No link between population isolation and speciation rate in squamate reptiles

Author List

Sonal Singhal[1*], Guarino R. Colli[2], Maggie R. Grundler[3,4], Gabriel C. Costa[5], Ivan Prates[6,7], Daniel L. Rabosky[6,7*]

[1]Department of Biology, CSU Dominguez Hills, Carson, California 90747
[2]Departamento de Zoologia, Universidade de Brasília, Brasília, DF, Brazil
[3]Department of Environmental Science, Policy, & Management, University of California Berkeley, Berkeley, CA 94720
[4]Museum of Vertebrate Zoology, University of California Berkeley, Berkeley, California, 94720
[5]Department of Biology and Environmental Sciences, Auburn University at Montgomery, Montgomery, Alabama
[6]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109
[7]Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109

* corresponding authors: sonal.singhal1@gmail.com and drabosky@umich.edu

Classification

Biological Sciences: Evolution

Keywords

Speciation, macroevolution, squamates, population isolation, gene flow

Abstract

Rates of species formation vary widely across the tree of life and contribute to massive disparities in species richness among clades. This variation can emerge from differences in metapopulation-level processes that affect the rates at which lineages diverge, persist, and evolve reproductive barriers and ecological differentiation. For example, populations that evolve reproductive barriers quickly should form new species at faster rates than populations that acquire reproductive barriers more slowly. This expectation implicitly links microevolutionary processes (the evolution of populations) and macroevolutionary patterns (the profound disparity in speciation rate across taxa). Here, leveraging extensive field sampling from the Neotropical Cerrado biome in a biogeographically-controlled natural experiment, we test the role of an important microevolutionary process – the propensity for population isolation – as a control on speciation rate in lizards and snakes. By quantifying population genomic structure across a set of co-distributed taxa with extensive and phylogenetically independent variation in speciation rate, we show that broad-scale patterns of species formation are decoupled from demographic and genetic processes that promote the formation of population isolates. Population isolation is likely a critical stage of speciation for many taxa, but our results suggest that interspecific variability in the propensity for isolation has little influence on speciation rates. These results suggest that other stages of speciation – including the rate at which reproductive barriers evolve and the extent to which newly formed populations persist – are likely to play a larger role than population isolation in controlling speciation rate variation in squamates.

Significance Statement

Speciation rate measures how quickly a species gives rise to new species, and this rate varies up to 50-fold across vertebrate groups. In this study, we explore one hypothesis that explains this variation: species that form geographically isolated populations more readily should also form new species more readily and thus should have higher speciation rates. This hypothesis links microevolutionary studies of speciation with macroevolutionary studies of biodiversity. We test this hypothesis using a diverse set of lizard and snake species found in the South America savannahs. We find no effect of geographic population isolation on speciation rates. Our results suggest that other stages in the speciation process are more important controls on speciation rate variation.

Introduction

Speciation rates vary widely across the tree of life, and this variation contributes to profound disparities in species richness among clades, across regions, and through time (1). Despite the importance of speciation rates for broad-scale diversity patterns, only recently have we been able to quantify and compare these rates across taxa. Ernst Mayr, who laid much of the conceptual foundation of modern speciation research, noted in his landmark book *Animal Species and Evolution* that "[t]here is perhaps no other aspect of speciation of which we know so little as its rate. We shall probably never have very accurate information on this phenomenon" (p. 575) (2). Yet today, our understanding of speciation rate variation has been fundamentally transformed by the ascendance of molecular-based phylogenetic inference. The

availability of time-calibrated phylogenetic trees (3, 4), in concert with analytical models of macroevolution (5), has demonstrated wide variation in the rate of speciation, even among closely-related groups of organisms (Fig. 1, Fig. S1; (6–11). In mammals, for example, the fastest speciating lineages (*Rattus* rats: 1.12) are estimated to form new species nearly 100-fold faster than the slowest speciating lineages (aardvarks: 0.013 species myr-1; 12).

To understand why speciation rate varies, we can deconstruct the speciation process into population-level factors that control, at least in principle, the observed rate of new lineage origination at the macroevolutionary scale (2, 13, 14). "Successful" speciation typically requires that new populations become geographically isolated (15–17) and that reproductive barriers evolve between the populations. Finally, these nascent species must persist while they acquire reproductive and potentially ecological isolation from other populations. Any factor modifying the rate of these processes can potentially influence speciation rate (14, 18, 19). For example, species with traits leading to high dispersal (e.g., marine organisms with long larval periods) are predicted to maintain panmictic populations connected by gene flow across large areas in the face of biogeographic processes that might otherwise fragment species with lower capacity for cohesion (20–22). If the formation of population isolates is a rate-limiting step for speciation, we would then predict that species with high dispersal should exhibit reduced speciation rates compared to species with low dispersal. Similarly, biogeographic and climatic factors — including the dynamic nature of some high elevation and high-latitude landscapes — can facilitate increased speciation rates by promoting population isolation in newly colonized geographic areas (23–25). This framework thus links the microevolutionary processes leading to species formation with macroevolutionary patterns of species diversification (26), including the speciation rates we typically estimate from phylogenetic trees or the fossil record (Fig. 1).

Directly estimating the extent to which variation in population processes affects speciation requires comparable datasets on the population processes themselves (26). These must be collected across a broad sample of taxa characterized by phylogenetically-independent variation in both evolutionary rates as well as the focal population processes. Collecting such datasets is a daunting task for most groups of organisms (27–29). For this reason, studies often use organismal traits as proxies for population processes, implicitly assuming that these proxies capture lineage-specific propensities towards reproductive isolation or population splitting (30, 31). For example, avian wing morphology has been used as a proxy for dispersal and population isolation (32, 33), and sexual dimorphism has been used as a proxy for the strength of sexual selection and thus the rate at which reproductive barriers evolve (34, 35). Results from this trait-based approach have generally been mixed. Traits often exhibit inconsistent associations with speciation across clades (36), and, even for a given clade, purported relationships vary when challenged with alternate methodologies or different datasets (37, 38). This inconsistency may partially reflect the differential effects of specific organismal traits upon diversification across clades. For example, the evolution of reproductive barriers might be a rate-limiting step in one clade but not another. However, this inconsistency might also indicate that these organismal traits are imperfect proxies for the population-level processes they are trying to capture.

Here, we apply a direct approach to determining the factors that predict speciation rate. We measure population isolation across multiple species and then test if population isolation predicts speciation rate, without relying on organismal traits as proxies for isolation. The formation of geographically isolated populations is an initial and essential step in most speciation events (15–17), and population structure has long been recognized as a potential source of incipient species, across analyses of phylogeographic structure (39, 40), studies of ecotypes and subspecies (41–43), and explorations of species richness in the fossil record (22, 44). In this study, we focus on the macroevolutionary consequences of population isolation in the lizards and snakes from a single geographic region: the Neotropical Cerrado, South America's second largest biome (45). The Cerrado savannas harbor one of the world's most diverse squamate reptile communities (46, 47), enabling us to sample phylogenetically-distinct lineages that span nearly the complete range of speciation rates as measured from fully-sampled squamate phylogenies (Fig. 2A). Because most of these species have large ranges that span the same geographic area (Fig. 2B, S2, S3), they likely experienced similar biogeographic histories. Thus, our analyses partially control for the effect of extrinsic factors like biogeographic changes and environmental conditions on speciation rates.

We characterized population isolation across 66 snake and lizard taxa from vouchered samples collected over 20+ years of fieldwork in the Cerrado, generating a genome-wide perspective on intraspecific variation (mean: 5019 loci per individual) from 398 individuals. We then measured how levels of genetic differentiation accumulate across geographic space (e.g., isolation-by-distance [IBD]). The slope of this IBD relationship both reflects population density of demes and the extent of gene flow among them (48) and serves as a reasonably direct measure of species cohesion, which we define here as the capacity of species' geographic ranges to resist allopatric fragmentation into new population isolates. We then tested a foundational hypothesis in speciation research — the formation of geographically isolated populations is often an initial and essential step in speciation — by determining if an increased propensity for population isolation correlates with higher speciation rates.

Results

We collected an average of 3.2 Mb across 5019 nuclear loci and 3.8 kb of the mitochondrial genome for the 398 individuals in this study (Fig. 2D, Fig. S4, Table S1). We deliberately sampled geographically-unique taxon-locality combinations to maximize spatial coverage, and thus our dataset spans 386 distinct localities. Using these high-coverage and high-quality data, we first delimited species-level taxa (operational taxonomic unit; OTU; see *Methods*) to ensure our analyses were not confounded by cryptic diversity or taxonomic uncertainty. Fifty-one of our 66 nominal species (76%) directly corresponded to an OTU; for the purposes of analysis, we revised 15 nominal species in total. Our final dataset consisted of 59 provisional OTUs spanning 375 individuals (Fig. 2C, Fig. S5, S6).

For each OTU, we used an average of 28.5K variant sites to estimate the slope at which genetic isolation (F_{ST}) accumulates across space (β_{IBD}, Fig. 3). On average, geographic distance explained 31% of the variation in F_{ST} across individuals within OTUs (Fig. S7). OTUs showed notable differences in β_{IBD}, with values ranging from -0.41 (indicating low levels of differentiation

across space) to 2.48 (indicating steep turnover of genetic variation across space; Fig. 3, Fig. 4A, Fig. S8). β_{IBD} exhibited moderate phylogenetic signal (λ = 0.24, p = 0.009, Fig. 4A), suggesting that the variation in β_{IBD} across taxa is phylogenetically conserved. β_{IBD} is expected to vary as a function of both population density of demes and levels of migration between them (48). Accordingly, we identified organismal traits that might serve as proxies for density or migration and tested if they could predict β_{IBD}. We find that β_{IBD} is correlated with a measure of body elongation (phylogenetic generalized least squares [PGLS], p = 0.004, Fig. 5, Fig. S9, Table S2) that likely affects both population density and migration. Animals that were more elongated and had a more snake-like form (e.g., snakes and legless lizards) had lower levels of differentiation across space than animals with more lizard-like forms. Finally, although the genetic loci used in this study are likely under purifying selection (49), estimates of genetic diversity and differentiation across these markers and putatively neutrally evolving loci were highly concordant (r = 0.81-0.97; Fig. S10). These results suggest our estimates of β_{IBD} are likely robust to the evolutionary history of the loci used for inference.

Using the most comprehensive squamate phylogeny available to date (50), we estimated speciation rates using both a model-based ($\lambda_{CLaDS;}$ 8, 51) and semi-parametric approach (λ_{DR}; 6). These estimates of speciation rate are highly correlated both across all squamates (n = 9755, r = 0.87) and across our focal taxa (n = 59, r = 0.87, Fig. S11). Given this high concordance, all subsequent results focus on model-based estimates of speciation rate. Speciation rates across all squamates ranged from 0.018 to 0.424 species myr^{-1} with the highest rates occurring in snakes (Fig. 2A). Speciation rates in our focal taxa ranged from 0.027 to 0.280 species myr^{-1}, encompassing nearly the full variation seen in squamates (Fig. 2A).

