iwasaki, A., Gesch, D. B., Oimoen, M. J., Zhang, Z., ... Carabajal, C. (2011). ASTER global digital elevation model version 2-summary of validation results. *NASA*.

Taylor, P. J., Stoffberg, S., Monadjem, A., Schoeman, M. C., Bayliss, J., & Cotterill, F. P. (2012). Four new bat species (*Rhinolophus hildebrandtii* complex) reflect Plio-Pleistocene divergence of dwarfs and giants across an Afromontane archipelago. *Plos One, 7*(9): e41744.

Teeling, E. C., Springer, M. S., Madsen, O., Bates, P., O'Brien, S. J., & Murphy, M. J. (2005). A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science, 307*, 580–584. https://doi.org/10.1126/science.1105113

Theissinger, K., Fernandes, C., Formenti, G., Bista, I., Berg, P. R., Bleidorn, C., ... Zammit, G. (2023). How genomics can help biodiversity conservation. *Trends in Genetics, 39*(7), 545-559.

Thompson, C. E., Pelletier, T. A., & Carstens, B. C. (2021). Genetic diversity of North American vertebrates in protected areas. *Biological Journal of the Linnean Society, 132*(2), 388-399. https://doi.org/10.1093/biolinnean/blaa195

Tuanmu, M. N., & Jetz, W. (2015). A global, remote sensing-based characterization of terrestrial habitat heterogeneity for biodiversity and ecosystem modelling. *Global Ecology and Biogeography, 24*(11), 1329-1339. https://doi.org/10.1111/geb.12365

Velazco, P.M., Soto-Centeno, J.A., Fleck, D.W., Voss, R.S. & Simmons, N.B. (2017). A new species of nectar-feeding bat of the genus *Hsunycteris*. *American Museum Novitates*, 1–26.

Voelker, G., Outlaw, R. K., & Bowie, R. C. (2010). Pliocene forest dynamics as a primary driver of African bird speciation. *Global Ecology and Biogeography, 19*(1), 111-121. https://doi.org/10.1111/j.1466-8238.2009.00500.x

Warwick, R. M., & Clarke, K. R. (1995). New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. *Marine ecology progress series, 129*, 301-305. https://doi.org/10.3354/meps129301

Watson, J. E., & Venter, O. (2017). A global plan for nature conservation. *Nature, 550*(7674), 48-49.

Webb, C. O. (2000). Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist, 156*(2), 145-155. https://doi.org/10.1086/303378

Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual review of ecology and systematics, 33*(1), 475-505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448

Weiskopf, S. R., Rubenstein, M. A., Crozier, L. G., Gaichas, S., Griffis, R., Halofsky, J. E., ... Whyte, K. P. (2020). Climate change effects on biodiversity, ecosystems, ecosystem services, and natural resource management in the United States. *Science of the Total Environment, 733*, 137782. https://doi.org/10.1016/j.scitotenv.2020.137782

Yi, X., Latch, E. K., Lim, B. K., Koroiva, R., Da Rocha, P. A., & Feijó, A. (2024). UCE-derived mitochondrial phylogeny reveals pervasive mito-nuclear discordances in serotine bats (genus *Eptesicus*) and complex evolutionary history in Eptesicus (Histiotus). *Mammalian Biology*, *104*(4), 417–430. https://doi.org/10.1007/s42991-024-00418-2

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics, 29*(22), 2869-2876. https://doi.org/10.1093/bioinformatics/btt499

Data availability statement

All intermediate datafiles and list of DNA sequence accessions are available as supplemental tables (Table S1-8). All data analysis scripts and supplemental files 1-12 are on Dryad: Reviewer URL: http://datadryad.org/stash/share/Wrp3Bem5LbdGpxC8bdzp8h0f3eyMJW0zoc6hNW11y8A.

Table 1. Percent variation explained for all random forest models and the number of predictor variables used in each model. Cutoff groups were based on correlation coefficients between all variables, removing variables that had 0.7, 0.8, or 0.9 correlation or higher to another variable in the set. BIOCLIM variables for present day (Current), the Last Glacial Maxima (LGM), the Last Interglacial (LIG), and the Pliocene (Plio) were grouped into their own categories. The final two categories were geographic variables (Geo) and miscellaneous variables (Misc). Each category was also run with the removal of highly correlated variables with a 0.9 cutoff (additional of .9 to each category).