We tested the prediction that OTUs with greater rates of population differentiation would show greater speciation rates, as would be expected if population differentiation is a rate-limiting step in speciation. We found no such pattern (PGLS r = 0.04, p-value = 0.77, Fig. 4). Rather, we recovered some evidence for an inverse relationship, which can be seen when comparing β_{IBD} and rates of speciation in snakes versus lizards (Fig. 5). Snakes have high speciation rates but low β_{IBD}, opposite to our predictions.

We confirmed that these patterns are robust to possible methodological and technical artifacts by conducting a series of analyses in which we accounted for limited sampling, measurement error, alternate measures of genetic differentiation across space, possible taxonomic error, and the effects of biogeography. Across all these analyses, we recover the same result: no evidence for a relationship between β_{IBD} and speciation rate (Fig. S12, Table S3). Most importantly, restricting our analyses to only those OTUs with significant β_{IBD} relationships (n = 29) did not change these results and, in fact, yielded a non-significant but negative correlation between β_{IBD} and speciation rate (r = -0.004, p = 0.98, Table S3). Alternate measures of genetic differentiation (e.g., d_{xy} across geographic space) are only weakly correlated with each other (Fig. S13, Fig. S14), and across all these robustness analyses, our finding that population structure and speciation rates are decoupled holds (Table S3). Further, although we sampled relatively few individuals for some OTUs, both subsampling and simulations show that few individuals can effectively recover β_{IBD} similar to that seen with larger datasets (Fig. S15 - 17). Finally, we conducted a power analysis to estimate our power to identify a correlation between

isolation and speciation rates, if one exists. We found the power to detect a true correlation was moderate if correlations were <0.5, and power was >65% if correlations were ≥0.5 (Fig. S18A). However, this analysis also suggests that the potential correlation between β_{IBD} and speciation rates is unlikely to be substantial – and is indeed equally likely to be negative or positive – given the small magnitude of our inferred correlation between β_{IBD} and speciation rates (mean r = 0.04; 95% range: -0.37 - 0.45, Fig. S18B).

Discussion

We recovered wide variation in population isolation, as indexed by β_{IBD}, across our focal taxa (Fig. 4). Moreover, we found evidence that β_{IBD} is an evolutionarily structured trait – close relatives of species that are prone to population isolation are themselves prone to isolation – and isolation varies with morphology, with more elongate species showing lower β_{IBD} than more lizard-like species (Fig. 4, 5). Elongate species might either experience greater dispersal or maintain greater population density, either of which would result in lower β_{IBD} (Fig. S19). Whatever the mechanism, these results suggest that species cohesion varies across taxa and that it is influenced by organismal traits, consistent with previous studies in squamates and other taxa (21, 28, 52).

Species cohesion is a fundamental property of species that influences the formation of new species, as described in both verbal and mathematical models (39, 53, 54). The homogenizing effects of gene flow (55) has long been thought to inhibit speciation by reducing the probability of population fission. We might thus predict those taxa that have a higher potential for isolation, as indexed by IBD slopes, will also have greater propensity to form new species (20, 55). Yet, despite our broad phylogenetic sampling, we found no correlation between the potential for isolation and speciation rates (Fig. 4).

Factors such as limited sampling and environmental and historical effects can make estimating IBD slopes challenging, potentially weakening our ability to recover a true underlying relationship between IBD slope and speciation rate. Our IBD slopes appear to be capturing a real property of species propensity to isolation, however. These slopes are phylogenetically structured and correlate with organismal morphology (Fig. 4, 5). Further, our results are qualitatively and quantitatively similar across a series of robustness analyses (Fig. S12, Table S3). It is unlikely that our results are solely due to low statistical power. Perhaps the most substantive and well-supported contrast in our dataset is the evolutionary split between lizards and snakes. This partition accounts for much of the variation in both speciation rate (ANOVA r^2 = 0.54) and IBD slope (ANOVA r^2 = 0.27) across our full dataset. Under our hypothesis, we would expect to see much higher patterns of IBD in snakes than lizards, given that speciation rates are markedly higher in snakes (lizards: average λ_{CLaDS} = 0.068; snakes: average λ_{CLaDS} = 0.16). However, we see the opposite pattern (Fig. 5). While this pattern reflects just a single evolutionary contrast, it is a major feature of our dataset and suggests our negative result is robust and biologically meaningful.

Most previous studies exploring the link between population isolation and diversification have used organismal traits related to dispersal as proxies for gene flow and population isolation. For example, in birds, species with more elongate, narrower wings are hypothesized to have higher dispersal (33), higher levels of gene flow, and thus reduced rates of population isolation. Lower rates of population isolation would then result in lower rates of diversification, as has been seen in some avian datasets (25, 56, 57). Other organismal traits that have been hypothesized to predict both patterns of genetic connectivity and broad-scale diversification include pollination strategy (58), length of pelagic dispersal in marine species (44, 59 but see 38), and seed size in angiosperms (60). These studies necessarily make several simplifying assumptions about how organismal traits impact dispersal and thus gene flow. For example, although levels of population isolation result both as a function of migration and local genetic drift, these analyses focus solely on the role of dispersal in determining levels of population isolation. However, as a whole, these studies confirm the potentially important role of variation in dispersal in driving broad-scale diversity patterns (but see 61).

Studies that have instead directly estimated levels of population isolation have found mixed results (21, 27–29, 52, 62). In particular, a previous study in lizards – focused solely on a ~300 species radiation endemic to Australia (28) – also found no connection between population isolation and speciation (). What might explain our results and the mixed conclusions of previous studies? First, population isolation can be quantified in many ways. Population isolation spans a continuum that ranges from isolation-by-distance to discrete phylogroups that experience little gene flow amongst themselves. In the present study, we chose to estimate population isolation directly by measuring patterns of IBD. A pattern of IBD reflects an early stage in population isolation and primarily captures the role of declining gene flow over distance in determining genetic differentiation. IBD is thus a continuous measure of genetic differentiation (63). How IBD interacts with landscape stability, geographic barriers, and environmental gradients then determines the likelihood of formation of discrete breaks characteristic of phylogeographic lineages (64, 65). We expect these continuous and discrete measures of population isolation to be correlated; species with greater IBD should also exhibit greater levels of phylogeographic structuring. With our current sampling, we cannot test this expectation, and it also remains relatively untested in the broader literature (26). If continuous and discrete measures of population isolation are uncorrelated, then using alternate metrics of population isolation like the number of genetic population clusters (63), phylogenetically distinct lineages within a species (27), or subspecies (66, 67) might better explain speciation rate. That said, despite its simplicity, IBD explains 31% of the average variation in patterns of genetic divergence across our focal species (Fig. S7), suggesting IBD remains a powerful approach for measuring population isolation. Further, exploration of alternate metrics of isolation – including levels of genetic differentiation at a fixed geographic distance – recover the same patterns seen when isolation is measured as IBD (Table S3).

Additionally, population isolation might impact speciation in ways more complicated than typically outlined in verbal and formal models (61). For example, an increased propensity to form population isolates might lead to smaller populations, which would be more vulnerable to demographic fluctuations, making them more likely to go locally extinct (44). In such a scenario, species cohesion — which is expected to increase with gene flow — may have opposing effects

on rate of population formation and population extinction. The net effect of increased population structure on speciation rate might then be neutral.

Alternatively, perhaps population isolation is not a significant factor in determining speciation rate variation at all. Indeed, our power simulations suggest that if such an effect exists, it is likely to be small (Fig. S18). Instead, other unmeasured intrinsic and extrinsic factors might be the source of speciation rate variation in squamates. Speciation consists of multiple stages, and variation in progress through any of these stages might influence speciation rate. For example, some species might evolve reproductive barriers more quickly than others, whether because they undergo more rapid body size divergence (11), more rapid changes in genome structure and organization (68), or more intense levels of sexual selection (34). If evolving reproductive barriers serves as a rate-limiting step on speciation, then species that evolve reproductive barriers more quickly will also have higher speciation rates (but see 69). Additionally, a variety of factors might influence the propensity for new species to persist through time (19, 70), including the evolution of ecological traits that reduce competition and facilitate range expansions and/or secondary sympatry (2, 14, 71–73). Further, historical changes in the environment and geography can have marked effects on population fission and thus speciation rate (23–25). In our focal system, however, where all species are found in a common geographic region, the impact of extrinsic forces on speciation rate variation is likely to be more uniform across taxa. Of course, multiple factors could reasonably be acting concurrently or unevenly to influence speciation rates across the squamate tree of life. Given the broad phylogenetic scale of this study (~50 species spanning all squamate diversity, representing ~180 million years of evolution), the forces that control speciation rate variation might be so variable across clades that the overall impact of any single force ultimately is rather diffuse.

Focused studies of species complexes — including many in lizards and snakes (74–77) — have revealed the central role of population isolation in the formation of new species, consistent with the conceptual model of speciation popularized by Ernst Mayr (2, 78). Given this, it is perhaps intuitive that population-level processes such as population isolation should predict broad-scale speciation patterns. Here, we have shown that “population isolation” is itself a trait that varies widely among clades, consistent with numerous studies that have directly or indirectly linked organismal traits to broad-scale patterns of population structure. And like most other groups of organisms (9–11), squamate reptiles from Neotropical savannas are characterized by extensive clade-specific variation in rates of species formation (Fig. 1). However, we find no support for the hypothesis that population isolation – as indexed by contemporary patterns of population structure – is predictably related to phylogenetic variation in the rate of species formation.

This study provides one of the most comprehensive tests of a single component of the speciation process as described by Mayr and others (2, 13, 14). Generating the comparative datasets to test the role of other steps in the speciation process will be challenging. In particular, we need to assemble datasets that measure the rate at which reproductive barriers evolve and the likelihood that populations persist across species that vary in their speciation rate. Relative to inferring present-day levels of population isolation, measuring reproductive barriers and species persistence is difficult. For example, quantifying reproductive barriers in even a single species or species complex can require decades of work for a single research group, in both

controlled and natural settings (79–83). To understand how processes identified from single species groups contribute to broad-scale diversity patterns, biologists studying a diversity of taxa will need to collaborate to measure population-level processes in standardized ways that can be collated across study systems and tested within a common analytical framework (84). Only then can we truly understand the causes of speciation and the extent to which those causes underlie the dynamics of biological diversity at the largest scales of time and space.