	MPD % Var		**MNTD % Var**		**# variables**	
Group	**25-eco**	**109-eco**	**25-eco**	**109-eco**	**25-eco**	**109-eco**
0.7 Cutoff	42.72	56.86	14.2	14.56	37	41
0.8 Cutoff	60.04	55.88	12.24	19.36	55	64
0.9 Cutoff	68.23	54.83	9.79	15	101	110
Geo	65.18	50.46	10.45	11.81	9	14
Geo.9	46.52	48.77	6.69	10.73	5	6
Current BIOCLIM	58.18	37.66	-4.37	6.66	38	43
Current BIOCLIM.9	58.71	41.63	-5.56	11.45	23	26
Misc	65.41	51.87	0.56	13.62	36	35
Misc.9	55.16	53.35	2.01	14.45	29	24
LGM BIOCLIM	57.81	41.09	11.86	2.53	57	62
LGM BIOCLIM.9	60.11	44.59	8.49	5.49	19	24
LIG BIOCLIM	46.33	37.29	8.5	7.39	57	62
LIG BIOCLIM.9	52.33	38.34	4.92	2.84	16	20
Plio BIOCLIM	57.1	34.27	16.48	0.51	42	47
Plio BIOCLIM.9	42.3	39.61	14.91	0.82	9	15

Figures and figure legends

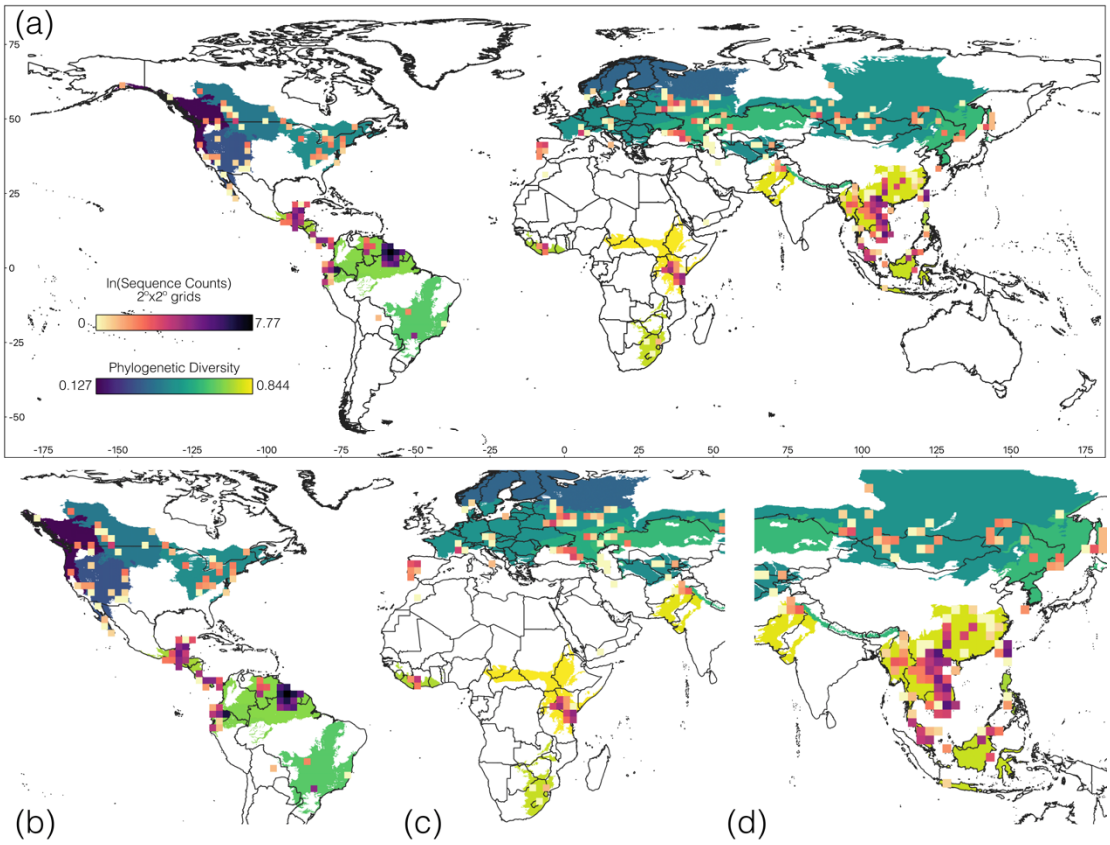

Figure 1: 25-ecoregions. (a) Phylogenetic diversity of bats Worldwide and number of DNA sequences (given in natural logarithm of actual sequence numbers in a 2x2 degree grid, for better visualization). (b) The Neotropics (c) East Africa (d) Southeast Asia

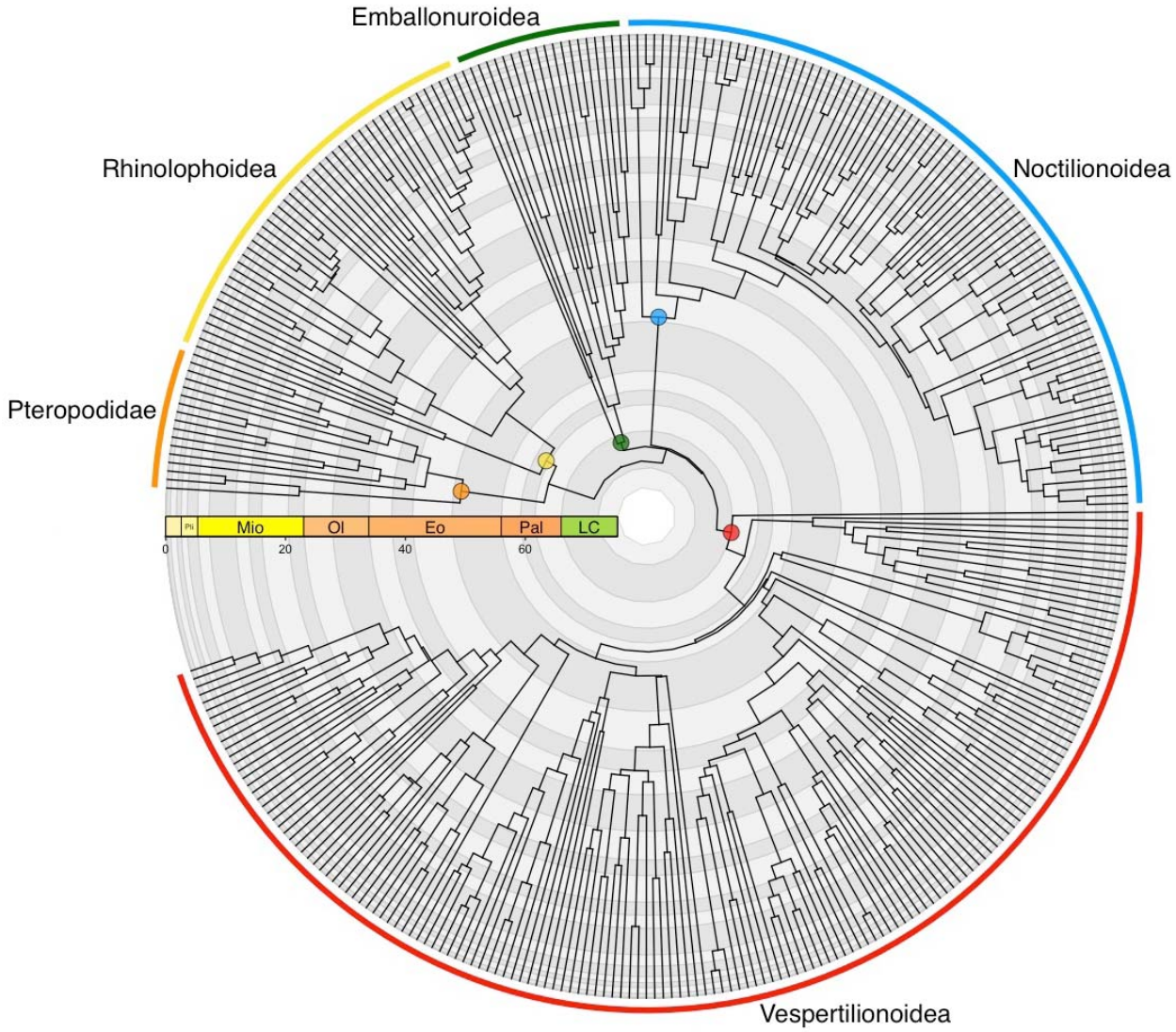

Figure 2: Time calibrated phylogeny of bats worldwide representing 17 families and all extant superfamilies. Estimates of divergence were performed in TreePL using the RaxML-NG COI phylogeny. Color coded nodes represent the superfamily and family taxonomic designations at tips of the phylogeny. Time scale represented in millions of years and following the geological time designations Pal= Paleogene, Eo = Eocene, Ol = Oligocene, Mio = Miocene, and Pli = Pliocene (the Pleistocene and Holocene not labeled).