Methods

Sampling and Genetic Data Collection

We sampled squamate (lizard and snake) taxa broadly across clades that differ in speciation rates, thus increasing the power of our study by maximizing the amount of phylogenetically-independent variation in both diversification rates and biological traits (Fig. 2A). Our study leverages over 20 years of field sampling in the Cerrado and adjoining biomes (46, 47, 85, 86). Most of these samples correspond to voucher specimens preserved at Coleção Herpetológica da Universidade de Brasília (CHUNB). Across all samples, we identified nominal species for which individuals were sampled at ≥4 locations. For each of those species, we included one individual per sampling locality. Given our levels of genomic sampling, sampling one individual is akin to sampling multiple individuals in a population for fewer loci (87). In total, we sampled 398 individuals from 66 nominal species (Fig. 2B, Fig. S2, S3).

We sequenced homologous loci across species using a target-capture approach. We used the Squamate Conserved Loci (SqCL) as our targets, a set of 5,462 loci that span ultraconserved elements (UCEs), anchored hybrid enrichment loci (AHEs), and single-copy genes used traditionally in phylogenetics (88–90). This bait set has been shown to work effectively across lizards and snakes (89, 91). For each individual, we extracted DNA from either tail or liver tissue using high-salt extraction (92). After shearing DNA to a modal length of ~350 base pairs using a sonicator, the commercial provider RAPiD Genomics (Gainesville, Florida, USA) then prepared doubly-indexed libraries following standard Illumina protocols. We created pools of 16 equimolar libraries, grouping closely related individuals into the same pool. The commercial provider Arbor Biosciences (Ann Arbor, Michigan, USA) used these pools for the target-capture experiment, following the myBaits v3 protocol. Finally, we combined six pools (or 96 libraries) and sequenced each set on one 125-bp paired-end lane of an Illumina HiSeq 4000.

Genetic Data Analysis

Data processing

For the raw sequencing reads per sample, we first trimmed adapters using Trimmomatic v0.39, removed low-quality sequences, and then merged overlapping paired reads using PEAR v0.9.11 (93, 94). Next, we then assembled cleaned reads using Trinity v2.1 (95) and annotated the resulting contigs by comparing them to our SqCL targets via blat v36x2 (96).

In target capture experiments, typically anywhere from 50 - 80% of sequencing reads map to targeted loci; the remaining reads are called by-catch. These by-catch reads can be used to assemble high-copy loci like the mitochondrial genome, which can be an important source of additional genetic data. We assembled mitochondrial genomes from our raw sequence reads using MITObim v1.9.1 (97). As a starting reference, we used the full mitochondrial genome from the most closely-related species available on NCBI GenBank. We then annotated each assembled mitochondrial genome for the 13 coding mitochondrial genes and ribosomal RNA 12S and 16S genes using exonerate v2.4.0 (98). Finally, we extracted these loci, aligned them using mafft v7.310 (99), and concatenated them to generate a mtDNA alignment across all individuals.

Species delimitation

Studies across lizards and snakes from both temperate and tropical regions have shown that nominal species often consist of multiple, cryptic lineages (100–102). Accordingly, we first analyzed our sampled individuals' genetic data to ensure that nominal species assignments reflected evolutionary relationships. We delimited lineages (operational taxonomic units; OTUs) based on evidence for evolutionary cohesion and independence. As detailed below, we used a combination of phylogenetic, population genetic, and spatial data to identify putative species-level taxa that form a monophyletic group and show a single and continuous isolation-by-distance (IBD) curve across their geographic range. We refrain from providing formal taxonomic recommendations given our study only collects genetic data.

To determine if individuals within nominal species form monophyletic groups, we first inferred a nuclear phylogeny across all individuals. We created locus-specific alignments using mafft, removing any loci sampled for <40% of individuals and any individuals sampled for <5% of loci. We then concatenated these alignments and used RAxML v8.2.11 (103) to infer a phylogeny across all individuals. To determine if individuals within nominal species show coherent IBD relationships, we calculated pairwise genetic divergence (d_{xy}) based on the concatenated alignment and determined geographic distances among individuals. For each nominal species, we plotted the monophyletic group spanning all the samples identified to that species, pairwise levels of genetic divergence (d_{xy}) by pairwise geographic distance, and spatial distribution of samples. Using this approach, we were able to identify samples that had been misidentified in the field and cases of likely cryptic speciation (Fig. S6). For a few species complexes that require major taxonomic revision (e.g., the snake taxon *Bothrops moojeni*), this approach provided a provisional taxonomy. In general, we took a conservative approach and only revised taxa in which monophyletic groups included individuals from other nominal species or where a taxon consisted of two or more deeply divergent lineages.

Measures of divergence

After defining provisional species limits, we then called genetic variants in each OTU. To do so, per OTU, we first created a reference set of loci, which comprised the longest assembly per locus across all individuals. We then mapped individual trimmed reads to the reference using bwa v0.7.15 (104). We called variants across all individuals using samtools v1.5 (105), filtered

variants to retain only those with coverage >20× and quality >20, and used this variant set to recalibrate alignments using GATK v4.1.8 (106). We then called genotypes across variable and invariable sites using samtools, removing genotypes with coverage <10× and sites with quality <20.

For all pairwise individual comparisons within an OTU, we calculated three metrics of genetic divergence: $F_{ST\ (107)}$, nuclear $d_{xy\ (108)}$, and mtDNA d_{xy}. We then used these metrics to calculate four different rates of differentiation. First, we calculated our focal metric for genetic differentiation (β_{IBD}), which is the slope of inverse F_{ST} across the log of geographic distance (109). We determined if β_{IBD} shows evidence for phylogenetic signal using Pagel's lambda (λ; 110) as implemented in the R package phytools (111). Further, β_{IBD} is predicted to vary as a function of both population density and dispersal (48, 109). Accordingly, we tested if β_{IBD} is predictable based on a few organismal traits that are proxies for these two properties: average genetic diversity (π), geographic range area, body mass, and elongation index. We calculated average genetic diversity across synonymous sites with >10x coverage and geographic range area using squamate geographic ranges presented in (112). Body mass and elongation index were based on published morphological data (113). We calculated the elongation ratio as a ratio of length to diameter, treating each individual as a cylinder with length equivalent to its snout-vent length. Correlations between β_{IBD} and these four organismal traits were tested using phylogenetic generalized least squares (PGLS) as implemented in the R package nlme (114).

We calculated four additional measures that might better capture how individuals differentiate across space: magnitude of inverse F_{ST} at 1000 km (which captures the effects of baseline differentiation), slope of nuclear d_{xy} across geographic distance, slope of mtDNA d_{xy} across geographic distance, and slope of F_{ST} across environmental distance. To calculate environmental distance, we extracted climatic data at each individual's location using the CHELSA climatic layers (115) and the R raster package (116). We then summarized these climatic data using a scaled and centered principal component analysis and calculated environmental distance between individuals as the summed Euclidean distance across all principal component axes. Finally, we assessed whether correlations between divergence and distance matrices were significant using Mantel tests using the R package ade4 (117).

Comparison to ddRAD

Many of the loci targeted in the SqCL dataset are highly-conserved loci assumed to be under purifying selection (49). Using such loci can bias demographic inference (118). To test this possibility, we compared SqCL-based estimates of genetic divergence with those derived from double-digest restriction-site aided DNA sequencing (ddRAD). ddRAD loci are randomly selected across the genome and are assumed to be mostly evolving neutrally (119). We used 61 individuals from 10 species of Australian lizards and snakes for which we have previously collected and analyzed both ddRAD and SqCL data (28, 120). We then calculated and compared the correlation in estimates of genetic diversity (π), genetic differentiation (F_{ST}), and rate of genetic differentiation (β_{IBD}) across these two marker sets.

Phylogenetic Framework and Speciation Rate Estimation

We estimated "recent" speciation rates (e.g., tip rates) across squamates, which index the instantaneous rate of lineage splitting in the limit as time approaches the present day (121). Although tip rates can be biased by incomplete taxon sampling, they are much less susceptible to biases that affect diversification rates as measured over deeper timescales and/or at earlier points in a clade's history (122, 123). The distribution of branch lengths in a reconstructed phylogenetic tree is a function of rates of lineage splitting and extinction (121); the net diversification rate is the difference between the rate at which lineages give rise to new species and the rates at which lineages become extinct. Although tip rate metrics have been interpreted as estimates of net diversification rate (e.g., DR statistic), theory and simulations show that branch lengths near the tips of the tree are much more highly correlated with speciation rate than with net diversification rate (5).

We estimated tip speciation rates across the phylogeny first published in Tonini et al. 2016 (50). This phylogeny is based on genetic data for 5415 of 9754 squamate taxa. Tonini et al. generated a pseudo-posterior distribution of 10,000 timetrees by adding the remaining 4339 taxa that lacked genetic data using a stochastic birth-death polytomy resolver (124); phylogenies constructed in this fashion yield conservative inferences about tip speciation rates and are more accurate than approaches that rely on analytical "sampling fraction" corrections (125).

For each of 100 trees sampled randomly from the full Tonini et al. distribution we inferred speciation rates across each tree sample using both semi-parametric and model-based approaches and then took the average tip rates across all trees to use in downstream analyses. For a semi-parametric estimate, we used the inverse-splits statistic (DR, λ_{DR}; 6, 126). The DR statistic is a weighted mean of inverse branch lengths along the path connecting a given tip to the root of the tree, with recent branches afforded proportionately more weight. Although the metric weights recent splitting events more, it estimates tip speciation rates with reasonable accuracy, even in the presence of extensive stochastic variation in branch lengths (5). For a model-based approach, we used CLaDS (51) to estimate speciation rates (λ_{CLaDS}). CLaDS applies a Bayesian approach to inferring speciation rates along a phylogeny and assumes that rates change after every speciation event. Both these approaches to measuring speciation rates appear capable of capturing fine-scale variation in speciation rates.

Our taxonomic modifications identified a few OTUs that do not occur in the tree. For these OTUs, in all comparative analyses, we mapped them to a closely-related congeneric taxon.

Test of Genetic Differentiation and Diversification

To test if rates of genetic differentiation (β_{IBD}) positively correlate with rates of speciation, we used two complementary approaches. First, we tested the correlation between β_{IBD} and speciation rate using a phylogenetic generalized least squares (PGLS) approach, implemented in the R package nlme (114). Here, we first took the natural log of both β_{IBD} and speciation rate and inferred the phylogenetic signal of the residuals under a Brownian motion model. Second,

we used the simulation-based approach ES-SIM (127), which tests for correlations between speciation rates and continuous traits. In this approach, traits are simulated along the given phylogeny, following the same Brownian motion model parameters inferred from the empirical trait data. The true correlation is then compared to the null distribution of correlations between the simulated traits and speciation rate. Under a range of trait evolution scenarios, ES-SIM has greater power than PGLS. We repeated these two tests for both speciation rates inferred using λ_{DR} and λ_{ClaDS}. Because of the sparseness of our IBD trait data (59 species-level taxa from a clade with ~10,000 species), we cannot use formal state-dependent models of evolutionary diversification (128, 129).

We additionally conducted a series of robustness analyses to ensure that our results were not due to technical or methodological artifacts. First, we investigated the effects of sampling. We iteratively dropped OTUs sampled for a minimum number of individuals and then tested the correlation between IBD slopes and speciation rates using these filtered datasets. Further, for those OTUs where we sampled >5 individuals, we subsampled all possible combinations of four individuals and measured IBD slopes to determine the potential for error due to small sample sizes. Additionally, we conducted individual-based forward genetic simulations using SLiM v3.6 (130) to determine how sample size affects error in IBD estimation (see *Supplemental Information* for full details). Then, we checked for the effects of error in IBD slope measurement in three ways. First, we repeated our analysis after removing OTUs with non-significant IBD slopes. Second, we iteratively dropped OTUs based on minimum r^2 and then tested the correlation between IBD slopes and speciation rates using these filtered datasets. Finally, we fit a phylogenetic mixed model with the squared standard error of the estimated slope as a factor in the model (MCMCglmm; 131); IBD slopes with greater error are weighted less in the model. Next, we used alternate rates of differentiation — absolute inverse F_{ST} at 1000 km, rate of nuclear d_{xy} divergence across geographic distance, rate of mtDNA d_{xy} divergence across geographic distance, and rate of F_{ST} divergence across environmental distance, because these alternate estimates might better capture the differentiation of these OTUs across space (132). Then, we repeated our analyses using species designations as recognized by current taxonomy to ensure our results were not affected by our revised species delimitations. Finally, a modest number of individuals were sampled outside of the core Cerrado region (Fig. S20); these individuals might have experienced a different biogeographic history. We removed these individuals, recalculated IBD slopes, and again tested our hypothesis.

Finally, we tested the power of our approach to recover a significant correlation between β_{IBD} and speciation rate, should one exist. Here, we simulated traits on our phylogeny under a Brownian motion of model, with a known correlation to speciation rate. For a range of correlations from 0 to 1.0, we simulated 100 datasets and determined how often both PGLS and ES-SIM recovered a significant correlation. Then, to determine the potential error in estimates of correlation, we simulated 100,000 datasets, drawing correlations from a uniform distribution of -1 to 1. We accepted correlations that deviated by less than 0.01 from the observed value (r = 0.04). This yielded a distribution of actual correlations between β_{IBD} and speciation rate that could potentially have generated our observed correlation.

References

1. D. Schluter, M. W. Pennell, Speciation gradients and the distribution of biodiversity. *Nature* **546**, 48–55 (2017).

2. E. Mayr, *Animal species and evolution* (Harvard University Press;, 1963).

3. F. Rutschmann, Molecular dating of phylogenetic trees: A brief review of current methods that estimate divergence times. *Divers. Distrib.* **12**, 35–48 (2006).

4. M. dos Reis, P. C. J. Donoghue, Z. Yang, Bayesian molecular clock dating of species divergences in the genomics era. *Nat. Rev. Genet.* **17**, 71–80 (2016).

5. P. O. Title, D. L. Rabosky, Tip rates, phylogenies and diversification: What are we estimating, and how good are the estimates? *Methods Ecol. Evol.* **10**, 821–834 (2019).

6. W. Jetz, G. H. Thomas, J. B. Joy, K. Hartmann, A. O. Mooers, The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).

7. D. L. Rabosky, *et al.*, An inverse latitudinal gradient in speciation rate for marine fishes. *Nature* **559**, 392–395 (2018).

8. O. Maliet, F. Hartig, H. Morlon, A model with many small shifts for estimating species-specific diversification rates. *Nat Ecol Evol* 3, 1086–1092 (2019).

9. E. P. Economo, N. Narula, N. R. Friedman, M. D. Weiser, B. Guénard, Macroecology and macroevolution of the latitudinal diversity gradient in ants. *Nat. Commun.* 9, 1778 (2018).

10. T. Varga, *et al.*, Megaphylogeny resolves global patterns of mushroom evolution. *Nat Ecol Evol* 3, 668–678 (2019).

11. C. R. Cooney, G. H. Thomas, Heterogeneous relationships between rates of speciation and body size evolution across vertebrate clades. *Nat Ecol Evol* 5, 101–110 (2021).

12. N. S. Upham, J. A. Esselstyn, W. Jetz, Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biol.* **17**, e3000494 (2019).

13. W. D. Allmon, “A causal analysis of stages in allopatric speciation” in *Oxford Surveys in Evolutionary Biology*, D. Futuyma, J. Antonovics, Eds. (Oxford University Press, 1992), pp. 219–257.

14. D. Schluter, Speciation, Ecological Opportunity, and Latitude (American Society of Naturalists Address). *Am. Nat.* **187**, 1–18 (2016).

15. D. J. Futuyma, G. C. Mayer, Non-Allopatric Speciation in Animals. *Syst. Biol.* **29**, 254–271 (1980).

16. A. L. Pigot, J. A. Tobias, Dispersal and the transition to sympatry in vertebrates. *Proc. Biol. Sci.* **282**, 20141929 (2015).

17. A. B. Phillimore, *et al.*, Sympatric speciation in birds is rare: insights from range data and simulations. *Am. Nat.* **171**, 646–657 (2008).

18. D. L. Rabosky, Reproductive isolation and the causes of speciation rate variation in nature. *Biol. J. Linn. Soc. Lond.* **118**, 13–25 (2016).

19. M. Dynesius, R. Jansson, Persistence of within-species lineages: a neglected control of speciation rates. *Evolution* **68**, 923–934 (2014).

20. A. J. Arnold, K. Fristrup, The theory of evolution by natural selection: a hierarchical expansion. *Paleobiology* 8, 113–129 (1982).

21. Y. Kisel, T. G. Barraclough, Speciation has a spatial scale that depends on levels of gene flow. *Am. Nat.* **175**, 316–334 (2010).

22. T. A. Hansen, Modes of larval development and rates of speciation in early tertiary neogastropods. *Science* **220**, 501–502 (1983).

23. L. P. Lagomarsino, F. L. Condamine, A. Antonelli, A. Mulch, C. C. Davis, The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytol.* **210**, 1430–1442 (2016).

24. M. G. Harvey, *et al.*, The evolution of a tropical biodiversity hotspot. *Science* **370**, 1343–1348 (2020).

25. B. T. Smith, *et al.*, The drivers of tropical speciation. *Nature* **515**, 406–409 (2014).

26. M. G. Harvey, S. Singhal, D. L. Rabosky, Beyond reproductive isolation: demographic controls on the speciation process. *Annu. Rev. Ecol. Evol. Syst.* **50**, 75–95 (2019).

27. M. G. Harvey, *et al.*, Positive association between population genetic differentiation and speciation rates in New World birds. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 6328–6333 (2017).

28. S. Singhal, *et al.*, Does population structure predict the rate of speciation? a comparative test across Australia's most diverse vertebrate radiation. *Am. Nat.* **192**, 432–447 (2018).

29. Y. Kisel, *et al.*, Testing the link between population genetic differentiation and clade diversification in Costa Rican orchids. *Evolution* **66**, 3035–3052 (2012).

30. T. M. Panhuis, R. Butlin, M. Zuk, T. Tregenza, Sexual selection and speciation. *Trends Ecol. Evol.* **16**, 364–371 (2001).

31. J. A. Coyne, H. A. Orr, *Speciation* (Sinauer Associates, 2004).

32. J. D. Kennedy, *et al.*, The influence of wing morphology upon the dispersal, geographical distributions and diversification of the Corvides (Aves; Passeriformes). *Proc. Biol. Sci.* **283** (2016).

33. C. Sheard, *et al.*, Ecological drivers of global gradients in avian dispersal inferred from wing morphology. *Nat. Commun.* **11**, 2463 (2020).

34. J. G. Cally, D. Stuart-Fox, L. Holman, J. Dale, I. Medina, Male-biased sexual selection, but not sexual dichromatism, predicts speciation in birds. *Evolution* **75**, 931–944 (2021).

35. D. M. Portik, *et al.*, Sexual Dichromatism Drives Diversification within a Major Radiation of African Amphibians. *Syst. Biol.* **68**, 859–875 (2019).

36. K. Kraaijeveld, F. J. L. Kraaijeveld-Smit, M. E. Maan, Sexual selection and speciation: the comparative evidence revisited. *Biol. Rev. Camb. Philos. Soc.* **86**, 367–377 (2011).

37. M. R. Foisy, L. P. Albert, D. W. W. Hughes, M. G. Weber, Do latex and resin canals spur plant diversification? Re-examining a classic example of escape and radiate coevolution. *J. Ecol.* **107**, 1606–1619 (2019).

38. P. J. Krug, *et al.*, Species selection favors dispersive life histories in sea slugs, but higher per-offspring investment drives shifts to short-lived larvae. *Syst. Biol.* **64**, 983–999 (2015).

39. J. C. Avise, *Phylogeography: The History and Formation of Species* (Harvard University Press, 2000).

40. G. M. Hewitt, Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Mol. Ecol.* **10**, 537–549 (2001).

41. J. Clausen, D. D. Keck, W. M. Hiesey, Regional differentiation in plant species. *Am. Nat.* **75**, 231–250 (1941).

42. L. R. Dice, Ecologic and genetic variability within species of peromyscus. *Am. Nat.* **74**, 212–221 (1940).

43. M. F. Smith, J. L. Patton, Subspecies of pocket gophers: causal bases for geographic differentiation in *Thomomys bottae*. *Syst. Biol.* **37**, 163–178 (1988).

44. S. M. Stanley, Population size, extinction, and speciation: the fission effect in Neogene bivalvia. *Paleobiology* **12**, 89–110 (1986).

45. J. M. C. D. Silva, J. M. Bates, Biogeographic patterns and conservation in the South American Cerrado: a tropical savanna hotspot. *Bioscience* **52**, 225–234 (2002).

46. G. C. Costa, C. Nogueira, R. B. Machado, G. R. Colli, Squamate richness in the Brazilian Cerrado and its environmental-climatic associations. *Divers. Distrib.* **13**, 714–724 (2007).

47. C. Nogueira, S. Ribeiro, G. C. Costa, G. R. Colli, Vicariance and endemism in a Neotropical savanna hotspot: distribution patterns of Cerrado squamate reptiles. *J. Biogeogr.* **38**, 1907–1922 (2011).

48. M. Slatkin, Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**, 264–279 (1993).

49. S. Katzman, *et al.*, Human genome ultraconserved elements are ultraselected. *Science* **317**, 915 (2007).

50. J. F. R. Tonini, K. H. Beard, R. B. Ferreira, W. Jetz, R. A. Pyron, Fully-sampled phylogenies of squamates reveal evolutionary patterns in threat status. *Biol. Conserv.* **204**, 23–31 (2016).