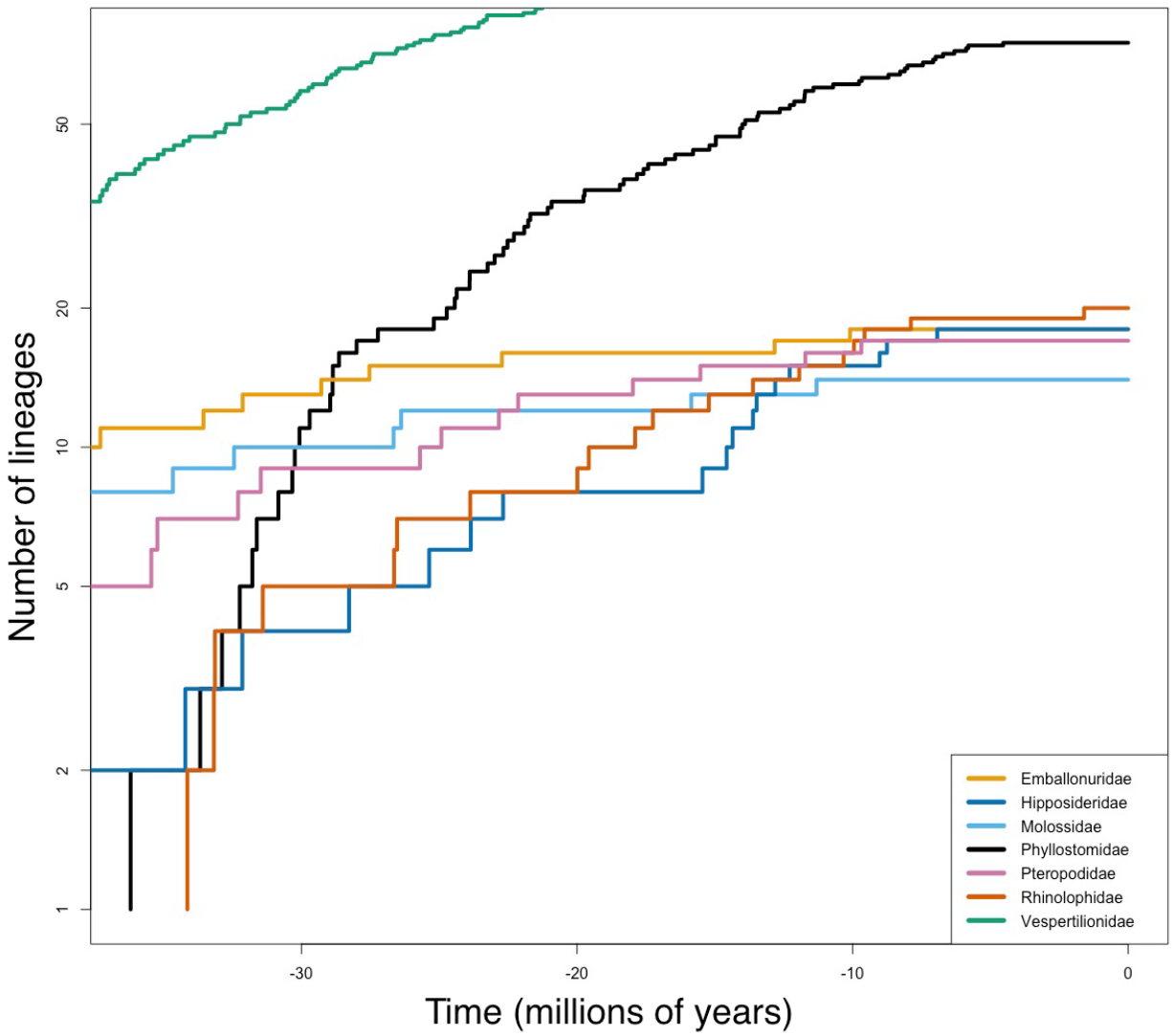

Figure 3: Lineage through time (LTT) plot based on the dated maximum likelihood RaxML-NG COI tree showing lineage accumulation in the last 30 million years for seven bat families. Plot scaled using the family Phyllostomidae (New World leaf-nosed bats, black line) as reference.