51. O. Maliet, H. Morlon, Fast and accurate estimation of species-specific diversification rates using data augmentation. *bioRxiv*, 2020.11.03.365155 (2020).

52. I. Medina, G. M. Cooke, T. J. Ord, Walk, swim or fly? Locomotor mode predicts genetic differentiation in vertebrates. *Ecol. Lett.* **21**, 638–645 (2018).

53. J. A. Endler, *Geographic variation, speciation, and clines* (Princeton University Press, 1977).

54. M. Kirkpatrick, V. Ravigné, Speciation by natural and sexual selection: models and experiments. *Am. Nat.* **159 Suppl 3**, S22–35 (2002).

55. D. Jablonski, Species selection: theory and data. *Annu. Rev. Ecol. Evol. Syst.* **39**, 501–524 (2008).

56. C. L. Salisbury, N. Seddon, C. R. Cooney, J. A. Tobias, The latitudinal gradient in dispersal constraints: ecological specialisation drives diversification in tropical birds. *Ecol. Lett.* **15**, 847–855 (2012).

57. J. Belliure, G. Sorci, A. P. Moller, J. Clobert, Dispersal distances predict subspecies richness in birds. *J. Evol. Biol.* **13**, 480–487 (2000).

58. G. L. Stebbins, Why are there so many species of flowering plants? *Bioscience* **31**, 573–577 (1981).

59. D. Jablonski, Larval ecology and macroevolution in marine invertebrates. *Bull. Mar. Sci.* **39**, 565–587 (1986).

60. J. Igea, E. F. Miller, A. S. T. Papadopulos, A. J. Tanentzap, Seed size and its rate of evolution correlate with species diversification across angiosperms. *PLoS Biol.* **15**, e2002792 (2017).

61. B. Ashby, A. K. Shaw, H. Kokko, An inordinate fondness for species with intermediate dispersal abilities. *Oikos* **129**, 311–319 (2020).

62. C. Riginos, Y. M. Buckley, S. P. Blomberg, E. A. Treml, Dispersal capacity predicts both population genetic structure and species richness in reef fishes. *Am. Nat.* **184**, 52–64 (2014).

63. G. S. Bradburd, G. M. Coop, P. L. Ralph, Inferring continuous and discrete population genetic structure across space. *Genetics* **210**, 33–52 (2018).

64. S. Manel, M. K. Schwartz, G. Luikart, P. Taberlet, Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol. Evol.* **18**, 189–197 (2003).

65. I. J. Wang, G. S. Bradburd, Isolation by environment. *Mol. Ecol.* **23**, 5649–5662 (2014).

66. L. van Holstein, R. A. Foley, Terrestrial habitats decouple the relationship between species and subspecies diversification in mammals. *Proc. Biol. Sci.* **287**, 20192702 (2020).

67. D. G. Haskell, A. Adhikari, Darwin's manufactory hypothesis is confirmed and predicts the extinction risk of extant birds. *PLoS One* 4, e5460 (2009).

68. M. N. Puttick, J. Clark, P. C. J. Donoghue, Size is not everything: rates of genome size evolution, not C-value, correlate with speciation in angiosperms. *Proc. Biol. Sci.* **282**, 20152289 (2015).

69. D. L. Rabosky, D. R. Matute, Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in *Drosophila* and birds. *Proc. Natl. Acad. Sci. U.*

649 *S. A.* **110**, 15354–15359 (2013).

650 70. E. B. Rosenblum, *et al.*, Goldilocks Meets Santa Rosalia: An Ephemeral Speciation Model
651 Explains Patterns of Diversification Across Time Scales. *Evol. Biol.* **39**, 255–261 (2012).

652 71. J. P. McEntee, J. A. Tobias, C. Sheard, J. G. Burleigh, Tempo and timing of ecological trait
653 divergence in bird speciation. *Nat Ecol Evol* 2, 1120–1127 (2018).

654 72. T. Price, *Speciation in birds* (W. H. Freeman, 2008).

655 73. D. L. Rabosky, Diversity-dependence, ecological speciation, and the role of competition in
656 macroevolution. *Annu. Rev. Ecol. Evol. Syst.* **44**, 481–502 (2013).

657 74. A. Camargo, B. Sinervo, J. W. Sites Jr, Lizards as model organisms for linking
658 phylogeographic and speciation studies. *Mol. Ecol.* **19**, 3250–3270 (2010).

659 75. R. C. Bell, *et al.*, Patterns of persistence and isolation indicate resilience to climate change
660 in montane rainforest lizards. *Mol. Ecol.* **19**, 2531–2544 (2010).

661 76. E. A. Myers, A. D. McKelvy, F. T. Burbrink, Biogeographic barriers, Pleistocene refugia, and
662 climatic gradients in the southeastern Nearctic drive diversification in cornsnakes
663 (*Pantherophis guttatus* complex). *Mol. Ecol.* **29**, 797–811 (2020).

664 77. A. D. Leaché, S. C. Crews, M. J. Hickerson, Two waves of diversification in mammals and
665 reptiles of Baja California revealed by hierarchical Bayesian analysis. *Biol. Lett.* 3, 646–650
666 (2007).

667 78. E. Mayr, *Systematics and the Origin of Species, from the Viewpoint of a Zoologist* (Harvard
668 University Press, 1942).

669 79. J. Ramsey, H. D. Bradshaw Jr, D. W. Schemske, Components of reproductive isolation
670 between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). *Evolution* **57**,
671 1520–1534 (2003).

672 80. K. M. Kay, Reproductive isolation between two closely related hummingbird-pollinated
673 neotropical gingers. *Evolution* **60**, 538–552 (2006).

674 81. T. Dobzhansky, *Genetics and the Origin of Species* (Columbia University Press, 1982).

675 82. K. L. Ostevik, R. L. Andrew, S. P. Otto, L. H. Rieseberg, Multiple reproductive barriers
676 separate recently diverged sunflower ecotypes. *Evolution* **70**, 2322–2335 (2016).

677 83. C. S. Price, C. H. Kim, C. J. Gronlund, J. A. Coyne, Cryptic reproductive isolation in the
678 *Drosophila simulans* species complex. *Evolution* **55**, 81–92 (2001).

679 84. D. R. Matute, B. S. Cooper, Comparative studies on speciation: 30 years since Coyne and
680 Orr. *Evolution* (2021) https:/doi.org/10.1111/evo.14181.

681 85. C. C. Nogueira, *et al.*, Atlas of Brazilian snakes: verified point-locality maps to mitigate the
682 Wallacean shortfall in a megadiverse snake fauna. *South American Journal of Herpetology*
683 **14**, 1–274 (2019).

684 86. C. Nogueira, G. R. Colli, M. Martins, Local richness and distribution of the lizard fauna in

natural habitat mosaics of the Brazilian Cerrado. *Austral Ecol.* **34**, 83–96 (2009).

87. A. G. Nazareno, J. B. Bemmels, C. W. Dick, L. G. Lohmann, Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Mol. Ecol. Resour.* **17**, 1136–1147 (2017).

88. A. R. Lemmon, S. A. Emme, E. M. Lemmon, Anchored hybrid enrichment for massively high-throughput phylogenomics. *Syst. Biol.* **61**, 727–744 (2012).

89. S. Singhal, M. Grundler, G. Colli, D. L. Rabosky, Squamate Conserved Loci (SqCL): A unified set of conserved loci for phylogenomics and population genetics of squamate reptiles. *Mol. Ecol. Resour.* **17**, e12–e24 (2017).

90. B. C. Faircloth, *et al.*, Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* **61**, 717–726 (2012).

91. S. Singhal, *et al.*, Congruence and conflict in the higher-level phylogenetics of squamate reptiles: an expanded phylogenomic perspective. *Syst. Biol.* **70**, 542–557 (2021).

92. S. M. Aljanabi, I. Martinez, Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* **25**, 4692–4693 (1997).

93. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

94. J. Zhang, K. Kobert, T. Flouri, A. Stamatakis, PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**, 614–620 (2014).

95. M. G. Grabherr, *et al.*, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).

96. W. J. Kent, BLAT—The BLAST-Like Alignment Tool. *Genome Res.* **12**, 656–664 (2002).

97. C. Hahn, L. Bachmann, B. Chevreux, Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res.* **41**, e129–e129 (2013).

98. G. S. C. Slater, E. Birney, Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 6, 31 (2005).

99. K. Katoh, G. Asimenos, H. Toh, Multiple alignment of DNA sequences with MAFFT. *Methods Mol. Biol.* **537**, 39–64 (2009).

100. S. Singhal, C. Moritz, Reproductive isolation between phylogeographic lineages scales with divergence. *Proc. Biol. Sci.* **280**, 20132246 (2013).

101. D. L. Rabosky, M. N. Hutchinson, S. C. Donnellan, A. L. Talaba, I. J. Lovette, Phylogenetic disassembly of species boundaries in a widespread group of Australian skinks (Scincidae: *Ctenotus*). *Mol. Phylogenet. Evol.* **77**, 71–82 (2014).

102. A. D. Leaché, M. K. Fujita, Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proc. Biol. Sci.* **277**, 3071–3077 (2010).

721 103. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of
722 large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

723 104. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
724 *arXiv [q-bio.GN]* (2013).

725 105. H. Li, *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,
726 2078–2079 (2009).

727 106. A. McKenna, *et al.*, The Genome Analysis Toolkit: a MapReduce framework for
728 analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).

729 107. D. Reich, K. Thangaraj, N. Patterson, A. L. Price, L. Singh, Reconstructing Indian
730 population history. *Nature* **461**, 489–494 (2009).

731 108. M. Nei, W. H. Li, Mathematical model for studying genetic variation in terms of restriction
732 endonucleases. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 5269–5273 (1979).

733 109. F. Rousset, Genetic differentiation and estimation of gene flow from F-statistics under
734 isolation by distance. *Genetics* **145**, 1219–1228 (1997).

735 110. M. Pagel, Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884
736 (1999).

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

739 112. U. Roll, *et al.*, The global distribution of tetrapods reveals a need for targeted reptile
740 conservation. *Nat Ecol Evol* 1, 1677–1682 (2017).

741 113. A. Feldman, N. Sabath, R. A. Pyron, I. Mayrose, S. Meiri, Body sizes and diversification
742 rates of lizards, snakes, amphisbaenians and the tuatara. *Glob. Ecol. Biogeogr.* **25**, 187–
743 197 (2016).

744 114. J. Pinheiro, nlme : Linear and nonlinear mixed effects models. *http://cran.r-*
745 *project.org/web/packages/nlme/* (2009) (June 25, 2020).

746 115. D. N. Karger, *et al.*, Climatologies at high resolution for the earth's land surface areas.
747 *Sci Data* 4, 170122 (2017).

748 116. R. J. Hijmans, *et al.*, Package "raster." *R package* **734** (2015).

749 117. S. Dray, A.-B. Dufour, Others, The ade4 package: implementing the duality diagram for
750 ecologists. *J. Stat. Softw.* **22**, 1–20 (2007).

751 118. G. B. Ewing, J. D. Jensen, The consequences of not accounting for background
752 selection in demographic inference. *Mol. Ecol.* **25**, 135–141 (2016).

753 119. K. R. Andrews, J. M. Good, M. R. Miller, G. Luikart, P. A. Hohenlohe, Harnessing the
754 power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* **17**, 81–92
755 (2016).

756 120. M. R. Grundler, S. Singhal, M. A. Cowan, D. L. Rabosky, Is genomic diversity a useful

proxy for census population size? Evidence from a species-rich community of desert lizards. *Mol. Ecol.* **28**, 1664–1674 (2019).

121. S. Nee, R. M. May, P. H. Harvey, The reconstructed evolutionary process. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **344**, 305–311 (1994).

122. S. Louca, M. W. Pennell, Extant timetrees are consistent with a myriad of diversification histories. *Nature* **580**, 502–505 (2020).

123. C. R. Marshall, Five palaeobiological laws needed to understand the evolution of the living biota. *Nat Ecol Evol* 1, 165 (2017).

124. G. H. Thomas, *et al.*, PASTIS: an R package to facilitate phylogenetic assembly with soft taxonomic inferences. *Methods Ecol. Evol.* 4, 1011–1017 (2013).

125. J. Chang, D. L. Rabosky, M. E. Alfaro, Estimating diversification rates on incompletely sampled phylogenies: theoretical concerns and practical solutions. *Syst. Biol.* **69**, 602–611 (2020).

126. D. W. Redding, A. Ø. Mooers, Incorporating evolutionary measures into conservation prioritization. *Conserv. Biol.* **20**, 1670–1678 (2006).

127. M. G. Harvey, D. L. Rabosky, Continuous traits and speciation rates: Alternatives to state-dependent diversification models. *Methods Ecol. Evol.* 9, 984–993 (2018).

128. W. P. Maddison, P. E. Midford, S. P. Otto, Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* **56**, 701–710 (2007).

129. J. M. Beaulieu, B. C. O'Meara, Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Syst. Biol.* **65**, 583–601 (2016).

130. B. C. Haller, P. W. Messer, SLiM 3: Forward Genetic Simulations Beyond the Wright–Fisher Model. *Mol. Biol. Evol.* **36**, 632–637 (2019).

131. J. D. Hadfield, MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* (2010).

132. B. Charlesworth, Measures of divergence between populations and the effect of forces that reduce variability. *Mol. Biol. Evol.* **15**, 538–543 (1998).

133. D. L. Rabosky, Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One* 9, e89543 (2014).

Data Availability

- All code used to analyze data and make figures is available at GitHub: https://github.com/singhal/brazil_IBD/
- Sequencing reads are available on NCBI Short Read Archive under BioProject PRJNA787529
- Variant Call Files per species-level taxa are available at DataDryad DOI: https://doi.org/10.5061/dryad.1zcrjdftd

Acknowledgements

For advice, we thank P. Moorjani; for technical and logistical support, we thank A. Devault, J. Enk, H. Huang, M. R. Marchan Rivadeneira, R. Murrell, G. Schneider, and the staff at University of Michigan ARC-TS. Funding was provided by NSF DEB 1754398 to DLR and SS; NSF DEB 1754425 to GCC; a Packard Fellowship to DLR; NSF DEB 1519732 to SS, RSCA and Norris Grant to SS. GRC thanks Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Apoio à Pesquisa do Distrito Federal (FAPDF) and the USAID's PEER program under cooperative agreement AID-OAA-A-11-00012 for financial support. For photos, we acknowledge R. Recoder and D. A. A. Souza, and for the data in Fig. 1 and S1, we thank C. Cooney.

Author Contributions

DLR and SS conceived and coordinated the study, DLR, SS, GCC, and GRC designed study, GRC, GCC, SS, and MRG collected data, SS and DLR analyzed data, SS and DLR drafted manuscript, SS, DLR, and IP drafted figures, and all authors edited and approved of the manuscript.

Competing Interests

The authors declare no competing interests.

Figure Legends

Figure 1: Speciation rate variation across major vertebrate clades. Speciation rates are estimated using the DR statistic (λ_{DR}; 5, 6); data from Cooney & Thomas (2020; 11). Speciation rate distributions are highly skewed and are thus presented on a log10 scale. Within each group, the slowest and fastest 2.5% of speciation rates differ anywhere from 14-fold to 46-fold. The sources of this variation remain largely unexplained. Speciation rates estimated using a model-based approach to speciation rate estimation (BAMM; 133) are presented in Fig. S1.

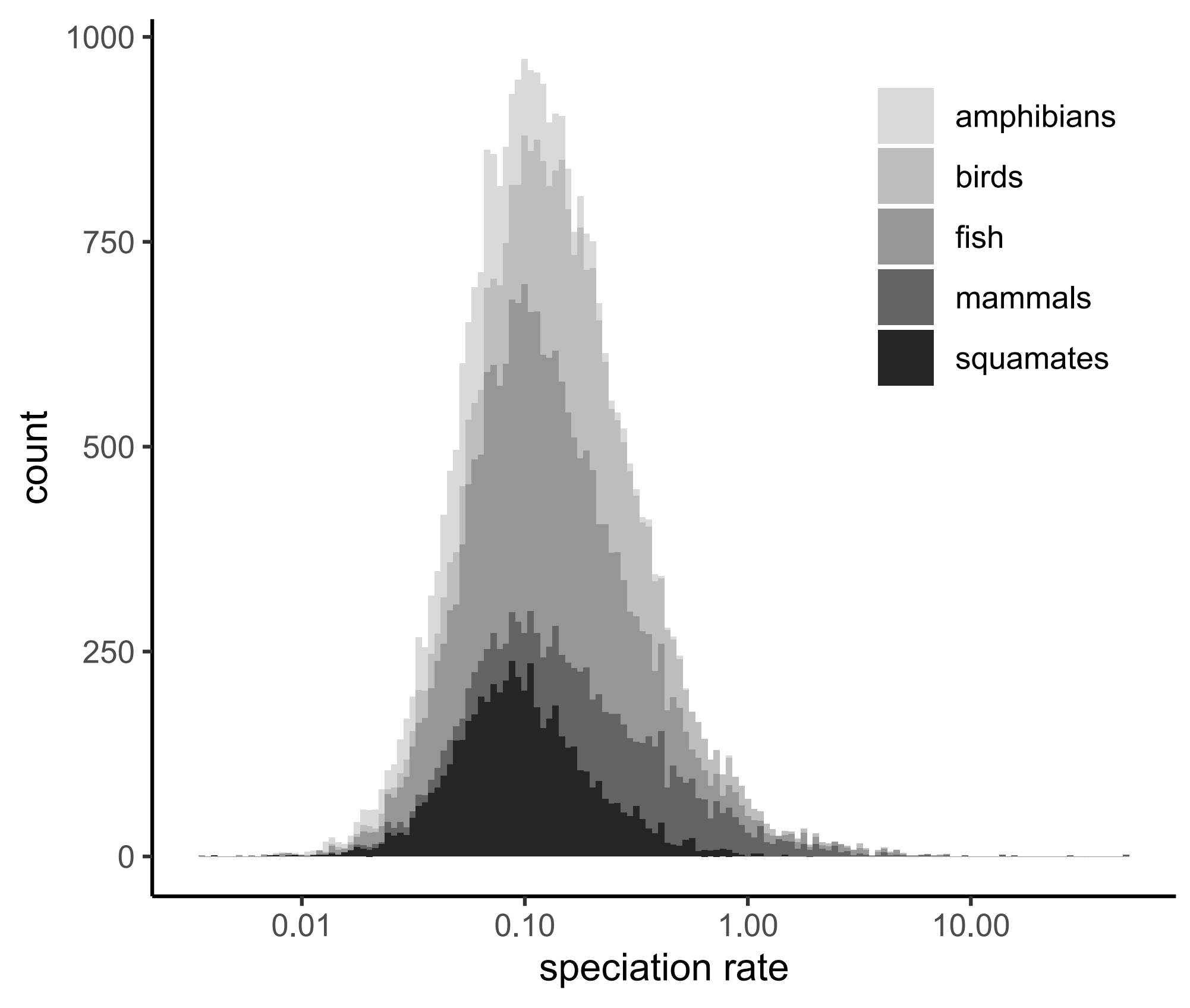
Figure 2: Speciation rates, geographic sampling, and genomic data coverage for focal lizard and snake taxa. (A) Speciation rate (λ_{ClaDS}) across all squamates for the Tonini et al. 2016 (50) phylogeny. Major squamate clades are labeled at nodes: Gekkota (G), Scincoidea (Sc), Lacertoidea (L), Iguania (I), and Serpentes (S). Phylogeny shown for a random 50% subsample of lineages to ease visualization. Focal taxa included in this study are labeled with circles at the tips, with fill color indicating tip-averaged speciation rate. A subset of tips are represented by photos of taxa; photos not to scale (photos by R. Recoder, GRC, and IP). Inset: rank-order plot with speciation rates of the focal taxa relative to the distribution of speciation rates of all squamate taxa. (B) Map of South America with sampling sites shown in white. (C) Concatenated phylogeny of 4796 loci and 375 individuals; colored tip groupings denote the 59 putative species-level taxa studied here. (D) Summary of genetic data for the 375 individuals included in this study. (E) Landscape images of the Cerrado, the grassland biome on which field sampling was centered (photos by D. A. A. Souza and GRC). This study's focal taxa span a wide phylogenetic distribution and encompass almost the full range of speciation rates seen in squamates.