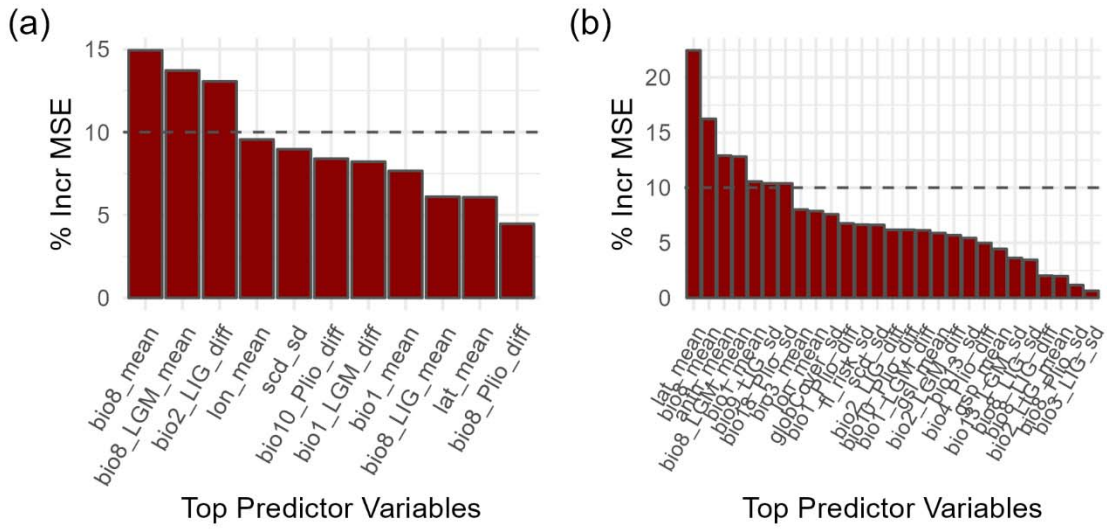

Figure 4: Bar plot of % increase in MSE for the top variable random forest for both (a) the 25-ecoregion dataset and (b) the 109-ecoregion dataset. The dashed line indicates the 10% increase in MSE cutoff used for determining top variables.

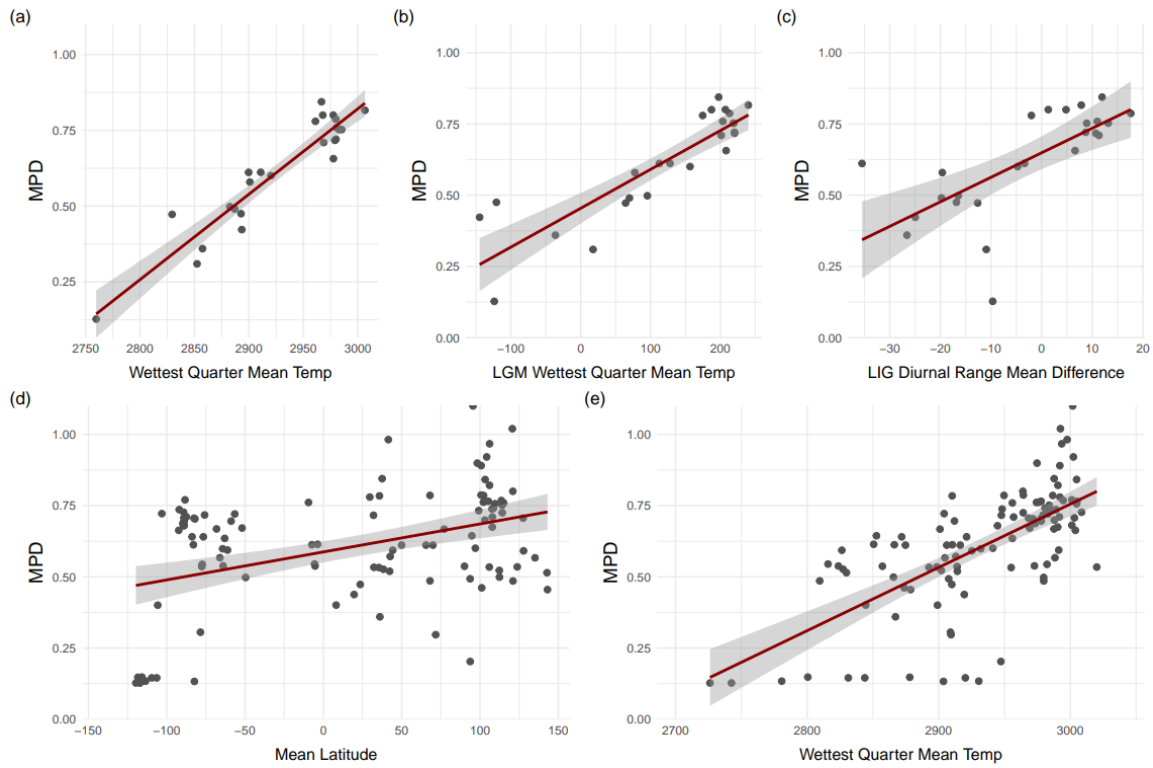

Figure 5: Scatterplots with linear regression lines for top variables for both the 25-ecoregion dataset (a, b, c) and the 109-ecoregion dataset (d, e).