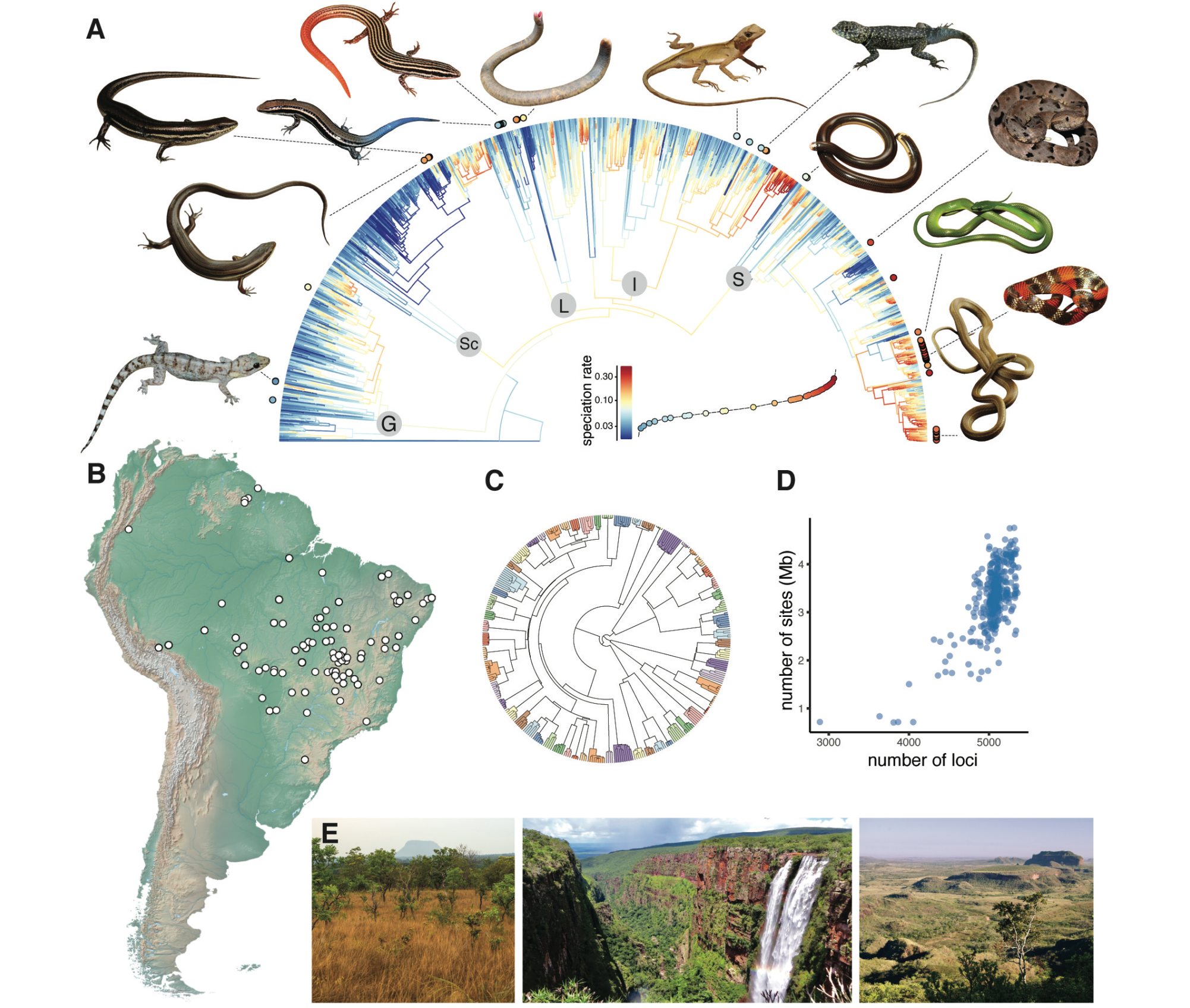
Figure 3: Illustrative isolation-by-distance slopes (β_{IBD}) generated for two species. These taxa exhibit strongly contrasting rates of population isolation: after controlling for geographic distance, the viper *Bothrops moojeni* (*n* = 10, top) has a low rate of divergence between populations, whereas the gymnophthalmid lizard *Micrablepharus atticolus* (*n* = 6, bottom) shows a high rate. Maps illustrate geographic ranges (see *Supplemental Information* for information on reconstruction) for each taxon along with sampling points. Scatter plots show how genetic differentiation (as measured by F_{ST}) accumulates across space. Figure S8 shows isolation-by-distance plots across all taxa.

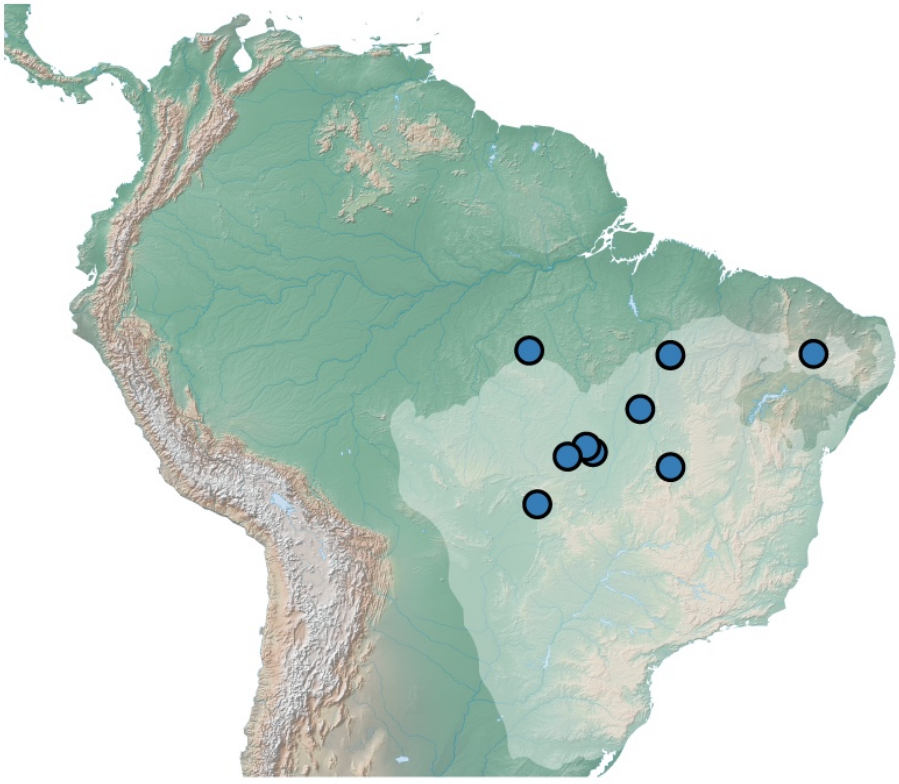
Figure 4: Correlation between population isolation (β_{IBD}) and speciation rates (λ_{ClaDS}) across squamate reptiles. (A) β_{IBD} exhibits phylogenetic signal, as might be expected if it tracks intrinsic organismal traits such as dispersal capacity. Snakes (panel A: node "S") are nested within squamates; all other lineages are traditionally known as "lizards". (B) β_{IBD} is not significantly correlated with speciation rate (PGLS $r = 0.04$, $p = 0.77$), and snakes and lizards show little overlap in the bivariate space defined by β_{IBD} and speciation rate.

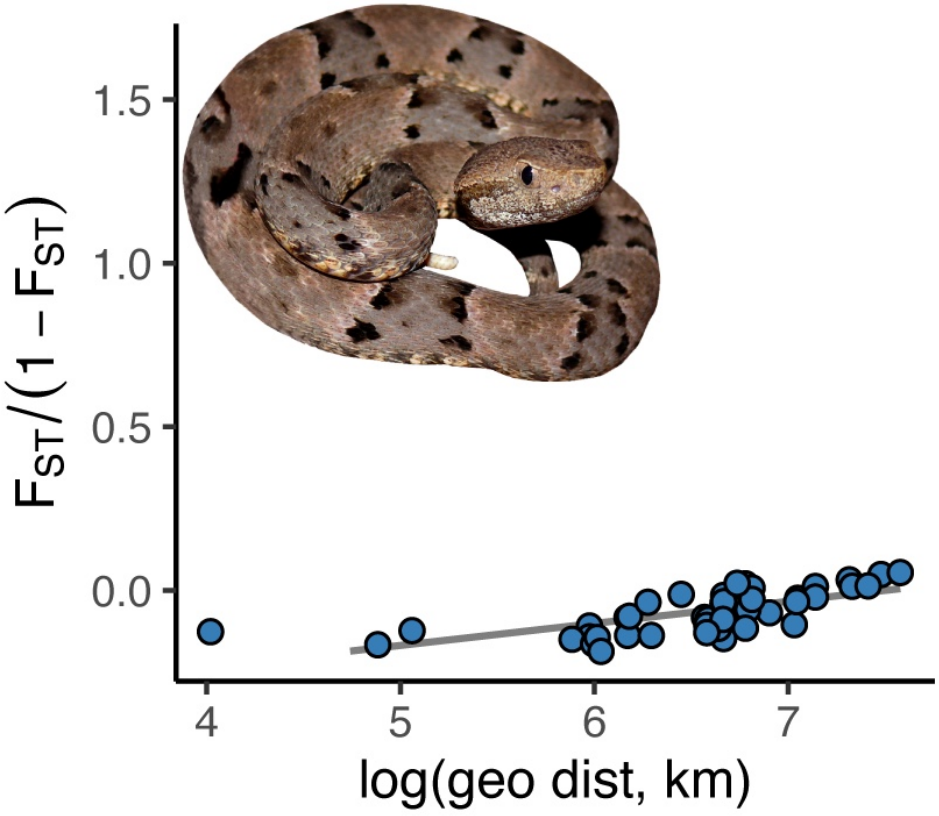
Figure 5: Speciation rate and population structure for taxa with lizard-like and snake-like body forms. More elongated animals (higher body elongation values) have lower rates of population isolation but most of this variation is captured by the evolutionary split between

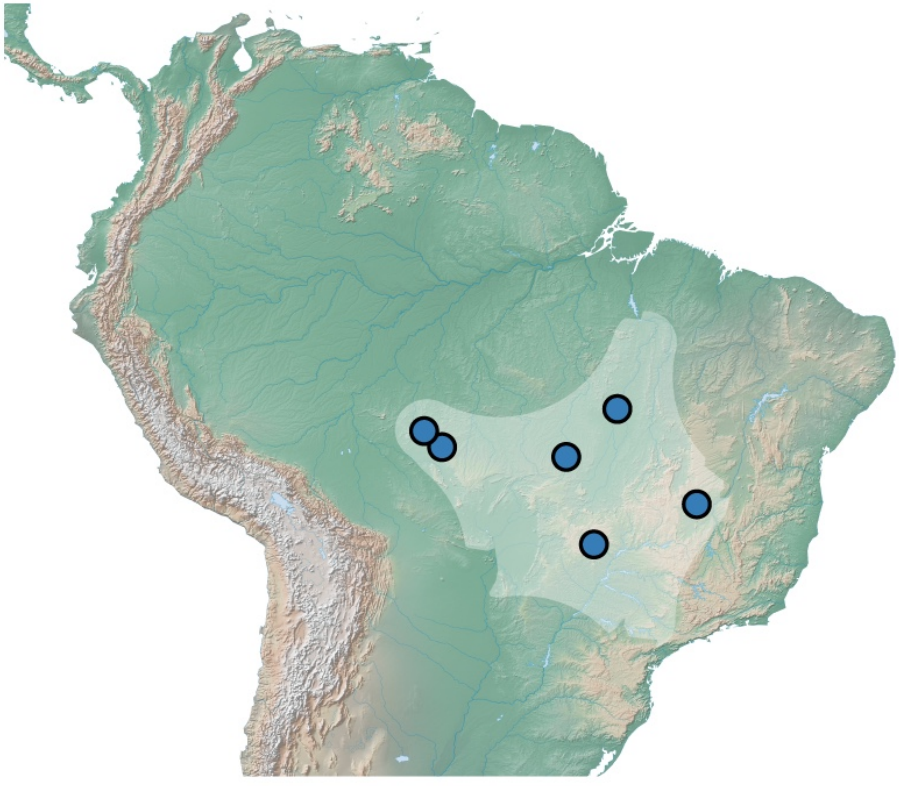
855 snakes and “lizards”. Body elongation was computed as the ratio of length to diameter, treating
856 each animal as an idealized cylinder. The focal lizards in this study (n = 24) exhibit higher levels
857 of population isolation than the focal snakes (n = 35). Despite exhibiting higher levels of
858 population isolation, lizards have overall lower speciation rates compared to snakes, contrary to
859 our central hypothesis.

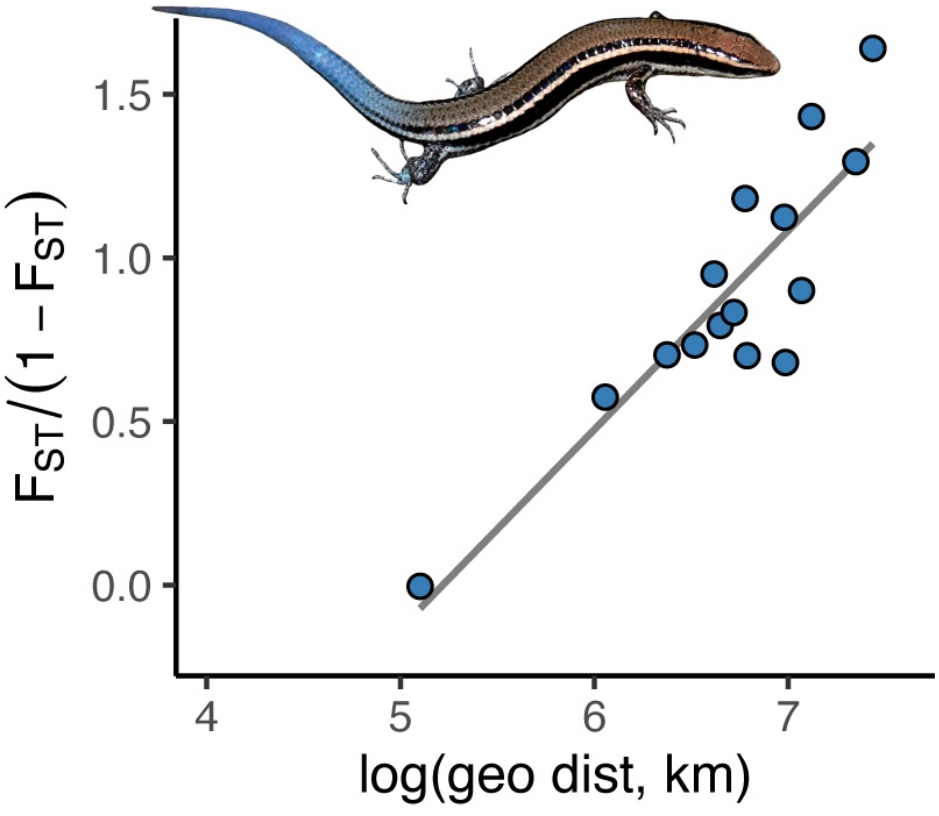

1000
750
500
250
0
count
0.01
0.10
1.00
10.00
speciation rate
amphibians
birds
fish
mammals
squamates

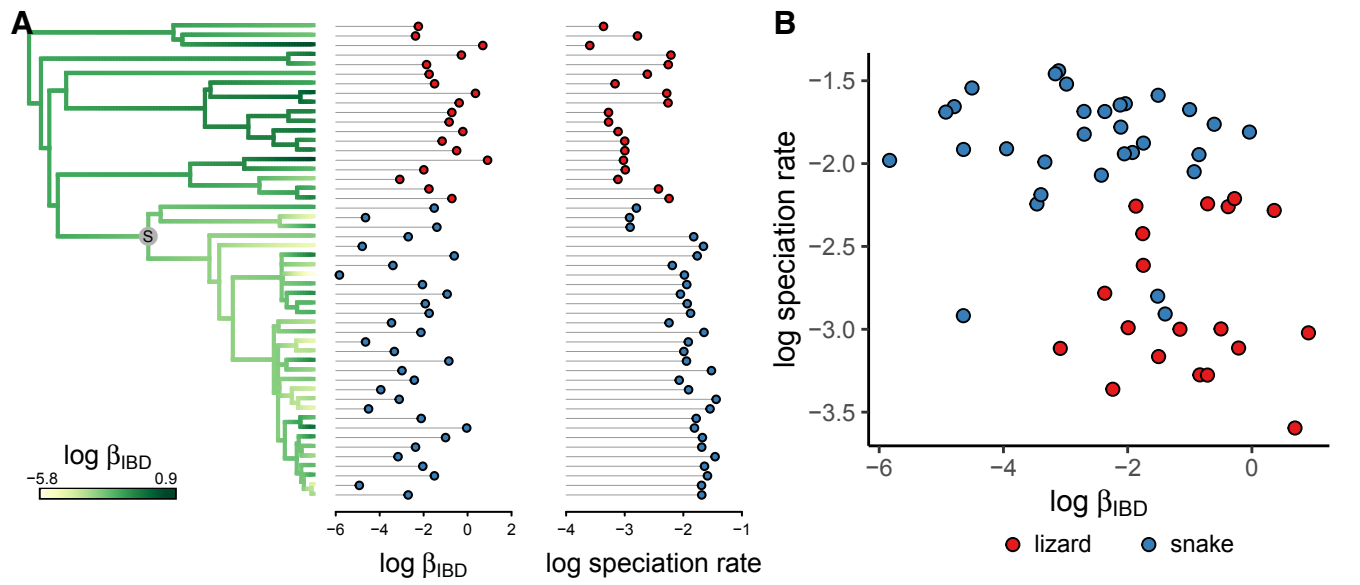

A
G
Sc
L
I
S
speciation rate
0.30
0.10
0.03
B
C
D
number of sites (Mb)
1
2
3
4
3000
4000
5000
number of loci
E

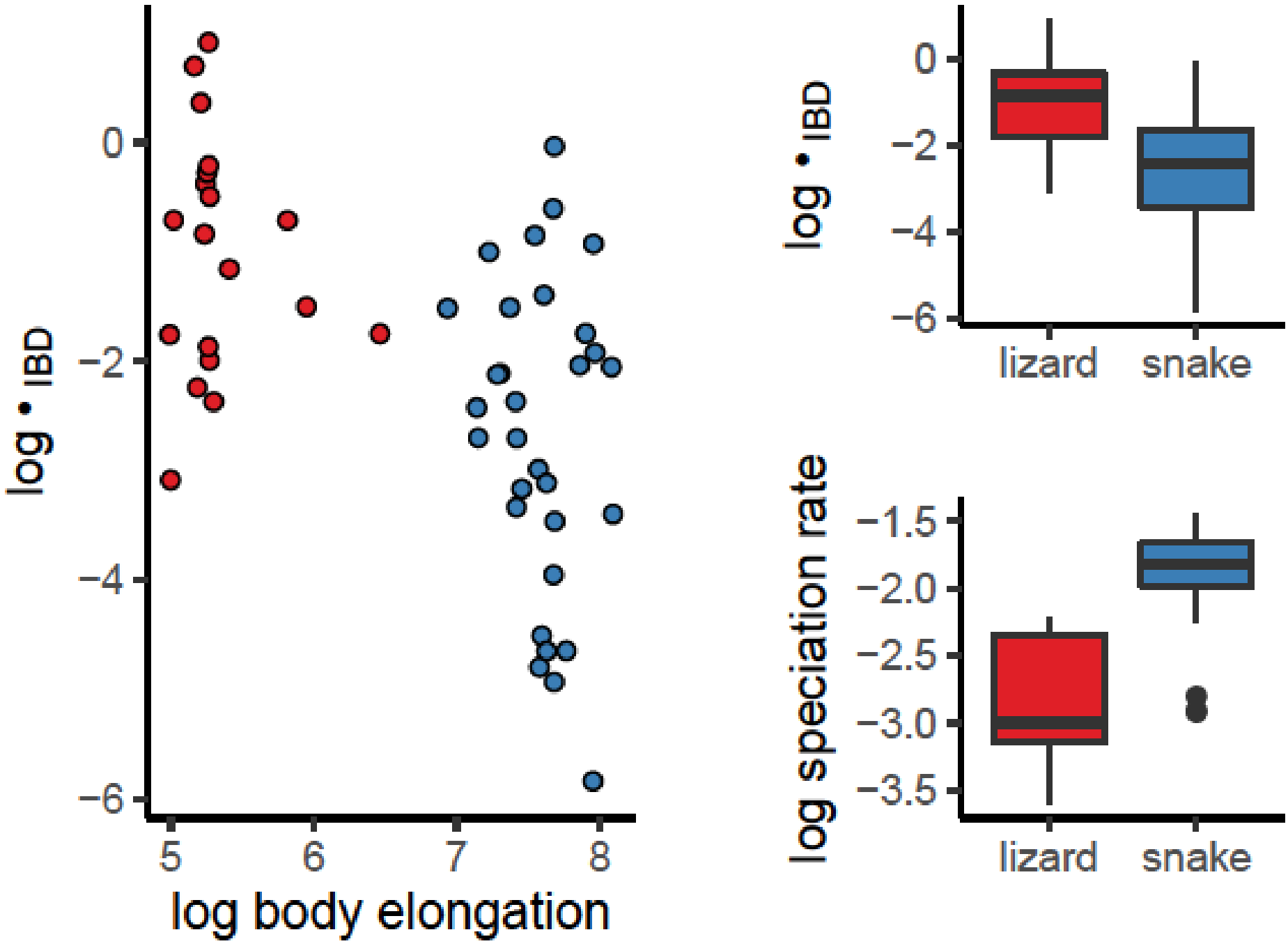



1.5
1.0
0.5
0.0
$F_{ST}/(1-F_{ST})$
4
5
6
7
log(geo dist, km)




1.5
1.0
0.5
0.0
$F_{ST}/(1-F_{ST})$
4
5
6
7
log(geo dist, km)


A
S
log β_{IBD}
−5.8
0.9
−6
−4
−2
0
2
log β_{IBD}
−4
−3
−2
−1
log speciation rate
B
log speciation rate
−1.5
−2.0
−2.5
−3.0
−3.5
−6
−4
−2
0
log β_{IBD}
lizard
snake


log • IBD
0
−2
−4
−6
5
6
7
8
log body elongation
log • IBD
0
−2
−4
−6
lizard
snake
log speciation rate
−1.5
−2.0
−2.5
−3.0
−3.5
lizard
snake